

Ecole Centrale des Arts & Manufactures

1^{ère} ANNÉE d'ETUDES

ANALYSE STATISTIQUE

M. GIRSCHIG

PROFESSEUR



1969

Tome I - ①

Introduction au cours d'analyse statistique et de calcul des probabilités

I. Statistique descriptive.

Le but de toute étude statistique est de décrire et de résumer les propriétés de collections d'individus, d'objets ou de faits classés en fonction d'un ou plusieurs des caractères qui leur sont communs.

Ces collections peuvent être complètement définies, comme l'est, à une époque donnée, la "population" d'une ville formée par l'ensemble de ses habitants. Par contre, la suite des résultats d'une partie indéfinie de "Pile ou Face" forme une collection qui ne peut être définie en extension, car une telle collection n'est jamais finie.

Il est habituel de désigner d'une manière très générale :

- par "individus" les éléments de la collection à étudier, que ces éléments soient ou non des êtres humains.
- par "population" la collection d'individus faisant l'objet d'une étude statistique.

Les moyens mis en œuvre pour l'étude d'une population sont :

- le recensement
- le classement, opération qui a pour but d'effectuer une partition de la population à étudier en classes d'équivalence contenant les individus qui possèdent un ou plusieurs caractères aux mêmes degrés d'intensité.

La description de la partition ainsi obtenue fait appel aux notions de fréquence et de distribution de ces fréquences en fonction des caractères de classement adoptés. On s'efforce, dans toute la mesure du possible, de représenter de telles distributions par des modèles mathématiques susceptibles de l'être avec une bonne approximation.

L'étude des distributions sera l'objet de la première partie de ce cours, consacrée à la statistique descriptive.

II. Événements à l'éatoire et probabilités.

Dans de nombreuses circonstances il nous est impossible d'effectuer une prévision rigoureuse des événements qui vont succéder à une situation donnée. Ces événements résultent, en effet, du jeu complexe d'un grand nombre de facteurs qui sont, pour nous, indéterminés et que, dans notre ignorance, nous assimilons au hasard.

Sous certaines conditions il est, cependant, possible de ramener l'incertitude à un niveau d'éventualités auxquelles nous pourront attribuer de plus ou moins grandes vraisemblances.

Dans ces circonstances, tantôt l'événements à l'éatoire de probabilités connues, le calcul des probabilités permet d'évaluer les chances de réalisation d'événements complexes qui sont la conséquence des premiers. Les modèles ainsi obtenus constitueront, de plus, des explications valables des phénomènes observés.

Historiquement le calcul des probabilités a été imaginé pour permettre l'étude des jeux de hasard. Son emploi s'est ensuite étendu à de nombreux domaines parmi lesquels on retiendra :

- les applications à la physique mathématique
(étude des populations de particules et de leurs interactions, étude des fluctuations, théorie cinétique des gaz...)
 - la théorie des assurances.
 - l'étude des processus aléatoires (lire d'attente, chaînes de Markoff...)
 - les études de fiabilité.
 - les méthodes de gestion des stocks.
- - - - -

L'étude du calcul des probabilités et de quelquesunes de ses applications sera l'objet de la seconde partie de ce cours.

III . Théorie de l'échantillonnage et tests statistiques.

La théorie de l'échantillonnage procède d'une démarche de pensée qui est, en quelque sorte, l'inverse de celle relative au calcul des probabilités. La théorie de l'échantillonnage se propose, en effet :

- soit d'estimer, avec une marge de risque spécifiée à l'avance, les caractères statistiques d'une population à partir du seul examen d'un "échantillon" d'effectif plus réduit, prélevé au hasard dans la population à étudier (calcul des "estimations par intervalle" ou "limites de confiance" des paramètres de la population échantillonnée).
- soit, d'après l'examen des résultats d'une série d'observations, d'accepter ou de rejeter une hypothèse considérée a priori, comme particulièrement plausible (théorie des tests statistiques).

Les applications de la théorie de l'échantillonnage et des tests statistiques sont fort nombreuses. A titre d'exemple on peut citer :

- l'estimation des caractéristiques moyennes et des tolérances d'un produit industriel
 - la surveillance de la régularité de l'fonctionnement d'un procédé de fabrication
 - les méthodes de réception sur prélevements
 - la vérification d'hypothèses scientifiques ou économiques.
- - - - -

IV Méthodes d'analyse des observations.

Les statisticiens ont élaborés des méthodes de plus en plus précises d'analyse des observations. Ces méthodes permettent de rechercher les facteurs qualitatifs ou quantitatifs qui exercent une influence non négligeable sur la qualité ou la variabilité d'un produit ou d'un résultat. Appliquées à l'analyse temporelle du déroulement d'un processus aléatoire, ces méthodes permettent d'établir une préiction d'un futur proche à partir d'observations se rattachant à un passé récent. Parmi les applications industrielles des méthodes d'analyse statistique on peut citer :

- l'identification des facteurs de qualité d'une fabrication et la mesure de leurs effets (analyse de la variance, analyse factorielle)
- la recherche des relations de dépendance entre les résultats d'un processus industriel et les variables qui peuvent en influencer le rendement (analyse des régressions, analyse de la covariance)
- l'élaboration de plans d'expérience permettant, aux moindres frais pour un degré de confiance fixé à l'avance, d'obtenir des résultats susceptibles de conduire à un modèle vraisemblable des phénomènes étudiés.
- l'ajustement, sur des observations antérieures, de séries temporelles suivant certains modèles à des processus stochastiques. En particulier, l'établissement de modèles relatifs à l'évolution de phénomènes économiques.

L'étude de quelquesunes des méthodes classiques d'analyse statistique des observations sera l'objet de la dernière partie de ce cours.

1^{re} Partie. Statistique descriptive

ch I. Distribution à un seul caractère.

Individus et populations.

La statistique descriptive a pour but de dénombrer un ensemble d'etui, d'objets ou de faits et de les classer selon un ou plusieurs de leurs caractères communs. Il est d'usage de distinguer sous le terme général de "Population" l'ensemble des éléments rassemblés en vue d'en effectuer un classement logique. La population à classer est formée d'éléments distincts que nous appellerons "individus".

Les individus d'une même population présentent en commun et à des degrés variables un certain nombre de caractères qui serviront de base au classement à effectuer. Dans cette première leçon, nous supposerons que la population étudiée est constituée par un ensemble d'individus possédant en commun un caractère quantitatif x , dont l'intensité peut varier d'un individu à l'autre.

I-1 Distribution d'une variable discrète

I-2

Ex. Classement de $N=100$ familles en fonction du nombre x_i de leurs enfants

- x_i : nombre d'enfants
- n_i : nombre de familles possédant x_i enfants
- f_i : fréquence relative des familles possédant x_i enfants
- F_i : fréquence cumulée des familles possédant x_i enfants
- $N = \sum n_i$: effectif global de la population étudiée

x_i	0	1	2	3	4	5	Totaux
n_i	5	20	40	20	10	5	$N = 100$
f_i	0,05	0,20	0,40	0,20	0,10	0,05	$\sum f_i = 1$
F_i	0,05	0,25	0,65	0,85	0,95	1	

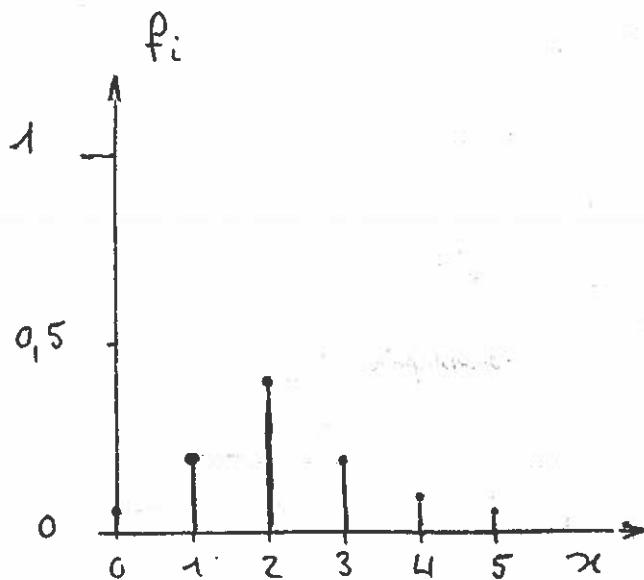


Fig I-1 Diagramme de "Distribution"

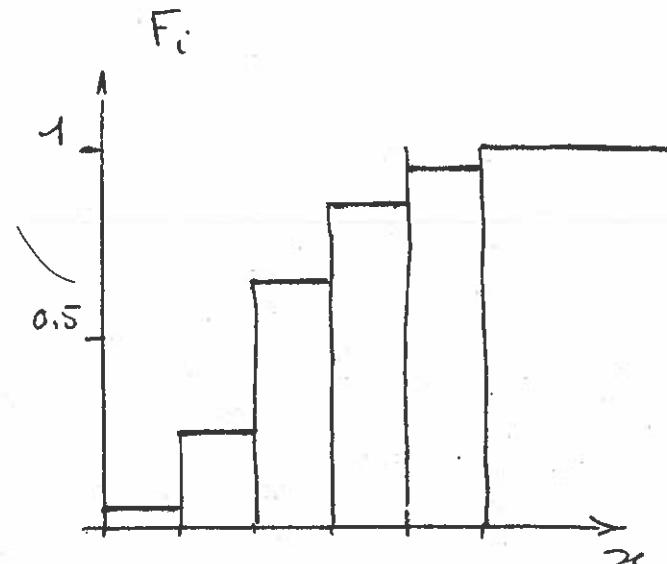


Fig I-2 Diagramme de "Répartition"

I - 2. Distribution d'une variable continue

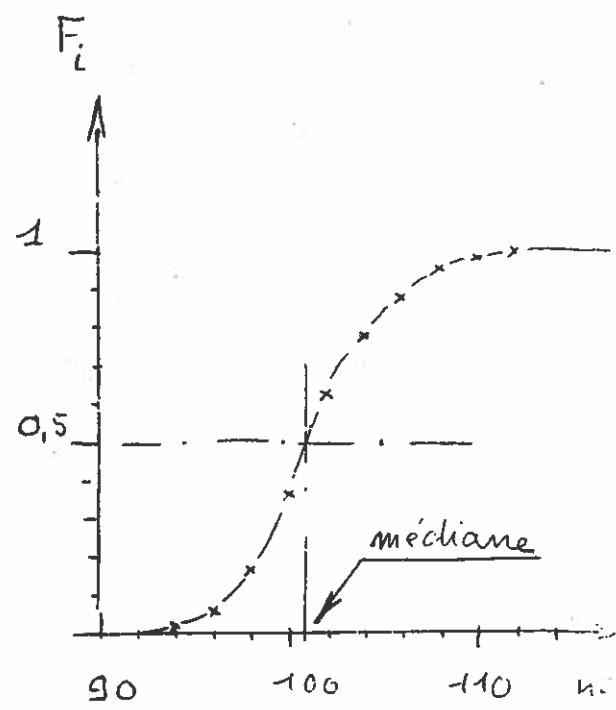
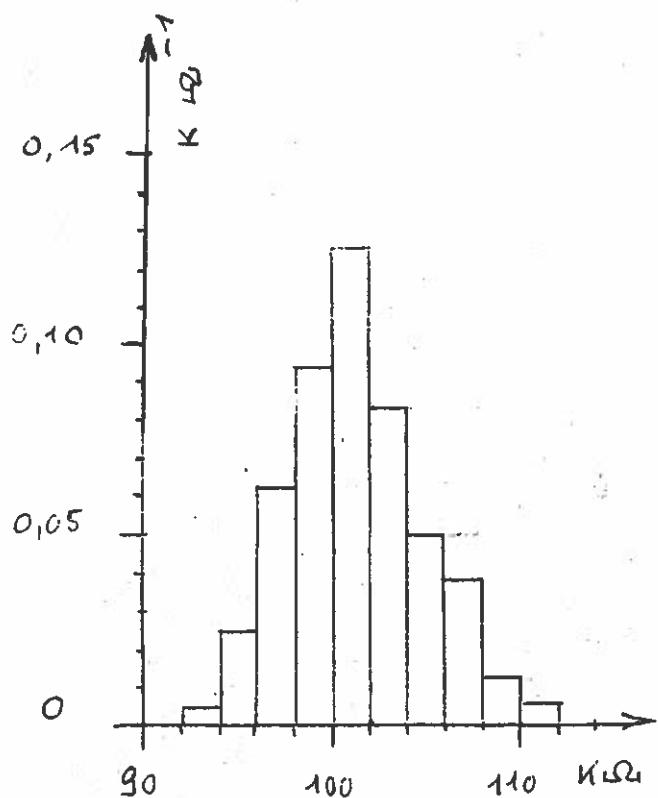
EX. Classement de $N = 400$ résistances électriques en fonction de leur résistance ohmique réelle.

Pour étudier la distribution d'une fabrication de résistances d'une valeur nominale de 100 K_Ω, on a contrôlé 400 d'entre elles à l'aide d'un ohmmètre de précision. On a reporté les résultats de ce contrôle en les groupant en classes d'une étendue arbitraire de 2 K_Ω, soit environ $6 \frac{1}{10}$ de la dispersion totale de l'échantillon contrôlé. Ce regroupement en classes permet de simplifier notablement la présentation des résultats, sans qu'il en résulte une perte notable d'information.

Pour une variable continue, le diagramme de la distribution sera présenté sous forme d'un "Histogramme" établi en reportant à l'abscisse de chaque classe un rectangle dont la surface est proportionnelle à la fréquence des résistances de la classe correspondante. Ce mode de représentation présente l'avantage d'être relativement indépendant des subdivisions adoptées, comme on peut, par exemple, s'en rendre compte en groupant en une seule classe deux classes voisines.

En portant à l'abscisse de la limite supérieure de chaque classe les fréquences cumulées correspondantes on tracera sous forme d'une courbe continue, le diagramme de répartition de la population étudiée.

classes	I	II	III	IV	V	VI	VII	VIII	IX	X	Total
limites (en K _Ω) des classes	92	94	96	98	100	102	104	106	108	110	
	94	96	98	100	102	104	106	108	110	112	
n_i	5	20	50	75	100	65	40	30	10	5	400
f_i	0,0125	0,05	0,125	0,1875	0,25	0,1625	0,10	0,075	0,025	0,0125	1
F_i	0,0125	0,0625	0,1875	0,375	0,625	0,7875	0,8875	0,9625	0,9875	1	



I-3 Localisation approximative d'une distribution

Pour localiser d'une manière approchée dans quelle zone de l'axe des x se situe une distribution, on utilise l'un ou l'autre des deux paramètres de position suivants

a) Médiane

Si l'on classe les valeurs de x par ordre croissant, la médiane correspond à la mesure de l'individu dont le rang correspond au milieu de cette suite. C'est aussi l'abscisse du point d'ordonnée $0,50$ sur le diagramme de répartition.

b) Moyenne arithmétique \bar{x}

La moyenne arithmétique est définie par la relation:

$$\bar{x} = \frac{\sum n_i x_i}{N} = \sum f_i x_i$$

I - 4 Ecart (par rapport à la moyenne arithmétique)

L'écart e_i de la mesure x_i d'un individu par rapport à la moyenne arithmétique de la population est défini par la relation

$$e_i = x_i - \bar{x}$$

On montre immédiatement que les écarts vérifient la relation

$$\sum f_i e_i = 0$$

I - V. Paramètres d'étalement

Pour caractériser l'étalement de la distribution on peut utiliser :

a) Soit son "étendue totale" ou différence entre les valeurs extrêmes de la variable x . L'étendue présente l'inconvénient d'être indépendante du mode de répartition des individus entre les limites extrêmes de la population.

b) Soient les écart moyens qui peuvent être (à défaut de l'écart algébrique moyen qui est nul) :

1°) soit l'écart absolu moyen

$$|\bar{e}| = \sum f_i |e_i|$$

abandonné en raison de sa fonction non analytique et dont les lois d'échantillonnage sont complexes ou inconnues

2°) soit l'écart quadratique moyen ou

écart-type σ , dont le carré

σ^2 est défini par la relation suivante :

$$s^2 = \frac{\sum n_i e_i^2}{N} = \sum f_i e_i^2$$

s'écrit encore

$$s^2 = \frac{\sum n_i (x_i - \bar{x})^2}{N}$$

I - 6 Calcul numérique des écarts-types

on n'utilise jamais la formule de définition pour calculer la valeur numérique de l'écart-type. En effet les écarts $e_i = x_i - \bar{x}$ ne sont pas, en général, des nombres simples. Leurs opérations au carré sont gênantes et sujettes à erreurs. En développant le numérateur de la formule de définition on obtient la formule suivante, qui est exclusivement utilisée pour les calculs numériques :

$$\begin{aligned} \sum n_i (x_i - \bar{x})^2 &= \sum n_i x_i^2 - 2\bar{x} \sum n_i x_i + \bar{x}^2 \sum n_i \\ &= \sum n_i x_i^2 - 2 \left(\frac{\sum n_i x_i}{N} \right)^2 + N \left(\frac{\sum n_i x_i}{N} \right)^2 \end{aligned}$$

s'écrit, en définitive

$$s^2 = \frac{1}{N} \left\{ \sum n_i x_i^2 - \frac{1}{N} \left(\sum n_i x_i \right)^2 \right\}$$

le terme $\Delta = \frac{1}{N} \left(\sum n_i x_i \right)^2$ s'appelle "terme correctif". Il tient compte du fait que l'origine à laquelle est rapportée la variable x ne coïncide pas, en général, avec la moyenne arithmétique.

Les calculs de la moyenne arithmétique et de l'écart-type sont, en outre, très simplifiés par des changements de variable (origine et unité) permettant d'utiliser des nombres plus simples.

Ex. Calcul de \bar{x} et s pour la population de 400 résistances du paragraphe I-2

I-7

Dans ce calcul on attribuera à toutes les résistances d'une même classe la valeur médiane de la classe à laquelle elles appartiennent. De plus on utilisera une variable auxiliaire θ définie arbitrairement par la relation

$$\theta = \frac{x - 101}{2}$$

La variable θ est, par suite, un nombre sans dimension qui varie d'une unité quand on passe d'une classe à la suivante. L'unité de la variable θ correspond à $2 \text{ K}\Omega$.

La présentation sous forme d'un tableau permet d'effectuer facilement les opérations.

Numéro du classe	1 Valeurs médianes des classes	Variable auxiliaire θ_i	Fréquences absolues n_i	$n_i \cdot \theta_i$	$n_i \cdot \theta_i^2$
I	93 K Ω	- 4	5	- 20	80
II	95 "	- 3	20	- 60	180
III	97 "	- 2	50	- 100	200
IV	99 "	- 1	75	- 75	75
V	101 "	0	100	0	0
VI	103 "	+ 1	65	+ 65	65
VII	105 "	+ 2	40	+ 80	160
VIII	107 "	+ 3	30	+ 90	270
IX	109 "	+ 4	10	+ 40	160
X	111 "	+ 5	5	+ 25	125
Totaux			400	+ 45	1315

Partant des résultats $N = 400$ $\sum n_i \theta_i = + 45$ $\sum n_i \theta_i^2 = 1315$
on en déduit :

$$\bar{\theta} = + \frac{1}{400} \times 45 = + 0,1125$$

$$\text{Termes correctif } \Delta = \frac{1}{400} \times \frac{45^2}{4} = 5,0625$$

$$\sum n_i (\theta_i - \bar{\theta})^2 = 1315 - 5,06 = 1309,94$$

$$\frac{1}{\theta} = \frac{1}{400} \times 1309,94 = 3,275$$

$$s_{\theta} = \sqrt{3,275} = 1,81$$

Il faut maintenant rapporter ces résultats
 $\bar{\theta} = +0,1125$ $s_{\theta} = 1,81$, relatifs à la variable
auxiliaire θ , aux résultats relatifs à la variable x .
On trouve immédiatement :

$$\left\{ \begin{array}{l} \bar{x} = 101 + 2 \times 0,1125 = 101,225 \text{ K}\Omega \\ s_x = 2 \times 1,81 = 3,62 \text{ K}\Omega \end{array} \right.$$

I-7 Modèle mathématique d'une distribution

On a cherché à représenter la fonction de répartition $F(x)$ d'une distribution par une expression mathématique ne dépendant que d'un petit nombre de paramètres. Si l'ajustement ainsi réalisé est assez exact, la connaissance de la distribution se ramènera à la connaissance des valeurs des paramètres caractéristiques. Parfois l'expression mathématique du modèle adopté rend compte du caractère ayant présidé à la genèse de la population étudiée (par ex. statistiques coïncidantes en physique mathématique)

Une distribution théorique (ou modèle mathématique) sera donc définie par l'association, à toutes les valeurs possibles de la variable x , d'une fonction $F(x)$: appelée "fonction de répartition" de la distribution.

La fonction $F(x)$ doit posséder l'ensemble de propriétés suivantes :

- $F(x)$ doit être une fonction uniforme et qui n'est jamais décroissante.

- 7) $F(x)$ doit être continue, sauf éventuellement en un nombre fini ou en une infinité dénombrable de points.
- c) $F(x)$ doit vérifier les égalités suivantes.

$$F(-\infty) = 0 \quad F(+\infty) = 1$$

- d) En un point de discontinuité, la fonction $F(x)$ doit être continue à gauche:

$$F(x) = F(x - \varepsilon)$$

La fonction $F(x)$ permet de calculer immédiatement la proportion ΔF d'individus dont la mesure x est comprise dans un intervalle fini (a, b) quelconque :

$$\Delta F = F(b) - F(a) \quad \text{pour } a \leq x < b$$

I-8 Densité d'une distribution (variable continue)

Lorsque la variable x est continue, la proportion dF d'individus dont la mesure est comprise entre x et $x + dx$ est :

$$dF = F(x + dx) - F(x)$$

Si la fonction $F(x)$ admet une dérivée $f(x) = \frac{dF}{dx}$ alors :

$$dF = f(x) dx$$

ce qui veut dire que la "fréquence élémentaire" dF est proportionnelle à l'élément dx de l'intervalle considéré, avec un coefficient variable de proportionnalité $f(x)$, appelé "densité de la distribution".

- ~ dF est un nombre pur (sans dimension)
- ~ dx est de dimension x
- ~ $f(x)$ est de dimension x^{-1} et ne peut donc être confondu avec une fréquence élémentaire

Les figures suivantes explicitent les relations entre la fonction $f(x)$ et $F(x)$:

Fig I-5 Courbe de "distribution" (densité) d'une variable continue

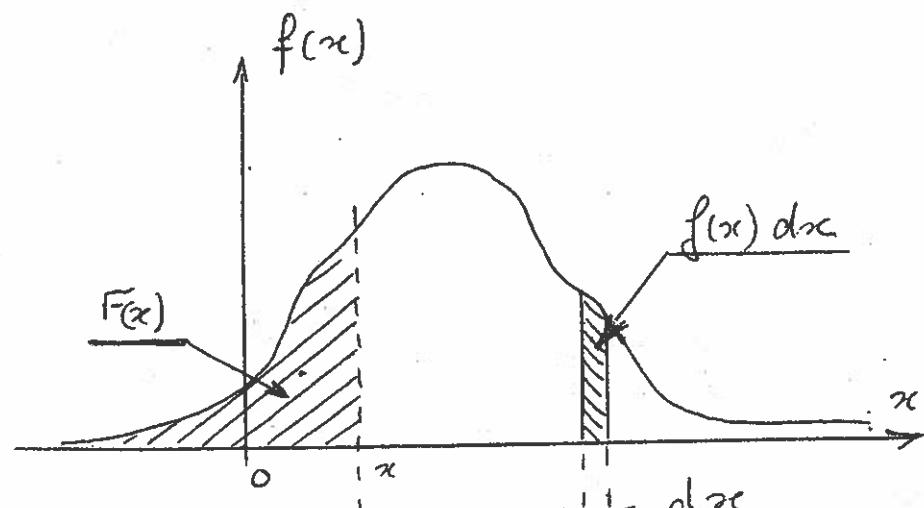
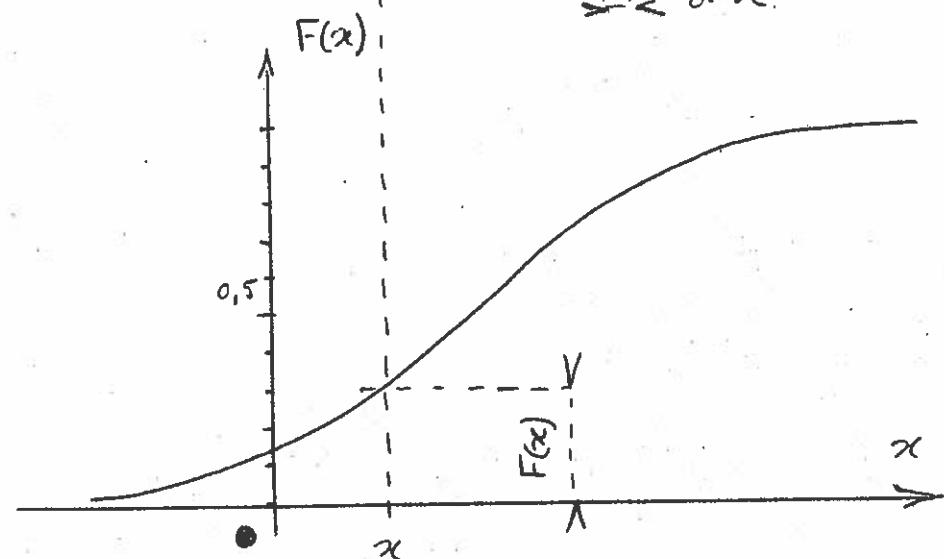


Fig I-6 Courbe de répartition



$$\left\{ \begin{array}{l} F(x) = \int_{-\infty}^x f(z) dz \\ \int_{-\infty}^{+\infty} f(x) dx = 1 \end{array} \right.$$

L'aire totale comprise entre l'axe des x et la courbe de distribution représente la somme des fréquences de tous les individus appartenant à la population et cette somme est évidemment égale à 1

I - 9 Modèle mathématique d'une distribution discrète

La variable x ne peut prendre qu'une suite discrète de valeurs distinctes x_1, x_2, x_3, \dots .
On peut définir une distribution théorique en associant à chacune des valeurs possibles de la variable x une fonction $p(x)$ désignant la probabilité de l'occurrence de la valeur x .

Annexe à la suite des nombres entiers positifs (zéro inclus) on peut associer la fonction

$$p(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (\text{Paramètre } \lambda \text{ positif constant})$$

qui définit la distribution de Poisson.

Les fonctions $p(x)$ choisis doivent être telles que leur somme étendue à toutes les valeurs possibles de la variable x , soit égale à l'unité. On peut, à titre d'exemple, vérifier que il en est bien ainsi pour la distribution de Poisson :

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \left\{ 1 + \frac{\lambda}{1} + \dots + \frac{\lambda^k}{k!} + \dots \right\} = e^{-\lambda} \times e^{\lambda} = 1$$

La fonction de répartition $F(x)$ est alors définie comme la somme de termes en nombre variable, soit pour la distribution de Poisson ; en supposant x compris entre les deux entiers successifs n et $n+1$:

$$F(x) = \sum_{z=c}^{z=n} e^{-\lambda} \frac{\lambda^z}{z!} \quad \text{pour } n < x \leq n+1$$

I - 10 Espérance mathématique d'une variable x

La moyenne arithmétique d'une distribution théorique s'appelle "Espérance mathématique" de la variable x . Le symbole correspondant s'écrit

$E[x]$ que l'on lit : "Espérance mathématique de x ou plus simplement $E[x]$ "

L'emploi du terme "espérance mathématique" permet de spécifier, sans ambiguïté, qu'il s'agit d'une moyenne théorique calculée pour une distribution mathématique et non de la moyenne arithmétique des individus d'une population concrète.

L'espérance mathématique s'exprime, dans tous les cas, par l'intégrale de Stieltjes

$$E[x] = \int_{-\infty}^{+\infty} x dF$$

En pratique cette intégrale se traduit

- a) pour une variable discrète par la somme, étendue à toutes les valeurs de x ,

$$E[x] = \sum_x x p(x)$$

- b) pour une variable continue dont la distribution possède une densité $f(x)$, par l'intégrale de Riemann :

$$E[x] = \int_{-\infty}^{+\infty} x f(x) dx$$

I - 11 Variable centrée

Si l'on rajoute une variable x , non pas à une origine quelconque mais à son espérance mathématique, on obtient une variable dite "variable centrée"

$$\xi = x - E[x]$$

on pourra également appeler la variable centrée ξ "écart théorique". Ce terme n'a pas de sens physique. Il est évident que l'espérance mathématique d'une variable centrée est nulle : $E[\xi] = 0$

I - 12. Ecart-type théorique σ . Variance de x

Le carré σ^2 de l'écart-type théorique σ est défini par la relation :

$$\sigma^2 = E[\xi^2] = \begin{cases} \sum_{-\infty}^{+\infty} (x - E[x])^2 p(x) & (\text{variable discrète}) \\ \int_{-\infty}^{+\infty} (x - E[x])^2 f(x) dx & (\text{variable continue}) \end{cases}$$

On appelle "variance de x " le carré σ^2 de l'écart-type σ .

I - 13. Variable réduite.

Il est naturel de choisir l'écart-type σ , caractéristique intrinsèque de la distribution, comme unité de mesure des écarts ξ par rapport à l'espérance mathématique $E[x]$. On obtient ainsi la variable réduite t qui a pour définition la relation :

$$t = \frac{\xi}{\sigma} = \frac{x - E[x]}{\sigma}$$

On vérifie facilement que l'espérance mathématique de la variable réduite est nulle et que son écart-type σ_t est égal à l'unité.

$$\left\{ \begin{array}{l} E[t] = 0 \\ \sigma_t = 1 \end{array} \right.$$

Il est important d'observer que la variable réduite t est un nombre sans dimension.

I - 14. Moments d'une distribution.

Si q étant un nombre positif quelconque, on appelle moment d'ordre q de la variable x "la quantité":

$$m_q = E[x^q] = \begin{cases} \sum x^q p(x) & (\text{variable discrète}) \\ \int_{-\infty}^{+\infty} x^q f(x) dx & (\text{variable continue}) \end{cases}$$

Par extension de cette définition on posera:

$$m_0 = E[x^0] = 1$$

Le moment d'ordre 1 n'est autre que l'espérance mathématique : $m_1 = E[x]$

I - 15 Moments rapportés à une origine quelconque ℓ

Les moments m'_q de la variable $x' = x - \ell$, rapportée à l'origine ℓ , s'expriment à partir des moments m_q de la variable x par la relation:

$$m'_q = E[(x')^q] = E[(x - \ell)^q] = \sum_{j=0}^{q-j} (-1)^j \binom{q}{j} \ell^{q-j} m_j$$

I - 16. Moments centrés μ_q

Si dans le changement de variable précédent on prend comme nouvelle origine l'espérance mathématique $\ell = E[x]$, on obtient les moments de la variable centrée \tilde{x} . Ces moments sont appelés les moments centrés de la distribution et désignés par la lettre grecque μ .

$$\mu_q = E\left[\left(\tilde{x}\right)^q\right] = E\left[\left(x - E[x]\right)^q\right] = \begin{cases} \sum [x - E[x]]^q p(x) \\ \int_{-\infty}^{+\infty} (x - E[x])^q f(x) dx \end{cases}$$

On notera, en particulier, que

I - 15

$$\left. \begin{array}{l} \mu_0 = 1 \\ \mu_1 = 0 \\ \mu_2 = \sigma^2 \text{ ou variance} \end{array} \right\}$$

La formule générale du paragraphe I - 15 permet, en particulier, d'établir les formules suivantes entre les moments m et les moments centrés μ :

$$\left. \begin{array}{ll} m_0 = 1 & \mu_0 = 1 \\ m_1 = m_1 & \mu_1 = 0 \\ m_2 = \mu_2 + m_1^2 & \mu_2 = m_2 - m_1^2 \\ m_3 = \mu_3 + 3\mu_2 m_1 + m_1^3 & \mu_3 = m_3 - 3m_1 m_2 + 2m_1^3 \\ m_4 = \mu_4 + 4\mu_3 m_1 + 6\mu_2 m_1^2 + m_1^4 & \mu_4 = m_4 - 4m_1 m_3 + 6m_1^2 m_2 - 3m_1^4 \\ \cdots & \cdots \end{array} \right\}$$

La formule $\mu_2 = m_2 - m_1^2$ montre que le moment centré d'ordre 2, est plus petit que tout moment d'ordre 2 rapporté à une origine différente ($m_1 \neq 0$) de l'espérance mathématique.

I - 17. Changement de variable dans le calcul des moments centrés.

Proposons nous de calculer le moment centré $(3) \mu_9$ de la nouvelle variable

$$Z = ax + b$$

On a.

$$(3) \mu_9 = E[(Z - E[Z])^9] = E[\{a(x - E[x])\}^9] = a^9 E[(x - E[x])^9]$$

soit

$$(3) \mu_q^q = a \times (\bar{x}) \mu_q^q$$

Cette relation montre que, au coefficient d'échelle a^q près, les moments centraux de la variable \bar{z} sont égaux à ceux de la variable \bar{x} . On dit que ces moments centraux sont des "semi-invariants" pour la transformation $\bar{z} = a\bar{x} + b$.

Si l'on considère, en particulier, la transformation $\bar{x} = m_1 + t\sigma$ qui relate la variable x à la variable réduite t , on trouve la relation:

$$(\bar{x}) \mu_q^q = \sigma^q \times (\bar{t}) \mu_q^q$$

Inégalité de Bienaymé - Tchebycheff

Soit une variable x admettant la loi de répartition $F(x)$.
 Soit $m_1 = E[x]$ et σ l'écart-type σ ,
 formons, de part et
 d'autre de la valeur
 $x = m_1$, deux segments
 égaux de longueur
 $t\sigma$, t étant un
 nombre positif quelconque.
 On définit ainsi un intervalle
 $[a, b]$ dont les limites ont pour
 valeurs:

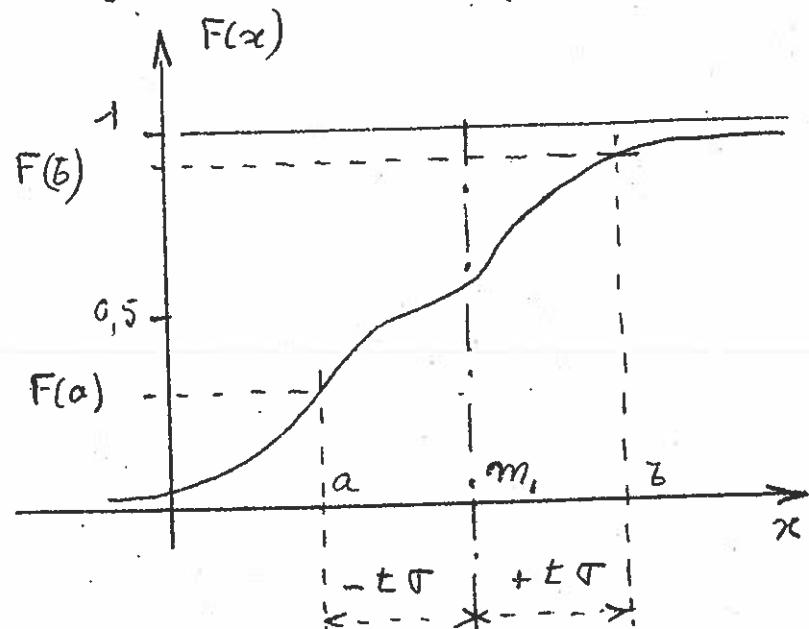


Fig I-7

$$\begin{cases} a = E[x] - t\sigma = m_1 - t\sigma \\ b = E[x] + t\sigma = m_1 + t\sigma \end{cases}$$

On sait que la proportion d'individus dont la mesure est comprise dans l'intervalle (a, b) a pour valeur exacte

$$\Delta F = F(b) - F(a)$$

Mais alors montrons que, quel que soit la loi de répartition $F(x)$, on peut affirmer que

$$\Delta F \geq 1 - \frac{1}{t^2}$$

Pour effectuer cette démonstration, nous utiliserons la variable centrée $\xi = x - m_1$ et nous partissons de la définition du moment centré d'ordre 2 $\mu_2 = \sigma^2$

$$\mu_2 = \sigma^2 = \int_{-\infty}^{+\infty} (x - m_1)^2 dF = \int_{-\infty}^{+\infty} \xi^2 dF$$

On peut alors écrire

$$\sigma^2 = \int_{-\infty}^{-t\sigma} \xi^2 dF + \int_{-t\sigma}^{+t\sigma} \xi^2 dF + \int_{+t\sigma}^{+\infty} \xi^2 dF \geq \int_{-\infty}^{-t\sigma} \xi^2 dF + \int_{+t\sigma}^{+\infty} \xi^2 dF$$

Car l'intégrale supérieure est toujours positive ou nulle. L'inégalité obtenue ne prouve seulement que les deux domaines d'intégration $(-\infty, -t\sigma)$ et $(+t\sigma, +\infty)$ dans lesquels $|\xi| \geq t\sigma$. On peut donc, à fortiori, écrire :

$$\sigma^2 \geq \int_{-\infty}^{-t\sigma} t^2 \sigma^2 dF + \int_{+t\sigma}^{+\infty} t^2 \sigma^2 dF = t^2 \sigma^2 \left\{ \int_{-\infty}^{-t\sigma} dF + \int_{+t\sigma}^{+\infty} dF \right\}$$

Soit, en divisant par σ^2 :

$$1 \geq t^2 \left\{ \int_{-\infty}^{-t\sigma} dF + \int_{+t\sigma}^{+\infty} dF \right\}$$

Or, par définition

$$\int_{-\infty}^{-t} dF = F(a) \quad \int_{+t}^{+\infty} dF = 1 - F(b)$$

On a donc bien :

$$1 \geq t^2 \{ F(a) + 1 - F(b) \} \text{ ou encore } F(b) - F(a) \geq 1 - \frac{1}{t^2}$$

On retiendra de cette inégalité que, quelle que soit la loi de répartition $F(x)$ d'une population, on peut toujours affirmer :

qui au moins 75% de la population sont compris dans l'intervalle $m_1 \pm 2\sigma$

qui au moins 89% de la population sont compris dans l'intervalle $m_1 \pm 3\sigma$

I-18 Fonction caractéristique d'une distribution

S'il est possible, comme a l'œil, de définir le modèle mathématique d'une distribution en spécifiant la fonction de répartition $F(x)$ correspondante, on démontre qu'il est également possible de définir cette même distribution "par la connaissance d'une fonction $\Phi(u)$, appelée "fonction caractéristique de la distribution." Les fonctions F et Φ forment un couple de fonctions associées en correspondance biunivoque.

La variable u étant une variable réelle⁽⁴⁾ auxiliaire indépendante de x , la fonction caractéristique $\Phi(u)$ est définie par la relation :

$$\Phi(u) = E[e^{iux}] = \int_{-\infty}^{+\infty} e^{iux} dF = \begin{cases} \sum_x e^{iux} p(x) \\ \int_{-\infty}^{+\infty} e^{iux} f(x) dx \end{cases}$$

(4) La variable u a pour dimension x^{-1} car ux doit être un nombre fini.

La convergence de l'intégrale définissant la fonction caractéristique est toujours assurée, du fait que $|e^{iux}| = 1$ et que l'intégral $\int dF = 1$

La fonction caractéristique est, en général, une quantité complexe $\Phi(u) = R(u) + i J(u)$ dont les parties réelles et imaginaires ont respectivement pour expressions :

$$\left\{ \begin{array}{l} R(u) = \int_{-\infty}^{+\infty} \cos(ux) dF \\ J(u) = \int_{-\infty}^{+\infty} \sin(ux) dF \end{array} \right.$$

La fonction caractéristique n'est réelle qu'à la condition que la distribution soit symétrique par rapport à l'origine de la variable x .

Les fonctions $R(u)$ et $J(u)$ étant respectivement paires et impaires par rapport à la variable u , on en déduit les relations suivantes :

$$\left\{ \begin{array}{l} R(u) = \frac{1}{2} \{ \Phi(u) + \Phi(-u) \} \\ J(u) = \frac{1}{2i} \{ \Phi(u) - \Phi(-u) \} \\ |\Phi(u)| = \sqrt{\Phi(u) \times \Phi(-u)} \end{array} \right.$$

I - 19. Développement en série de Mac-Laurin de la fonction caractéristique

Si la fonction caractéristique $\Phi(u)$ admet, au voisinage de $u = 0$, des dérivées successives, on sait qu'elle est alors dérivable en série de Mac-Laurin, selon la formule :

$$\Phi(u) = 1 + \Phi'(0)u + \dots + \frac{u^k}{k!} \left(\frac{d^k \Phi}{du^k} \right) + \dots$$

Or, si l'on calcule la forme dérivée par rapport à u de la fonction Φ , on trouve :

$$\frac{d^k \Phi}{du^k} = \int_{-\infty}^{+\infty} (ix)^k e^{iux} dF$$

$$\text{On voit donc que } \left(\frac{d^k \Phi}{du^k} \right)_{u=0} = i^k \int_{-\infty}^{+\infty} x^k dF = i^k m_k \quad \boxed{m_k = (-i)^k \left[\frac{\partial \Phi}{\partial u^k} \right]_{u=0}} \quad \text{I-20}$$

L'existence des dérivées successives de la fonction caractéristique est liée à l'existence des moments du même ordre. Dans le cas où ces moments existent, la fonction caractéristique admet alors le développement

$$\Phi(u) = 1 + m_1 iu - \frac{m_2}{2!} u^2 + \dots - \frac{(iu)^k}{k!} m_k + \dots$$

On obtiendrait de même, pour la variable centrée $\xi = x - m_1$, l'expression

$$\Phi(u) = E[e^{iu(\xi - m_1)}] = 1 - \frac{\mu_2}{2} u^2 + \dots - \frac{(iu)^k}{k!} \mu_k + \dots$$

La fonction caractéristique, dont le calcul ne demande qu'une seule intégration, permet donc de déterminer la suite infinie des moments d'une distribution.

I-20 Deuxième fonction caractéristique.

On utilise fréquemment une deuxième fonction caractéristique $\Psi(u)$ définie par la relation

$$\Psi(u) = \log_e (\Phi(u))$$

En développant en série la relation

$$\Psi(u) = \log_e \left\{ 1 + m_1 iu - \frac{m_2}{2!} u^2 + \dots - \frac{(iu)^k}{k!} m_k + \dots \right\}$$

on démontre facilement que la seconde fonction caractéristique admet le développement en série suivant :

$$\Psi(u) = m_1 iu - \frac{u^2}{2} \mu_2 + \frac{(iu)^3}{3!} \mu_3 + \frac{(iu)^4}{4!} \mu_4 + \dots - \frac{(iu)^q}{q!} \mu_q + \dots$$

exprimer dans l'équation les coefficients μ_i , appelés semi-invariants de la distribution, ne sont fonction que des moments centres μ_i , même si l'origine choisie pour la variable se trouve dans pas avec son origine mathématique. On note, en particulier que

$$\left\{ \begin{array}{l} X_4 = \mu_4 - 3\mu_2^2 \\ X_5 = \mu_5 - 10\mu_3\mu_2 \\ X_6 = \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3 \end{array} \right. \quad \left\{ \begin{array}{l} \mu_4 = X_4 + 3X_2^2 \\ \mu_5 = X_5 + 10X_3X_2 \\ \mu_6 = X_6 + 15X_4X_2^2 + 10X_3^2 + 15X_2^3 \end{array} \right.$$

L'exemple de la seconde fonction caractéristique permet donc d'obtenir les deux deux premiers moments centres sans avoir à effectuer le changement de variable $\tilde{x} = x - m_1$ correspondant à la variable centrée.

I-21 Changement de variable et fonctions caractéristiques.

Soit la nouvelle variable $\tilde{x} = ax + b$, la fonction caractéristique $\underline{\Phi}_{\tilde{x}}(u)$ de la variable \tilde{x} s'écrit:

$$\underline{\Phi}_{\tilde{x}}(u) = E[e^{iu(a\tilde{x}+b)}] = e^{iub} \times E[e^{iu(ax)}]$$

Notre équation.

$$\left\{ \begin{array}{l} \underline{\Phi}_{\tilde{x}}(u) = e^{iub} \times \underline{\Phi}_x(au) \\ \psi_{\tilde{x}}(u) = iub + \psi_x(au) \end{array} \right.$$

Si l'on considère, en particulier, la transformation $x = m_1 + t\sigma$ qui relate la variable x à la variable réduite t , on voit que:

$$\left\{ \begin{array}{l} \underline{\Phi}_x(u) = e^{ium_1} \times \underline{\Phi}_t(u\sigma) \\ \psi_x(u) = ium_1 + \psi_t(u\sigma) \end{array} \right.$$

I-22. Passage de la fonction caractéristique à la distribution I-22

Nous admettrons sans démonstration que le procédé de calcul suivant permet de passer de la fonction caractéristique $\Phi(u)$ à la distribution.

1°) On calculera tout d'abord l'intégrale:

$$S(x) = \lim_{u \rightarrow \infty} \left\{ \frac{1}{2\pi} \int e^{-iux} \Phi(u) du \right\}$$

Deux cas sont à considérer

- 1a. $S(x)$ est nulle pour toute valeur de x , sauf un nombre fini ou infini de valeurs, notées x_1, x_2, \dots, x_k , pour lesquelles l'intégrale $S(x)$ prend des valeurs comprises entre 0 et 1.

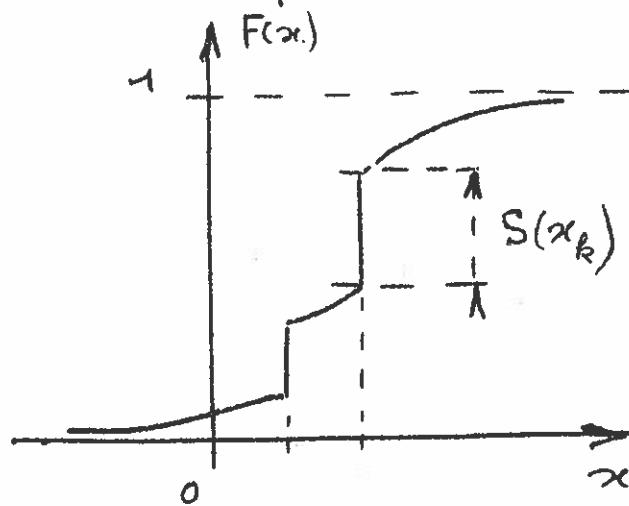


Fig I-8

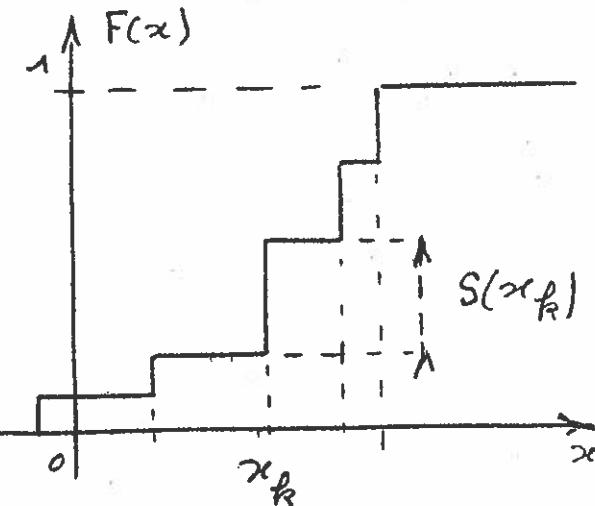


Fig I-9

On doit se concéder que la fonction de répartition $F(x)$ de la distribution présente, pour les valeurs x_1, x_2, \dots, x_k , des ressorts (Fig I-8) dont les amplitudes sont précisément égales aux valeurs correspondantes $S(x_1), S(x_2), \dots, S(x_k)$ de l'intégrale $S(x)$.

En général cette situation inclut que la distribution est une distribution discrète. La variable x ne pouvant prendre que des valeurs discrètes $x_1, x_2, \dots, x_k, \dots$ telle que celles sont associées les fréquences $S(x_1), S(x_2), \dots, S(x_k)$ (comme de répartition "éteinte" Fig I-9). Dans ce cas la somme de tous les ressorts est égale à l'unité: $\sum S(x_k) = 1$

S'il n'en est pas ainsi, la distribution

présente un caractère mixte (Fig I-8) de concentration I-23
Sur les valeurs discrètes x_1, x_2, \dots, x_k et
de répartition continue dont la densité devra
être calculée selon le méthode du paragraphe 2:

I-7 $S(x)$ est nulle pour toute valeur de x

On doit en conclure que la fonction
de répartition $F(x)$ est continue en tous points
et ne porte de sauts au aucun point.

Si la fonction $F(x)$ est déivable, la
densité $f(x)$ de la distribution peut alors être
calculée comme suit:

- 2°) Dans le cas d'une distribution continue admettant
une densité $f(x)$, cette densité est donnée
par l'intégrale:

$$f(x) = \lim_{n \rightarrow \infty} \left\{ \frac{1}{2\pi} \int_{-n}^{+n} e^{-iux} \hat{\phi}(u) du \right\}$$

Rév. Les formules permettant le passage de la fonction
caractéristique à la distribution sont appelées
"Formules de réciprocité de Fourier". Notons
au passage que le diagramme de distribution
constitue en fait le "Spectre de fréquence"
de la fonction caractéristique $\hat{\phi}(u)$, si
l'on attribue la variable u au temps et
la variable x à la pulsation $W = 2\pi u$
d'un mouvement sinusoïdal de fréquence variable v .

I-23 Suite de distributions. Recherche de la distribution limite

Dans l'étude des problèmes d'échantillonnage,
nous aurons à considérer des distributions, dont
les fonctions de répartition $F_n(x)$ dont l'expression
analytique dépend d'une manière connue de
l'effectif n de l'échantillon prisé.

Si l'on considère des échantillons dont les effectifs
 n croissent selon la suite des nombres entiers,
nous leur ferons correspondre une suite

de fonctions de répartition $F_1(x), F_2(x) \dots F_n(x) \dots$ auxquelles nous pouvons associer les fonctions caractéristiques correspondantes $\Phi_1(u), \Phi_2(u) \dots \Phi_n(u) \dots$

Effectif de l'échantillon	1	2	3	n	∞
Fonction de répartition	$F_1(x)$	$F_2(x)$	$F_3(x)$	$F_n(x)$	$F_\infty(x)$
Fonction caractéristique	$\Phi_1(u)$	$\Phi_2(u)$	$\Phi_3(u)$	$\Phi_n(u)$	$\Phi_\infty(u)$

On démontre que :

Si la suite $F_n(x)$ des fonctions de répartition converge uniformément vers une fonction de répartition limite $F_\infty(x)$, la suite correspondante des fonctions caractéristiques converge vers une fonction limite $\Phi_\infty(u)$ qui s'identifie avec la fonction caractéristique de la $F_\infty(x)$ de répartition limite.

Le théorème précédent étant évidemment vrai on voit que l'on peut obtenir le $F_\infty(x)$ en calculant tout d'abord la fonction caractéristique limite $\Phi_\infty(u)$ puis en utilisant les formules de réciproque pour déterminer $F_\infty(x)$.

Cette méthode est en général plus simple que le calcul direct de $F_\infty(x)$. Ce résultat justifie pleinement l'emploi de la fonction caractéristique, tout comme les résultats du calcul symbolique justifient l'emploi de la transformation de Laplace.

I.-24. Population normale. Loi de Laplace-Gauss

Pour illustrer les notions précédemment acquises, nous terminerons ce chapitre en étudiant les propriétés de la loi de Laplace-Gauss qui constitue l'un des modèles mathématiques les plus utilisés pour représenter des distributions continues.

Cette loi doit principalement son intérêt au fait qu'elle constitue la forme de loi limite de nombreuses distributions rencontrées dans la théorie de l'échantillonnage. La loi de Gauss est, en particulier, la loi limite des moyennes arithmétiques.

La loi normale est définie par la relation :

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

La variable x pouvant varier de $-\infty$ à $+\infty$

On vérifie facilement que

$$\int_{-\infty}^{+\infty} dF = 1$$

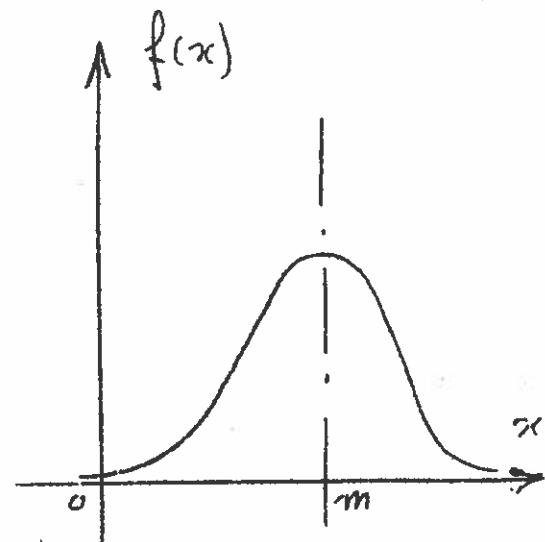


Fig I-10. Courbe de distribution de la loi Normale

La courbe de distribution (Fig I-10) est visuellement symétrique par rapport à l'axe $x = m$. On a donc $E[x] = m$. L'étude de la fonction caractéristique permet de démontrer que le paramètre σ est bien l'écart-type de la variable normale.

I-25 Loi normale réduite

$$\text{Si l'on pose } t = \frac{x-m}{\sigma}$$

La loi normale réduite s'écrit

$$dF = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} t^2} dt$$

I-26 Fonction caractéristique de la loi normale

On calculera tout d'abord la fonction caractéristique de la variable réduite t

$$\Phi(u) = E[e^{iut}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{iut} e^{-\frac{1}{2}t^2} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{t^2 - 2iut\}} dt$$

$$\underline{\Phi}_t(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{t^2 - 2iut + i^2u^2 - i^2u^2\}} dt$$

$$\underline{\Phi}_t(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{t^2 - 2iut + \frac{i^2u^2}{2}\}} e^{-\frac{1}{2}\{t - iu\}^2} dt$$

$$\underline{\Phi}_t(u) = e^{-\frac{1}{2}u^2} \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{t - iu\}^2} dt$$

On peut également écrire

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{t - iu\}^2} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}t^2} dt = 1$$

et, par suite :

$$\begin{cases} \underline{\Phi}_t(u) = e^{-\frac{1}{2}u^2} \\ \underline{\psi}_t(u) = -\frac{1}{2}u^2 \end{cases}$$

Si on separe à la variable $x = m + t$ on a

$$\begin{cases} \underline{\Phi}_x(u) = e^{ium} \times e^{-\frac{1}{2}u^2\sigma^2} \end{cases}$$

$$\begin{cases} \underline{\psi}_x(u) = ium - \frac{1}{2}u^2\sigma^2 \end{cases}$$

I-27. Moments centrés de la loi normale

I-27

Pour obtenir directement les moments centrés de la variable x , on partira du développement en série de la première fonction caractéristique de la variable centrée $\xi = x - m$, que l'on identifiera avec le développement en série de Mac Laurin calculé au paragraphe I-19. On trouve ainsi

$$\left\{ \begin{array}{l} \Phi(u) = e^{-\frac{1}{2}u^2\sigma^2} = 1 - \frac{1}{2}\sigma^2 u^2 + \frac{1}{8}\sigma^4 u^4 + \dots \quad \frac{(-1)^k}{2^k k!} u^{2k} + \dots \\ \Phi(u) = 1 - \frac{1}{2}\mu_2 u^2 + \frac{1}{3!} \mu_3(u)^3 + \dots \quad \frac{1}{(2k)!} (iu)^{2k} \mu_{2k} + \dots \end{array} \right.$$

Par identification, on en déduit immédiatement :

1°) que tous les moments centrés d'ordre impair sont nuls (propriété caractéristique de distribution symétrique)

$$\mu_{2k+1} = 0$$

2°) que les moments d'ordre pair ont pour valeurs :

$$\left\{ \begin{array}{l} \mu_2 = \sigma^2 \\ \mu_4 = 3\sigma^4 \\ \mu_6 = 15\sigma^6 \\ \mu_8 = 85\sigma^8 \\ \vdots \\ \vdots \\ \mu_{2k} = \frac{(2k)!}{2^k k!} \sigma^{2k} \end{array} \right. \quad (\Gamma représente donc bien l'écart type?)$$

relations qui peuvent se résumer dans la formule de récurrence

$$\mu_{2k+2} = (2k+1)\sigma^2 \mu_{2k}$$

I-28 Table de la loi normale réduite

I-28

Pour les applications pratiques, il importe de connaître les valeurs de la fonction de répartition $F(t)$ de la variable normale réduite.

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{1}{2}z^2} dz$$

ou "ogive de Galton"

Cette intégrale ne pouvant s'exprimer à partir des fonctions tabulées dans les manuels universitaires, a été calculée numériquement. Ces calculs ont permis d'établir la table II qui donne les valeurs de la fonction $F(t)$ (qui figurent sur les pages de la table) en fonction de l'argument t qui figure dans le corps de la table. Savoir que le signe de t n'est pas indiqué.

La courbe de Gauss réduite étant symétrique, on voit sur la fig I-13 que

$$\begin{cases} \text{Pour } t < 0 & F(t) < 0,50 \\ " \quad t = 0 & F(0) = 0,50 \\ " \quad t > 0 & F(t) > 0,50 \end{cases}$$

$$F(-t) = 1 - F(t)$$

Il est enfin utile de se rappeler que pour une distribution normale

95% des individus ont leurs mesures comprises dans l'intervalle $m \pm 1,96\sigma$

99,8% des individus ont leurs mesures comprises dans l'intervalle $m \pm 3,09\sigma$
(à comparer avec les résultats du paragraphe

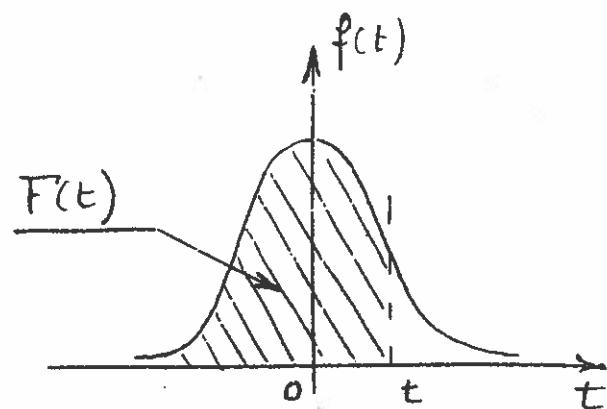


Fig I-11 Courbe de Gauss réduite

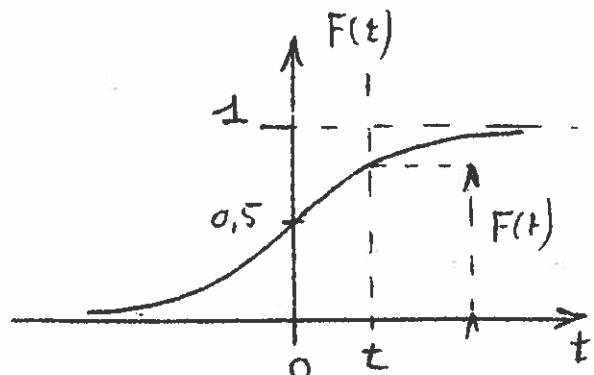


Fig I-12 Oogive de Galton

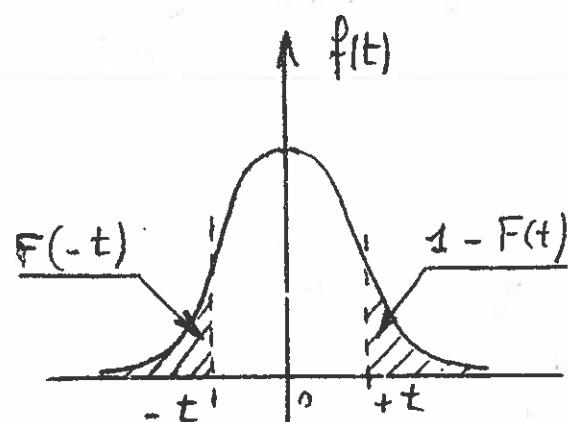


Fig I-13

Application numérique.

En admettant que la distribution des 400 résistances du paragraphe I - 2 puisse être assimilée à une loi normale, calculer, dans cette hypothèse, la proportion théorique de résistances dont la valeur serait comprise dans l'intervalle

$$x_1 = 98 \text{ K}\Omega \quad x_2 = 104 \text{ K}\Omega$$

Les étapes du calcul sont les suivantes.

1°) Choix des paramètres de la loi normale

Faute de connaître les valeurs exactes des paramètres m et σ de la loi normale, on adoptera pour ces paramètres les valeurs calculées au paragraphe I - 6, à savoir

$$\begin{cases} m = 101,2 \text{ K}\Omega \\ \sigma = 3,62 \text{ K}\Omega \end{cases}$$

2°) Calcul, en KΩ, des écarts correspondant aux limites de l'intervalle x_1 , x_2

$$\begin{cases} z_1 = x_1 - m = 98 - 101,2 = - 3,2 \text{ K}\Omega \\ z_2 = x_2 - m = 104 - 101,2 = + 2,8 \text{ K}\Omega \end{cases}$$

3°) Transformation des écarts en KΩ en écarts zérois

$$\begin{cases} t_1 = \frac{z_1}{\sigma} = - \frac{3,2}{3,62} = - 0,885 \\ t_2 = \frac{z_2}{\sigma} = + \frac{2,8}{3,62} = + 0,775 \end{cases}$$

4°) Consultation des tables

$$F_1 = F(-0,885) = 0,188 \quad F_2 = F(+0,775) = 0,781$$

5°) Calcul du résultat définitif

$$\Delta F = F_2 - F_1 = 0,781 - 0,188 = 0,593$$

sont 59,3% des résistances dans l'intervalle 98 KΩ - 104 KΩ

I - 30 Droite de Henry

En modifiant le mode de graduation de l'axe vertical de Po fig I - 12, représentant l'échelle de Galton, on peut établir un graphique dans lequel la loi de répartition de la distribution normale soit représentée par une droite.

Il suffit pour cela de prendre l'axe vertical en fonction de l'argument θ défini par l'équation

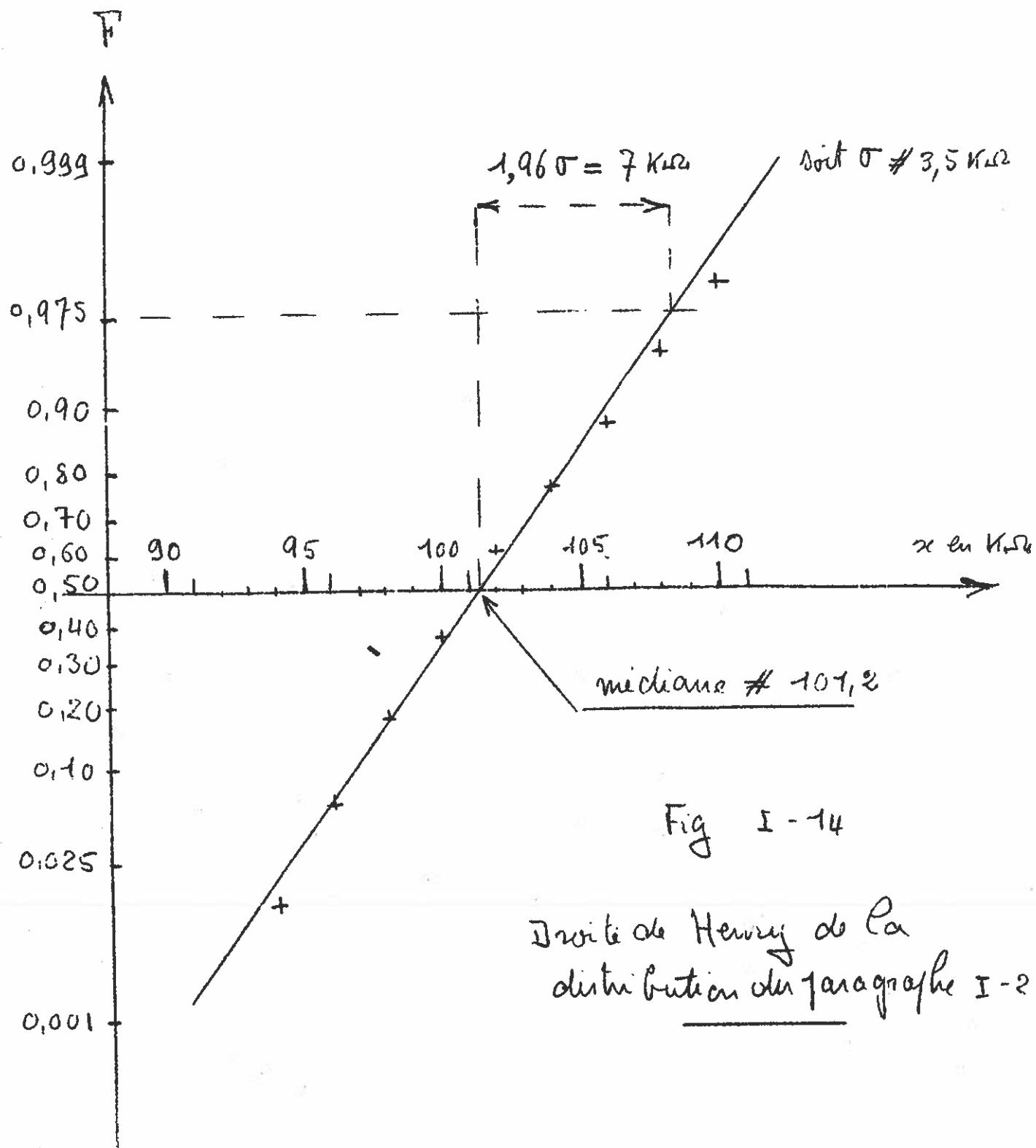
$$F = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta(F)} e^{-\frac{1}{2}z^2} dz$$

Il existe dans le commerce des jauge dont l'axe vertical est gradué selon cette méthode⁽¹⁾. Si l'on rapporte sur un tel jauge les points représentatifs d'une courbe expérimentale de fréquence cumulée, ces points doivent se situer approximativement sur une droite, si la population étudiée peut être assimilée à une population normale. Cette droite est appelée "droite de Henry" du nom de son inventeur, le capitaine Henry.

La fig I - 13 représente Po droite correspondant à Po distribution des 100 résistances du paragraphe I - 2, dont les fréquences cumulées avaient été calculées et reportées sur le tableau qui figure dans ce paragraphe.

On remarquera qu'il est possible de lire directement sur l'abaque de Henry des valeurs approchées de l'espérance mathématique et de l'écart type de la distribution (E_{avg}) est peu différent de Po médiane et Po centre de Po droite est immédiatement proportionnel à σ !

(1) L'échelle verticale de l'abaque IX (fonction de répartition de la loi de Poisson) correspond à ce type de graduation.



Dwite de Henry ob Pa
distribution von Paragraphen I - 2

Ch II. Distribution à deux caractères

II-1 Le concept de relation fonctionnelle entre deux variables

Les études classiques nous ont familiarisé avec la notion de relation fonctionnelle, qui apparaît comme le mode le plus souhaitable d'expression des lois physiques. Ainsi, la loi d'Ohm ne traduit-elle pas une relation linéaire entre la différence de potentiel aux bornes d'une résistance et l'intensité du courant qui la traverse. La vérification expérimentale de la loi d'Ohm est excellente et la répétition des essais ne permet pas de mettre en évidence de différences appréciables entre les résultats obtenus.

Il en résulte de même si, disposant d'un lot important de billes de roulement de diamètres différents, nous nous proposons de constater simultanément le diamètre x de ces billes et leur poids y . Nous trouvons que les couples (x_i, y_i) correspondant à chaque bille peuvent être représentés par des points se situant d'une manière parfaite sur une parabole de brûque (Fig 2-1)

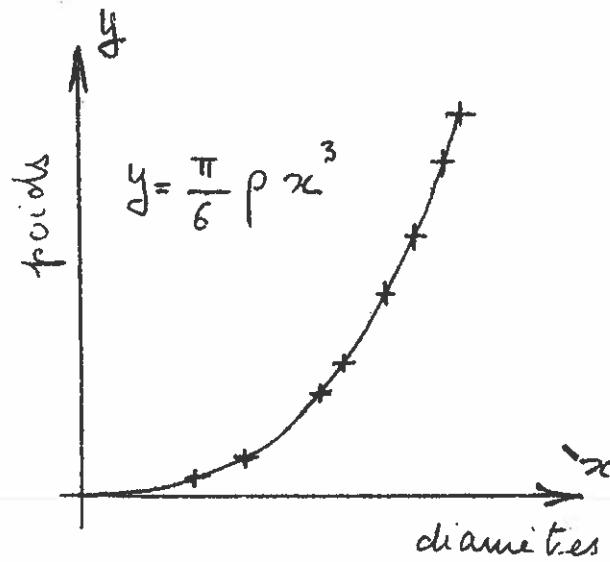


Fig 2-1. Relation fonctionnelle

Il est évident qu'en de telles circonstances la ménée de l'une, au choix, des deux variables x ou y permet de caractériser sans ambiguïté l'individu correspondant. L'une des deux variables est secondaire par rapport à l'autre. On dit aussi que la corrélation entre les variables x et y est parfaite.

II-2 Influence du concept de relation fonctionnelle

pour représenter une loi plus générale de dépendance entre deux variables

Avant de nous occuper du classement des t. tis, en fonction, si un élément, de leur diamètre et de leur poids, proposent nous maintenant de classer des hommes adultes d'une même race, en fonction de leur taille x et de leur poids y . Nous constaterons immédiatement que la notion de "relation fonctionnelle" ne permet plus de discuter avec f de l'ordre des observations reçues ci-dessous.

Pourtant, en effet, en abscisse x ,
taille x du individus.
et en ordonnée leur
poids y . Nous
constatons alors
(fig 2-2) que ces points
figurés ne se situent
pas sur une courbe
mais se groupent en
un nuage assez étendu.
On observe que pour une
même taille x deux
individus dont les tailles
différent peuvent
présenter des différences
importantes et que,
contrairement, deux
individus du même poids
 y peuvent avoir des
tailles très diverses.

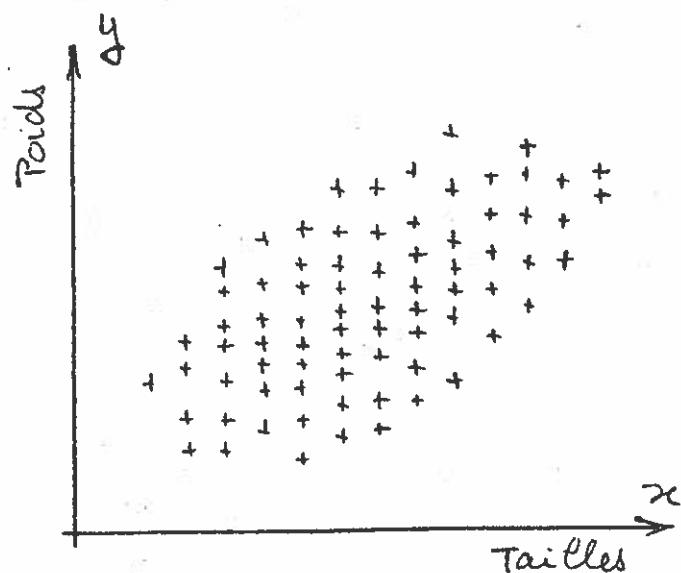


Fig 2-2. Corrélation

On ne peut plus songer à représenter par une courbe la relation entre les variables x et y qui sans être totalement indépendantes (les individus de petite taille sont, en moyenne, plus légers que ceux de grande taille) ne sont plus en relation fonctionnelle. On dit que les variables x et y sont liées par une "corrélation" dont nous nous proposons, dans ce chapitre, de définir les caractères principaux.

II-3. Étude quantitative d'une corrélation

S'il n'est plus possible, dans le cas d'une corrélation, de caractériser par une relation fonctionnelle la liaison entre les variables x et y , on peut cependant estimator les caractères moyens de cette liaison

en définissant
deux combinaisons
l'on appelle
"Combinaisons de régression"

Si l'on considère,
en effet, tous les
individus d'une
même taille x , on
peut alors définir
le poids moyen $Y(x)$
des individus de
cette taille. La combi-
naison de régression de y en x
n'est autre que la
combinaison de variation
des poids moyens $Y(x)$
en fonction de la
taille x des individus
correspondants.

On définit, de même,
la combinaison de régression de x en y , représentant la
variation des tailles moyennes $X(y)$ des individus
de même poids y .

Les deux combinaisons de régression sont, en général, distinctes
l'une de l'autre et ne passent pas par le centre de
gravité G de la distribution, sauf dans le cas où
elles sont toutes deux rectilignes.

Si lorsque la différence des valeurs de y correspondant
à une même valeur de x (et réciproquement) tend
vers zéro, les deux combinaisons de régression se rapprochent
l'une de l'autre pour se confondre en une combinaison
unique, qui traduit alors la relation fonctionnelle
littérale entre les variables x et y .

II - 4. Table de contingence

On présente en général sous forme d'une table
à double entrée les résultats d'un classement statistique
selon deux caractères. Cette table porte le nom de
"Table de contingence"

Soyons, par exemple, à représenter le classement des
 N individus en fonction de leur taille x et de leur
poids y . Les deux variables étant continues, on subdivise

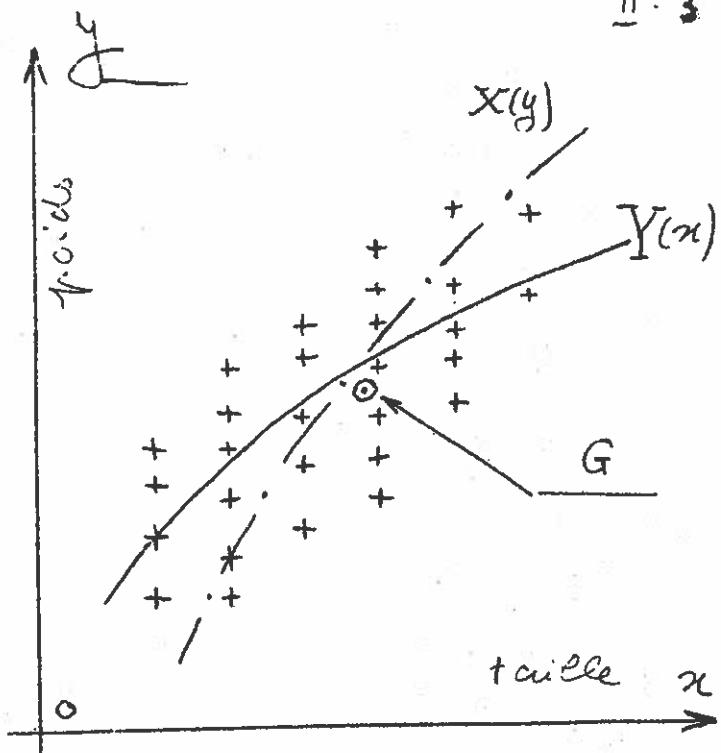


Fig 2-3. Combinaisons de régression

II-1. Table de Contingence

Poids y

Taille x	y_1	y_2	y_j	y_q	Totaux
x_1	n_{11}	n_{12}	n_{1j}	n_{1q}	$n_{1\cdot}$
x_2	n_{21}	n_{22}	n_{2j}	n_{2q}	$n_{2\cdot}$
\vdots					
x_i	n_{i1}	n_{i2}	n_{ij}	n_{iq}	$n_{i\cdot}$
\vdots					
x_p	n_{p1}	n_{p2}	n_{pj}	n_{pq}	$n_{p\cdot}$
Totaux	$n_{\cdot 1}$	$n_{\cdot 2}$	$n_{\cdot j}$	$n_{\cdot q}$	N

l'ens intervalles respectifs de variation en classes d'étatues école. Les volets mesurés de chaque classe seront respectivement désignés par $x_1, x_2 \dots x_p$ et $y_1, y_2 \dots y_q$, que l'on prend comme entrées de la table de contingence. On figure dans chacune des cases de la table le nombre n_{ij} d'individus ayant à la fois un tailles x_i et un poids y_j . En sommant par ligne et par colonne on obtiendra les nombres $n_{i\cdot}$ d'individus de taille x_i et $n_{\cdot j}$ d'individus de poids y_j . Enfin, l'^e effectif de chaque ligne et de chaque colonne et l'^e effectif global N vérifiant la relation

$$\left\{ \begin{array}{l} n_{i\cdot} = \sum_j n_{ij} \\ n_{\cdot j} = \sum_i n_{ij} \\ N = \sum_{ij} n_{ij} = \sum_i n_{i\cdot} = \sum_j n_{\cdot j} \end{array} \right.$$

5. Fréquence globale. Histogramme.

Si la fréquence $x_i y_j | H$ ⁽⁴⁾ des individus de taille x_i et de poids y_j est égale au quotient

$$x_i y_j | H = \frac{n_{ij}}{N}$$

que l'on désigne sous le nom de "fréquence globale" en ce sens qu'elle se rapporte à l'ensemble de la distribution.

On peut obtenir une représentation visuelle de la distribution en élévant, à l'échelle de chaque ligne de cases de la table de contingence, un pyramide dont l'aire proportionnelle à la fréquence globale de la case correspondante.

On obtient ainsi l'histogramme de la distribution.

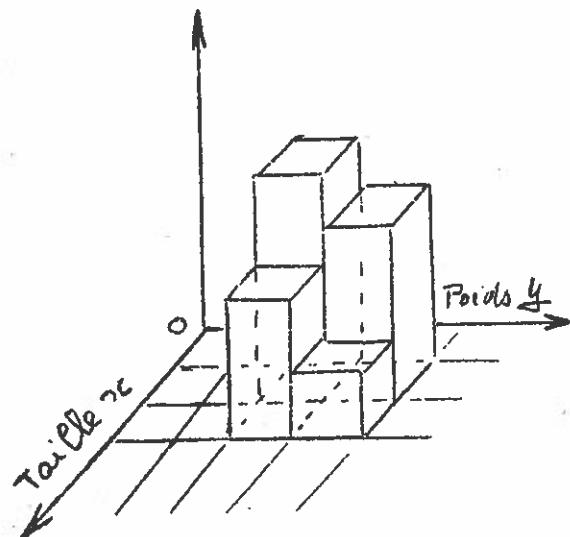


Fig II-4 Histogramme

II 6. Fréquences marginales.

On peut facilement déduire de la table de contingence la distribution des individus en fonction de leur taille, quel que soit leur poids. Cette distribution est, en fait, représentée par les sommes partielles, qui figurent sur la marge de droite du tableau. Les fréquences correspondant à cette distribution marginale sont égales au quotient :

(4) Le symbole $x_i y_j | H$ se lit "x_i et y_j dans le champ H".

Le champ H désigne la totalité de la table de contingence et est, en fait, synonyme de l'ensemble de la distribution.

$$x_i | H = \frac{n_{i \cdot}}{N}$$

On définit de même les fréquences marginales correspondant au classement des individus en fonction de leur poids, quelle que soit leur taille. Ces fréquences ont pour expressions

$$y_j | H = \frac{n_{\cdot j}}{N}$$

II - 7 Fréquences liées

Il peut être intéressant d'étudier la distribution, en fonction de leur taille, des individus d'un poids y_j donné. Cette distribution est représentée, sur le tableau de contingence, par la colonne y_j comportant un nombre N_j , tel que $n_{\cdot j}$ d'individus de poids y_j . Cette distribution s'appelle "distribution du volume de ce liées par y_j ". Les fréquences correspondantes se notent $y_j | x_i H$ et ont pour valeurs :

$$x_i | y_j H = \frac{n_{ij}}{n_{\cdot j}}$$

On définit de même la distribution, en fonction de leur poids, des individus d'une même taille x_i , ou distribution des y_j liés par x_i . Celle distribution correspond à la ligne x_i de ce tableau de contingence et ces fréquences liées correspondantes ont pour expressions :

$$y_j | x_i H = \frac{n_{ij}}{n_{i \cdot}}$$

- (a) Le symbole $x_i | y_j H$ se lit " x_i dans le champ $y_j H$ ". Ce champ réduit $y_j H$ représente la colonne y_j de la table de contingence c'est à dire, en fait, l'ensemble des individus de poids y_j .

Expression de la fréquence globale en fonction des fréquences marginales et liées.

Les identités arithmétiques suivantes.

$$\frac{n_{ij}}{N} = \frac{n_{i.}}{N} \times \frac{n_{.j}}{n_{i.}} \equiv \frac{n_{i.}}{N} \times \frac{n_{i.}}{n_{.j}}$$

montrent que la fréquence globale peut s'exprimer en fonction des fréquences marginales et liées par les deux expressions:

$$\left\{ \begin{array}{l} x_i y_j | H = x_i | H \times y_j | x_i H \\ x_i y_j | H = y_j | H \times x_i | y_j H \end{array} \right.$$

II. Variables indépendantes

On dit que deux variables x et y sont statistiquement indépendantes lorsque leurs distributions liées sont toutes indépendantes de la valeur du caractère de liaison. Il en serait ainsi dans l'exemple précédent si les distributions liées des poids des individus étaient indépendantes de la taille des individus considérés. Malheureusement cette définition n'exprime pas les conditions suivantes.

$$\left\{ \begin{array}{l} x_i | y_j H = x_i | y_k H \quad \text{quelque soient } x_i, \\ \qquad \qquad \qquad y_j \text{ et } y_k \\ y_j | x_i H = y_j | x_\ell H \quad \text{quelque soient } \\ \qquad \qquad \qquad y_j \text{ et } x_\ell \end{array} \right.$$

On démontre facilement que ces conditions exprimées par la seconde ligne sont compatibles avec celles exprimées par la première et qu'elles entraînent les conséquences suivantes:

$$\left\{ \begin{array}{l} x_i | y_j H = x_i | H \quad \text{quelque soit } y_j \\ y_j | x_i H = y_j | H \quad \text{quelque soit } x_i, \\ x_i y_j | H = x_i | H \times y_j | H \end{array} \right.$$

Pourtant, en effet, de la première condition

$$x_i | y_{ij} H = x_i | y_{ik} H \quad \text{qui fait l'écrire}$$

$$x_i | y_{ij} H = \frac{n_{ij}}{n_{\cdot j}} = \frac{n_{ik}}{n_{\cdot k}} = \dots = \frac{\sum n_{ij}}{\sum n_{\cdot j}} = \frac{n_{i\cdot}}{N} = x_i | H$$

On en déduit que la fréquence obtenue
 n_{ij} fait l'écrire

$$n_{ij} = n_{\cdot j} \times x_i | H = n_{\cdot j} \times \frac{n_{i\cdot}}{N}$$

Calculons, dans ces conditions, la fréquence
 $y_j | x_i H$. On trouve

$$y_j | x_i H = \frac{n_{ij}}{n_{i\cdot}} = \frac{n_{\cdot j} \times n_{i\cdot}}{n_{i\cdot} \times N} = \frac{n_{\cdot j}}{N} = y_j | H$$

qui prouve bien que la seconde condition est compatible avec la première.

On achèvera tout particulièrement de cette démonstration que lorsque deux variables x et y sont indépendantes, leur fréquence globale est égale au produit des fréquences marginales correspondantes.

$$x_i | y_j | H = x_i | H \times y_j | H$$

Exemple de variables indépendantes

Considérons l'ensemble de nombres entiers de 01 à 99, auquel nous ajoutons le nombre 00. Appelons x le chiffre des dizaines et y celui des unités.

Dans cet exemple :

a) les fréquences globales de tous les

numbers x et y sont toutes égales à 0,1

II - 9

$$x_i y_j | H = 0,01 \quad \text{quel que soit } x_i \text{ et } y_j$$

b) les fréquences marginales des x et des y sont toutes égales à 0,1

$$\begin{cases} x_i | H = 0,1 & \text{quel que soit } x_i \\ y_j | H = 0,1 & \text{quel que soit } y_j \end{cases}$$

c) les fréquences liées des x par le y ou des y par les x sont elles aussi toutes égales à 0,1. Pour vérifier, il suffit d'obtenir que la fréquence d'occurrence d'un chiffre quelconque de unités, le par exemple, est le même dans toutes les dizaines, et réciproquement. Les deux conditions

$$\begin{cases} x_i | y_j | H = x_i | H \\ y_j | x_i | H = y_j | H \end{cases}$$

se trouvent vérifiées. Les variables x et y sont indépendantes. On constate bien que la fréquence est égale à 0,1 pour tous les valeurs.

$$\begin{cases} x_i y_j | H = x_i | H \times y_j | H & \text{soit ici} \\ 0,01 = 0,1 \times 0,1 \end{cases}$$

II - 10 Modèle mathématique.

Tout comme pour les distributions à une seule variable, on s'est efforcé de représenter les distributions à deux caractères par des modèles mathématiques, l'ajustant le mieux possible sur les distributions observées. Nous nous limitons, dans notre exposé, à l'étude du cas où les variables x et y sont continues, en supposant de plus qu'à des distributions peuvent être définies en trois points par la densité (fréq.)

Dans ces conditions la fréquence dF des individus dont les mesures sont comprises à l'intérieur du parallélélogramme élémentaire $x, x+dx$ et $y, y+dy$ est évidemment proportionnelle à l'aire de ce parallélélogramme élémentaire $dx dy$ de ce paralléléogramme élémentaire aussi qui a la densité $f(x,y)$ de la distribution au point:

La fréquence globale $xy/H = dF$ que, par analogie avec la notation précédente, nous pourrons enfin désigner par le symbole xy/H a donc pour valeur.

$$xy/H = dF = f(x,y) dx dy$$

On peut obtenir une représentation visuelle de la distribution en construisant (fig II - 5) un modèle de la surface $Z = f(x,y)$ représentant les variations de la densité $f(x,y)$ de la distribution. La fréquence globale xy/H est alors représentée par le volume du prisme élémentaire élévé à l'apogée du paralléléogramme élémentaire $dx dy$ et limité à sa partie supérieure par la surface $Z = f(x,y)$. Le volume total compris entre le plan $Z = 0$ et la surface Z représente le somme des fréquences de Z tous les individus appartenant à la population, somme qui est évidemment égale à l'unité:

$$\iint_{-\infty}^{+\infty} f(x,y) dx dy = 1$$

II - 11 Fréquences marginales

Par définition la fréquence marginale $x/H = f(x) dx$ correspond à la proportion d'individus dont la mesure x est comprise dans l'intervalle $x, x+dx$, quelle que soit leur mesure y .

Si l'on se reporte à la fig. (II - 5) on voit que la fréquence marginale x/H est représentée par le volume de la tronche oblique dans la surface $Z = f(x,y)$ par les deux plans verticaux $x = \text{cste}$ et $x+dx = \text{cste}$, parallèles au plan $3oy$, soit:

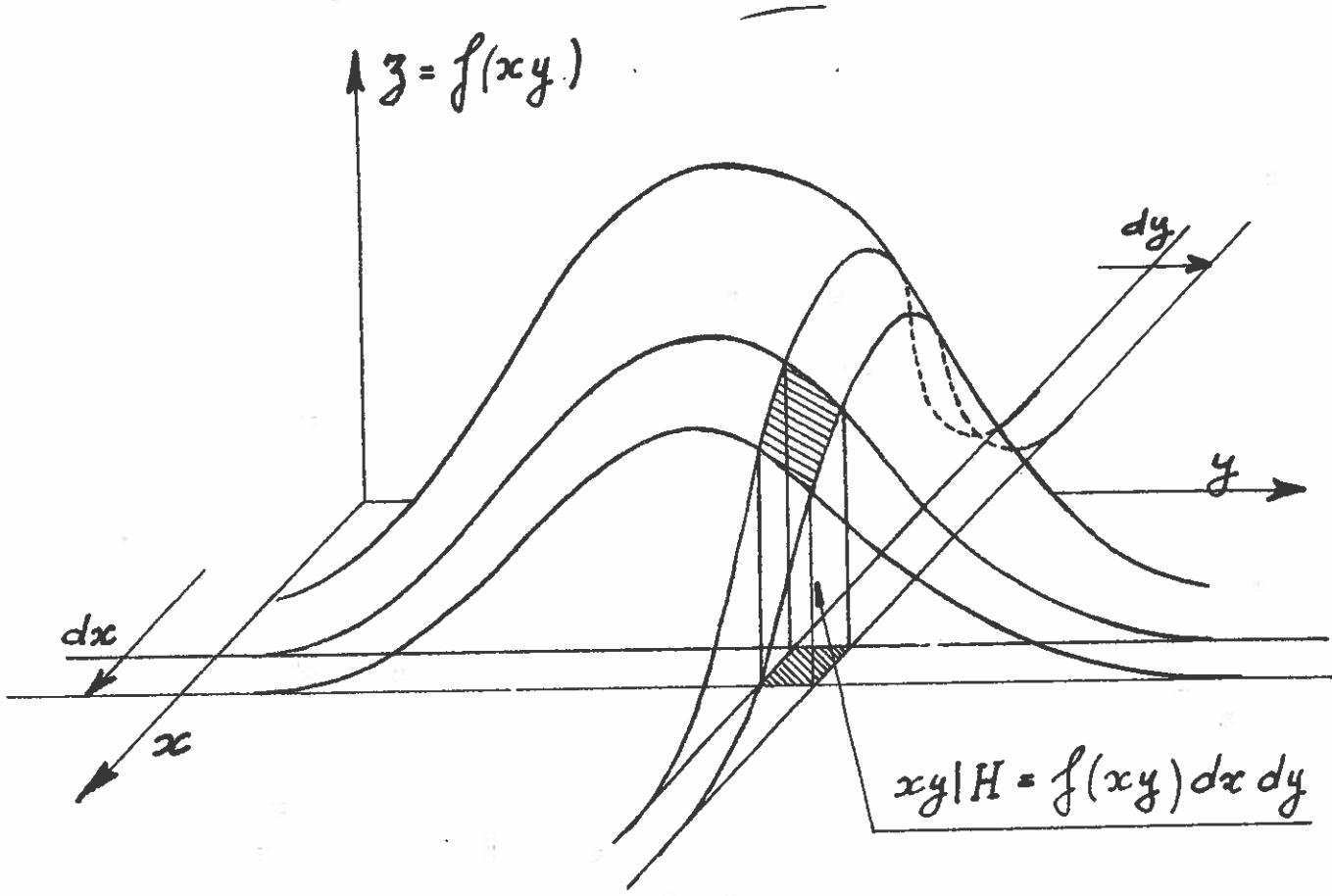


Fig: II - 5

Fréquence globale $xy | H = f(x,y) dx dy$

Fréquences marginales

$$\left\{ \begin{array}{l} x | H = f_1(x) dx = dx \int_{-\infty}^{+\infty} f(x,y) dy \\ y | H = f_2(y) dy = dy \int_{-\infty}^{+\infty} f(x,y) dx \end{array} \right.$$

Fréquences liées

$$\left\{ \begin{array}{l} x | y H = h_1(x,y) dx = \frac{f(x,y) dx}{\int_{-\infty}^{+\infty} f(x,y) dx} \\ y | x H = h_2(y,x) dy = \frac{f(x,y) dy}{\int_{-\infty}^{+\infty} f(x,y) dy} \end{array} \right.$$

$$x|H = f_1(x) dx = dx \int_{-\infty}^{+\infty} f(xy) dy$$

On définirait de même la seconde fréquence marginale par la relation:

$$y|H = f_2(y) dy = dy \int_{-\infty}^{+\infty} f(xy) dx$$

II-12. Fréquences liées

Lorsque l'on se fixe la valeur de la variable x , qui doit rester comprise dans l'intervalle $x, x+dx$, on définit un sous-ensemble dont la distribution correspond à la tranche découpée dans la surface $Z=f(xy)$ par les plans $x=cte$ et $x+dx=cte$.

Dans ce sous ensemble, la fréquence $y|xH$ des individus dont la mesure y est comprise dans l'intervalle $y+dy$ est donnée par le quotient du volume du parallélépipède élémentaire par le volume de la tranche:

$$y|xH = h_2(y|x) dy = \frac{\text{volume du parallélépipède}}{\text{volume de la tranche}} = \frac{f(x,y) dx dy}{dx \int_{-\infty}^{+\infty} f(xy) dy}$$

Soit, en définitive

$$y|xH = h_2(y,x) dy = \frac{f(xy) dy}{\int_{-\infty}^{+\infty} f(x,y) dy}$$

On définirait de même la seconde distribution liée par la relation:

$$x | y H = h_1(x, y) dx = \frac{\int f(x, y) dx}{\int_{-\infty}^{+\infty} f(x, y) dx}$$

On montre facilement que les densités de distribution jointe, marginales et conditionnelles vérifient les relations :

$$\begin{cases} f(x, y) = f_1(x) \times h_2(y, x) \\ f(x, y) = f_2(y) \times h_1(x, y) \end{cases}$$

qui nous font trouver la formule établie au paragraphe II-8

II-13. Équations des combes de régression

Si l'on se reporte aux définitions données au paragraphe II-3, on constate que l'ordonnée $y(x)$ de la courbe de régression de y en x n'est autre que la valeur moyenne de la variable y lorsque l'on s'est fixé x .

En d'autres termes, l'ordonnée $y(x)$ est égale à l'espérance mathématique de la variable y_1 , calculée dans la distribution de y liée par x , équation qui permet d'écrire immédiatement l'équation de la courbe de régression de y en x .

$$y(x) = \int_{-\infty}^{+\infty} y h_2(y, x) dy = \frac{\int_{-\infty}^{+\infty} y f(x, y) dy}{\int_{-\infty}^{+\infty} f(x, y) dy}$$

On trouverait de même, pour la seconde courbe de régression

$$x(y) = \int_{-\infty}^{+\infty} x h_1(x, y) dx = \frac{\int_{-\infty}^{+\infty} x f(x, y) dx}{\int_{-\infty}^{+\infty} f(x, y) dx}$$

II - 14. Variables indépendantes

La relation $x_i, y_j | H = x_i | H \times y_j | H$, établie au paragraphe II-9, montre que, lorsque les variables x et y sont indépendantes, la densité $f(x,y)$ de la distribution prend la forme d'un produit d'une fonction de x par une fonction de y

$$f(x,y) = f_1(x) \times f_2(y)$$

Les distributions liées sont alors égales aux distributions marginales correspondantes et les combes de régression sont des droites parallèles aux axes de coordonnées.

La réciproque n'est pas obligatoirement vraie, car combes de régression pouvant être des droites parallèles aux axes, sans pour cela que les variables x et y soient indépendantes.

II - 15. Espérances mathématiques des variables x et y

Une distribution se trouve en général groupée au voisinage de son centre de gravité, dont les coordonnées $E[x]$ et $E[y]$ sont les espérances mathématiques des variables x et y .

$$\left\{ \begin{array}{l} E[x] = \iint_{-\infty}^{+\infty} x f(x,y) dx dy \\ E[y] = \iint_{-\infty}^{+\infty} y f(x,y) dx dy \end{array} \right.$$

Notons que les termes $E[x]$ et $E[y]$ sont également les espérances mathématiques des distributions marginales correspondantes. On peut, en effet, écrire

$$\left\{ \begin{array}{l} E[x] = \int_{-\infty}^{+\infty} x dx \int_{-\infty}^{+\infty} f(x,y) dy = \int_{-\infty}^{+\infty} x f_1(x) dx \\ E[y] = \int_{-\infty}^{+\infty} y dy \int_{-\infty}^{+\infty} f(x,y) dx = \int_{-\infty}^{+\infty} y f_2(y) dy \end{array} \right.$$

II - 16. Variables centrées.

II - 15

Si on rapporte une distribution à son centre de gravité, les nouvelles variables ξ et η obtenues sont appelées "variables centrées".

$$\begin{cases} \xi = x - E[x] \\ \eta = y - E[y] \end{cases}$$

II - 17 Moments d'une distribution

On appelle moment d'ordre q qui appartient à la variable x et d'ordre r qui appartient à la variable y la quantité m_{qr} définie comme l'espérance mathématique du produit $x^q y^r$:

$$m_{qr} = \iint_{-\infty}^{+\infty} x^q y^r f(x,y) dx dy$$

Notons que les moments des deux distributions marginales correspondent respectivement aux moments m_{q0} (marginal en x) et m_{0r} (marginal en y).

II - 18 Moments centrés.

Les moments centrés μ_{qr} de la distribution sont les moments des variables centrées ξ et η .

$$\mu_{qr} = E[\xi^q \eta^r] = \iint_{-\infty}^{+\infty} (x - E[x])^q (y - E[y])^r f(x,y) dx dy$$

II - 19 Moments centrés d'ordre 2

Les moments centrés d'ordre 2 sont au nombre de trois, qui portent les dénominations particulières suivantes:

variance x : $\mu_{20} = E[\bar{x}^2] = \iint_{-\infty}^{+\infty} (x - E[x])^2 f(x,y) dx dy$

variance y : $\mu_{02} = E[\bar{y}^2] = \iint_{-\infty}^{+\infty} (y - E[y])^2 f(x,y) dx dy$

covariance xy : $\mu_{11} = E[\bar{x}\bar{y}] = \iint_{-\infty}^{+\infty} (x - E[x])(y - E[y]) f(x,y) dx dy$

On pose également

$$\begin{cases} \mu_{20} = \sigma_x^2 \\ \mu_{02} = \sigma_y^2 \end{cases} \quad \mu_{11} = \sigma_{xy} = \sigma_{yx} \Rightarrow \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix}$$

définissant ainsi les écarts-types théoriques des variables x et y .
Je fais l'hypothèse que les variances sont des quantités essentiellement positives, alors que la covariance peut être positive, négative ou nulle.

II-20 Covariance de deux variables indépendantes

Supposons les variables x et y sont statistiquement indépendantes, leur covariance est nulle. En effet, la variance de la distribution pouvant se mettre sous la forme $f_1(x) \times f_2(y)$, la covariance μ_{11} s'exprime par l'intégrale

$$\mu_{11} = \iint_{-\infty}^{+\infty} (x - E[x])(y - E[y]) f_1(x) dx f_2(y) dy \text{ soit encore}$$

$$\mu_{11} = \int_{-\infty}^{+\infty} (x - E[x]) f_1(x) dx \times \int_{-\infty}^{+\infty} (y - E[y]) f_2(y) dy = 0 \times 0$$

La réciproque n'est pas vraie car le fait que la covariance soit nulle n'entraîne pas obligatoirement que les variables x et y soient indépendantes.

II - 21 Coefficient de corrélation.

On s'est efforcé de définir un coefficient numérique sans dimension qui, dans une certaine mesure, puisse servir à chiffrer l'intensité de la liaison des variables x et y , ce coefficient devrait être nul lorsque les variables sont indépendantes et varier de $+1$ si il existe une liaison fonctionnelle entre les variables x et y .

Le coefficient ρ choisi, appelé coefficient de corrélation, est défini par la relation:

$$\rho = \frac{\text{cov. } xy}{\sqrt{\text{var } x \cdot \text{var } y}} = \frac{\mu_{11}}{\sigma_x \cdot \sigma_y}$$

Nous avons démontré (paragraphe II-20) que la covariance μ_{11} de deux variables indépendantes était nulle. Il en est donc de même du coefficient de corrélation, sous que la réciproque soit également vraie.

Il nous reste à calculer dans quel intervalle peut varier le coefficient de corrélation. Nous calculerons, dans ce but, l'expérience mathématique du cas d'une fonction linéaire des variables centées \bar{x} et \bar{y} soit:

$$\begin{aligned} E\left[\left(\frac{\bar{x}}{\sigma_x} - \rho \frac{\bar{y}}{\sigma_y}\right)^2\right] &= E\left[\left(\frac{\bar{x}^2}{\sigma_x^2} - 2\rho \frac{\bar{x}\bar{y}}{\sigma_x \sigma_y} + \rho^2 \frac{\bar{y}^2}{\sigma_y^2}\right)\right] \\ &= \frac{\sigma_x^2}{\sigma_x^2} - 2\rho \frac{\text{cov}(xy)}{\sigma_x \sigma_y} + \rho^2 \frac{\sigma_y^2}{\sigma_y^2} \\ &= 1 - 2\rho \cdot \rho + \rho^2 = 1 - \rho^2 \end{aligned}$$

$$E\left[\left(\frac{\bar{x}}{\sigma_x} - \rho \frac{\bar{y}}{\sigma_y}\right)^2\right] = 1 - \rho^2$$

Le premier membre de cette dernière relation
représente le biais moyen d'un caractère. Ce biais
moyenne ne peut être que positive ou nulle. Il
est nécessaire donc que

$$1 - \rho^2 \geq 0$$

surtout

$$-1 \leq \rho \leq +1$$

Mais enfin que le coefficient de corrélation ne
peut atteindre le valeur +1 qu'à la condition
que tous les caractères

$$\left(\frac{\bar{x}}{\sigma_x} - \frac{\bar{y}}{\sigma_y} \right)^2$$

sont nuls, c'est à dire que les variables
 x et y sont liées par la relation fonctionnelle
linéaire

$$y = E[y] + \frac{\sigma_y}{\sigma_x} (x - E[x])$$

On montrerait de même que $\rho = -1$
lorsque.

$$y = E[y] - \frac{\sigma_y}{\sigma_x} (x - E[x])$$

II - 22 Matrice des variances et covariances.

On peut grouper les moments centris d'une
distribution en une matrice carree symétrique
 V_{100} , appelée "matrice des variances", qui se présente
sous la forme :

$$[V] = \begin{bmatrix} \mu_{00} & \mu_{10} \\ \mu_{01} & \mu_{02} \end{bmatrix} = \begin{bmatrix} \sigma_x^2 & \rho \sigma_x \sigma_y \\ \rho \sigma_x \sigma_y & \sigma_y^2 \end{bmatrix}$$

dont le déterminant a pour valeur :

$$|V| = (1-\rho^2) \sigma_x^2 \sigma_y^2$$

La matrice des variances est donc régulière,
sauf pour $\rho = \pm 1$

La matrice inverse, lorsqu'on peut la définir,
a pour expression :

$$[V]^{-1} = \frac{1}{(1-\rho^2)} \begin{bmatrix} \frac{1}{\sigma_x^2} & -\frac{\rho}{\sigma_x \sigma_y} \\ -\frac{\rho}{\sigma_x \sigma_y} & \frac{1}{\sigma_y^2} \end{bmatrix}$$

II - 23 Ellipse indicatrice

Nous supposons, dans ce paragraphe, que la distribution est rapportée à un système d'axes encliqueté c. les goniomètres $G\bar{x}, G\bar{y}$ passant par son centre de gravité et parallèle aux axes x, y .
 \bar{x} et \bar{y} sont donc les variables centrées obtenues au paragraphe II-16.

Considérons,
dans le plan \bar{x}, \bar{y} ,
un vecteur GU ,
de direction orthogonale
mais dont les composantes u et v vérifient
l'équation quadratique
suivante, dont les coefficients sont ceux
de la matrice V :

$$u^2 \sigma_x^2 + 2uv\rho \sigma_x \sigma_y + v^2 \sigma_y^2 = 1$$

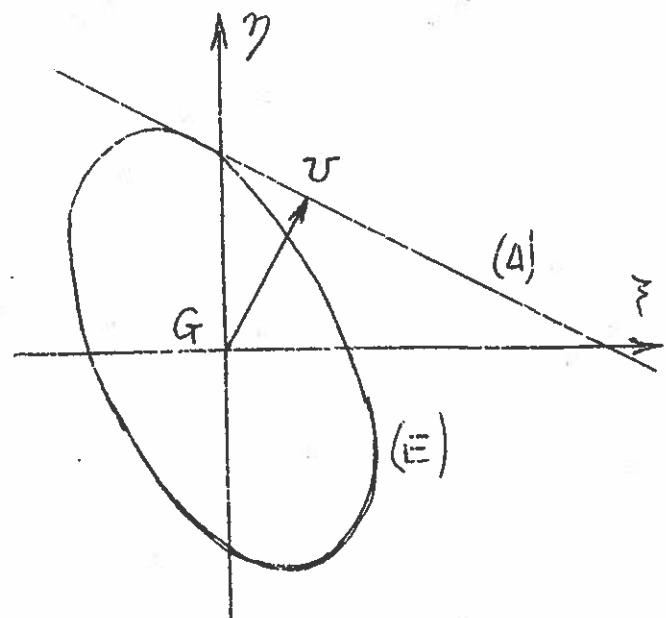


Fig II-6. Ellipse indicatrice

lorsque la direction du vecteur v varie,

la droite (A) d'équation $u\bar{x} + v\bar{y} = 1$ enroule une ellipse (E) qui est, par définition, l'ellipse indicatrice de la distribution. L'équation tangentielle de cette ellipse est donnée par la relation précédente.

Or, tant que son équation quadratique se déduit des coefficients de la matrice inverse $[V]^{-1}$. En résumé, l'ellipse indicatrice est définie par l'intersection des deux équations suivantes :

$$\left\{ \begin{array}{l} \sigma_x^2 u^2 + 2 \rho \sigma_x \sigma_y u v + \sigma_y^2 v^2 = 1 \quad (\text{forme tangentielle}) \\ \frac{1}{1-\rho^2} \left\{ \frac{\bar{x}^2}{\sigma_x^2} - 2\rho \frac{\bar{x}}{\sigma_x} \frac{\bar{y}}{\sigma_y} + \frac{\bar{y}^2}{\sigma_y^2} \right\} = 1 \quad (\text{forme quadratique}) \end{array} \right.$$

Il est facile de voir pourquoi l'ellipse indicatrice s'inscrit dans un rectangle (fig II-7) dont les côtés sont égaux à $2\sigma_x$ et $2\sigma_y$.

Les points de contact de l'ellipse avec le rectangle circulaire ont pour coordonnées

$$\left\{ \begin{array}{l} \pm \rho \sigma_x \\ \pm \sigma_y \end{array} \right.$$

$$\left\{ \begin{array}{l} \pm \sigma_x \\ \pm \rho \sigma_y \end{array} \right.$$

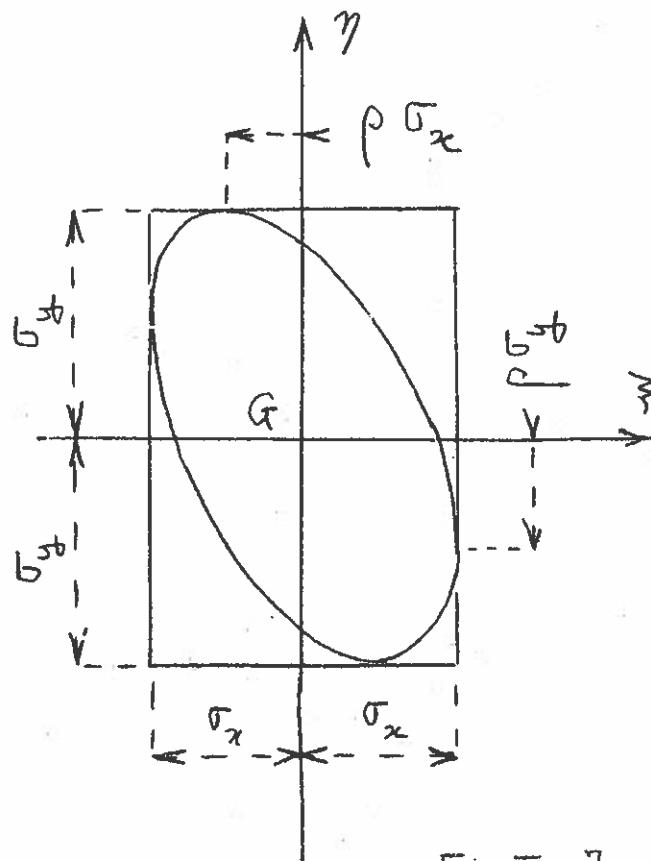


Fig II - 7

II - 24 Variabilités réduites

Tout comme dans le cas d'une distribution à un seul caractère, il est naturel de choisir les écart-types σ_x et σ_y comme unités de mesure des écarts x et y de variables x et y qui s'offre à leurs esperances mathématiques respectives. On obtient ainsi les variables réduites qui ont pour espérances:

$$\left\{ \begin{array}{l} \theta_1 = \frac{x}{\sigma_x} = \frac{x - E[x]}{\sigma_x} \\ \theta_2 = \frac{y}{\sigma_y} = \frac{y - E[y]}{\sigma_y} \end{array} \right.$$

Les variables réduites ont leurs espérances mathématiques nulles:

$$E[\theta_1] = 0 \quad E[\theta_2] = 0$$

Leur matrice des covariances a pour expression:

$$\begin{bmatrix} V_R \end{bmatrix} = \begin{bmatrix} 1 & p \\ p & 1 \end{bmatrix}$$

II - 25. Indicateur réducte

Dans le plan des variables réduites, l'indicateur a pour équation tangente à et perpendiculaire les expressions suivantes:

$$\left\{ \begin{array}{l} u^2 + 2puv + v^2 = 1 \\ \frac{1}{1-p^2} \left\{ \theta_1^2 - 2p\theta_1\theta_2 + \theta_2^2 \right\} = 1 \end{array} \right.$$

On peut joindre et montrer qu'en
conformité cartésienne, les demi-axes de
l'indicatrice réduite ont pour longueur.

$$\left\{ \begin{array}{l} \frac{1}{2} \text{ Grand axe} : \sqrt{1+\rho} \\ \frac{1}{2} \text{ Petit axe} : \sqrt{1-\rho} \end{array} \right.$$

Lorsque le coefficient de corrélation ρ
varie, l'ellipse indicatrice se déforme comme suit
à ρ_0 [fig. II-8]. En particulier, pour $\rho = \pm 1$, l'ellipse
se confond avec les diagonales du carré concerné.

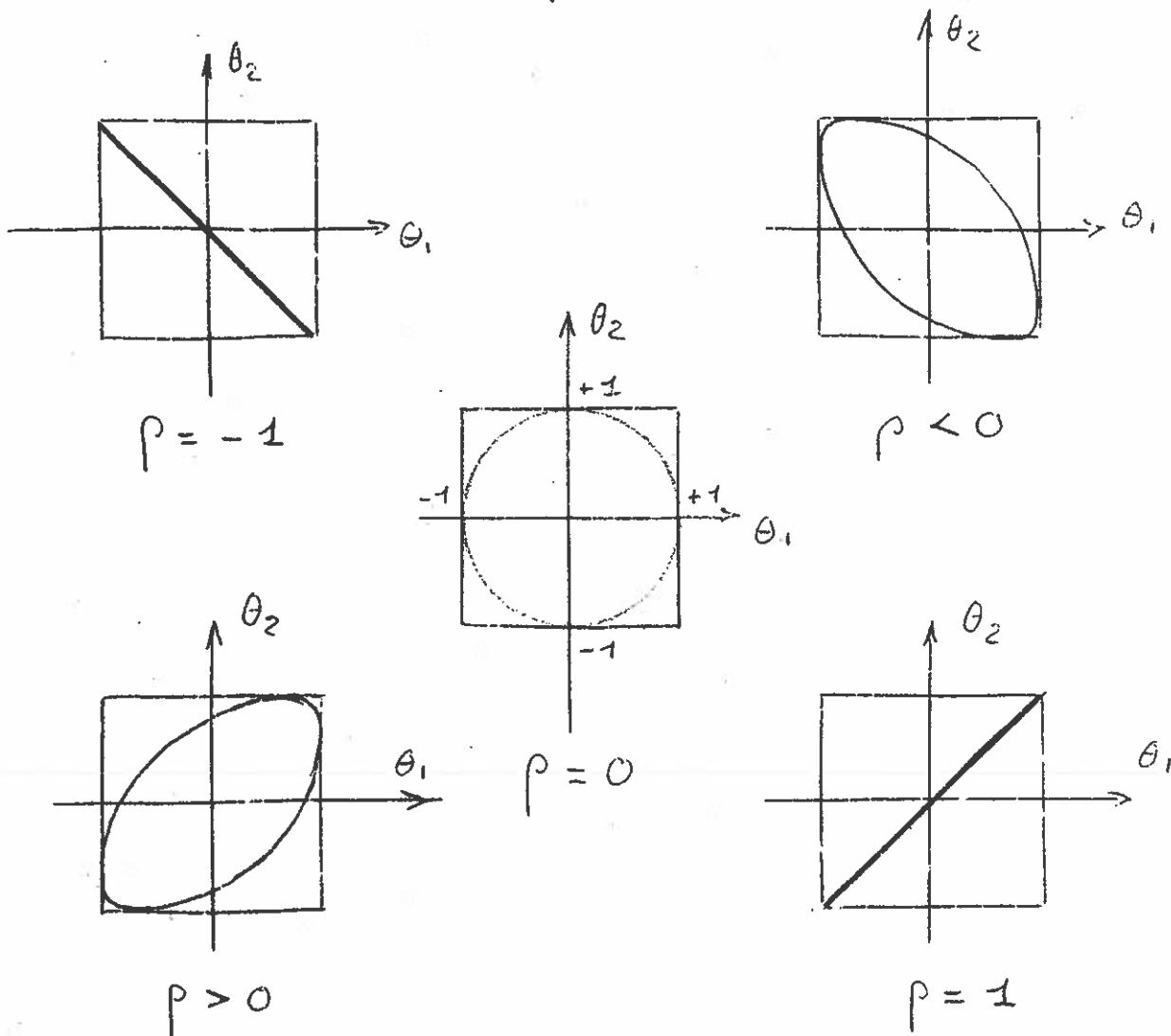


Fig II-8. Variations de forme de
l'ellipse indicatrice réduite

II - 28 Ellipses de concentration. Généralisation de l'inégalité de Bienaymé-Tchebychev

II-23

χ étant un paramètre pouvant prendre des valeurs positives quelconques, considérons la famille d'ellipses représentées par l'équation générale :

$$\frac{1}{1-p^2} \left\{ \frac{\xi^2}{\sigma_x^2} - 2p \frac{\xi}{\sigma_x} \frac{\eta}{\sigma_y} + \frac{\eta^2}{\sigma_y^2} \right\} = \chi^2$$

Ces ellipses, homothétiques à l'ellipse indicative, sont appelées "Ellipses de concentration"

Considérons l'une d'entre elles, correspondant à la valeur t du paramètre χ et définie par l'équation

$$\frac{1}{1-p^2} \left\{ \frac{\xi^2}{\sigma_x^2} - 2p \frac{\xi}{\sigma_x} \frac{\eta}{\sigma_y} + \frac{\eta^2}{\sigma_y^2} \right\} = t^2$$

Cette ellipse divise le plan xoy en deux domaines (fig II-9)

{ l'un \bar{I} intérieur à l'ellipse

{ l'autre \bar{I} extérieur à l'ellipse

Si nous proposons de démontrer que, quelle que soit la loi de distribution $f(x,y) dx dy$, il existe

une proportion au moins égale à $1 - \frac{2}{t^2}$

d'individus dont les mesures x,y correspondent à des points situés dans le domaine hachuré \bar{I} :

$$\iint_{\bar{I}} f(x,y) dx dy \geq 1 - \frac{2}{t^2}$$

Inégalité qui généralise la formule de Bienaymé-Tchebychev.

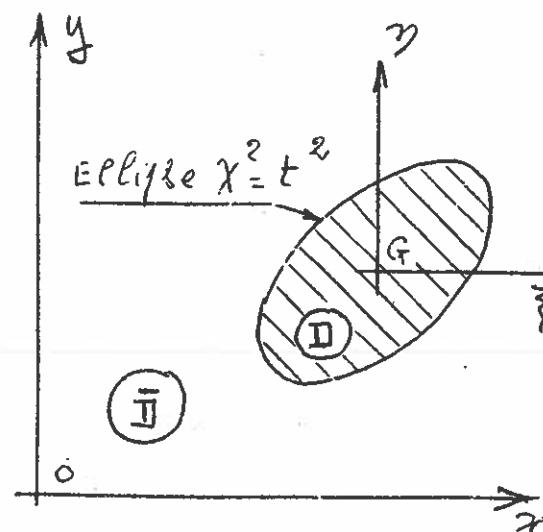


Fig II-9

La démonstration comprend deux parties

II - 24

10) On calculera, en premier lieu, l'espérance mathématique de χ^2

$$E[\chi^2] = \frac{1}{1-p^2} E\left\{\frac{z^2}{\sigma_x^2} - 2p \frac{z}{\sigma_x} \frac{y}{\sigma_y} + \frac{y^2}{\sigma_y^2}\right\} = \frac{1}{1-p^2} \left\{1 - 2p + p + 1\right\}$$

soit :

$$E[\chi^2] = 2$$

2^e) L'intégrale de définition de $E[\chi^2]$, que l'on étend à tout le plan \mathbb{R}^2 , peut être séparée en deux parties, dont l'une s'étend sur le domaine $\bar{\Omega}$ et l'autre au domaine complémentaire $\bar{\Omega}$. Dans ces conditions nous pouvons écrire :

$$E[\chi^2] = \iint_{-\infty}^{+\infty} \chi^2 f(x,y) dx dy = \iint_{\bar{\Omega}} \chi^2 f(x,y) dx dy + \iint_{\bar{\Omega}} \chi^2 f(x,y) dx dy$$

L'intégrale étendue au domaine $\bar{\Omega}$ ne pouvant être que positive ou nulle, on a écrit d'écrire :

$$E[\chi^2] \geq \iint_{\bar{\Omega}} \chi^2 f(x,y) dx dy$$

Le domaine d'intégration est maintenant réduit au domaine $\bar{\Omega}$, dans lequel les valeurs de χ^2 sont toujours supérieures à t^2 . On a déduit que :

$$\iint_{\bar{\Omega}} \chi^2 f(x,y) dx dy \geq \iint_{\bar{\Omega}} t^2 f(x,y) dx dy = t^2 \iint_{\bar{\Omega}} f(x,y) dx dy$$

Tenant compte du fait que $E[\chi^2] = 2$, nous avons donc démontré que :

$$2 \geq t^2 \iint_{\bar{\Omega}} f(x,y) dx dy = t^2 \left\{1 - \iint_{\bar{\Omega}} f(x,y) dx dy\right\}$$

soit finalement que

$$\iint_{\mathbb{D}} f(x,y) dx dy \geq 1 - \frac{2}{t^2}$$

Les échelles de concentration s'ajustent si plus en plus lorsque le coefficient de corrélation se rapproche de ± 1 , ce qui renoue que le coefficient de corrélation constitue bien un indice valable de l'intensité de la liaison entre les variables x et y .

II-27 Droites des moindres carrés.

Par un point M quelconque de la distribution (fig II-10), menons une parallèle à l'axe ou parallèle qui rencontre en un point P_2 une droite arbitraire (A_2) .

Nous nous proposons de rechercher s'il est possible de choisir la droite (A_2) de manière à rendre minimum l'espérance mathématique Z_2 des carrés des longueurs des segments MP_2 , soit la quantité :

$$Z_2 = E[\overline{MP}_2^2]$$

Afin de simplifier les calculs, nous écrivons l'équation de la droite (A_2) sous la forme :

$$Y - E[y] = \alpha_2 + \beta_2(x - E[x])$$

équation dans laquelle les paramètres α_2 et β_2 doivent être déterminés de manière à rendre Z_2 minimum.

Le segment \overline{MP}_2 a pour longueur

$$\overline{MP}_2 = Y - y = \alpha_2 + \beta_2(x - E[x]) - (y - E[y])$$

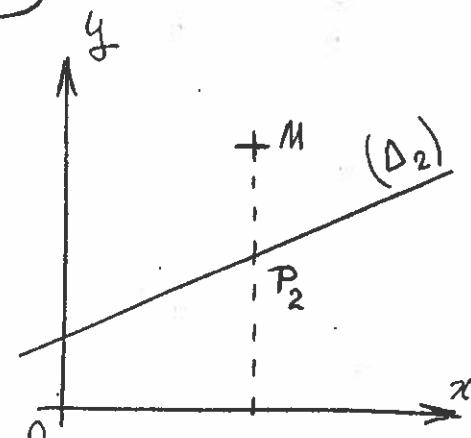


Fig II-10

soit encore, en utilisant les variables centrées

II-26

$$\xi = x - E[x] \quad \eta = y - E[y]$$

$$\overline{MP}_2 = \alpha_2 + \beta_2 \xi - \eta$$

$$\text{d'où} \quad Z_2 = E[(\alpha_2 + \beta_2 \xi - \eta)^2]$$

Pour déterminer les valeurs des paramètres α_2 et β_2 rendant Z_2 maximum, il nous suffit d'annuler les dérivées partielles de Z_2 par rapport à α_2 et β_2 :

$$\frac{\partial Z_2}{\partial \alpha_2} = E[2(\alpha_2 + \beta_2 \xi - \eta)] = 2\alpha_2 = 0$$

$$\frac{\partial Z_2}{\partial \beta_2} = E[2\xi(\alpha_2 + \beta_2 \xi - \eta)] = 2\left\{\beta_2 \sigma_x^2 - \rho \sigma_x \sigma_y\right\} = 0$$

d'où l'on déduit:

$$\begin{cases} \alpha_2 = 0 \\ \beta_2 = \rho \cdot \frac{\sigma_y}{\sigma_x} \end{cases}$$

La droite (D_2) , appelée droite des moindres carrés de y en x , a donc pour équation:

$$(D_2) \quad \boxed{Y - E[y] = \rho \frac{\sigma_y}{\sigma_x} (x - E[x])}$$

Si l'on se reporte à l'ellipse indicatrice (paragraphe II-23), on vérifiera facilement que la droite (D_2) est un diamètre conjugué de la direction de l'axe oy (fig II-11).

On peut enfin calculer la valeur minimum de Z_2 , correspondant à la droite des moindres carrés (D_2) :

$$(Z_2)_{\min} = E\left[\left(\rho \frac{\sigma_y}{\sigma_x} \xi - \eta\right)^2\right] = E\left[\rho^2 \frac{\sigma_y^2}{\sigma_x^2} \xi^2 - 2\rho \frac{\sigma_y}{\sigma_x} \xi \eta + \eta^2\right]$$

$$= \rho^2 \frac{\sigma_y^2}{\sigma_x^2} \sigma_x^2 - 2\rho \frac{\sigma_y}{\sigma_x} \times \rho \sigma_x \sigma_y + \sigma_y^2$$

$$\boxed{(Z_2)_{\min} = (1 - \rho^2) \sigma_y^2}$$

On définit de même la droite (D_1) des moindres carrés de x en y , pour laquelle (Fig II-11) est minimum l'espérance mathématique Z_1 des carrés des distances MP_1^2 , comptées parallèlement à l'axe des x . Cette droite (D_1) a pour équation :

$$(D_1) \quad X - E[x] = \rho \frac{\sigma_x}{\sigma_y} (y - E[y])$$

(D_1) est un diamètre de l'ellipse indicatrice conjuguée de la direction de l'axe Ox . Z_1 a pour valeur minimum :

$$(Z_1)_{\min} = (1 - \rho^2) \sigma_x^2$$

Remarque Lorsque les courbes de régression (paragraphes II-3 et II-13) d'une distribution sont rectilignes, elles coïncident alors avec les droites des moindres carrés correspondantes. Ce cas particulier ne justifie pas l'emploi fréquent du terme "droites de régression" pour ces droites des moindres carrés lorsque les courbes de régression ne sont pas rectilignes.

Les coefficients β_1 et β_2 , qui ont pour valeurs :

$$\left\{ \begin{array}{l} \beta_1 = \rho \frac{\sigma_x}{\sigma_y} \\ \beta_2 = \rho \frac{\sigma_y}{\sigma_x} \end{array} \right.$$

sont parfois appelés "coefficients de régression". Ils vérifient la relation :

$$\beta_1 \beta_2 = \rho^2$$

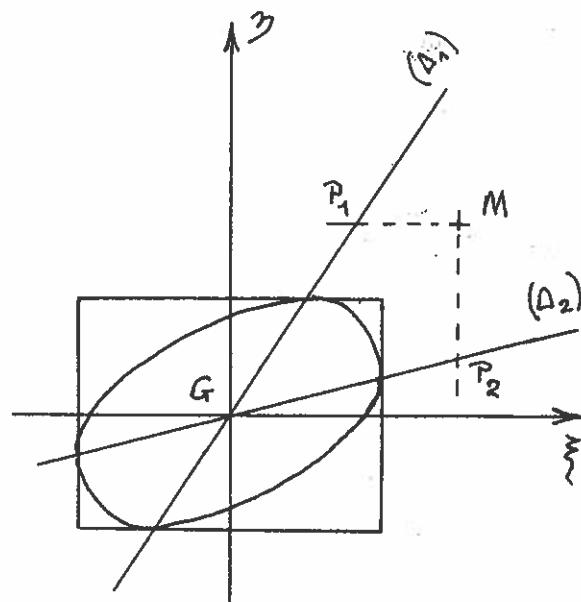


Fig II-11

Dans toute la suite de ce chapitre, nous abandonnons la désignation ox_1, ox_2 des axes pour adopter la notation ox_1, ox_2 qui permet de mieux mettre en évidence la symétrie des formules.

Soient donc x_1 et x_2 les coordonnées d'un point M de la distribution, supposée représentée dans un espace euclidien à deux dimensions. Cet espace est rapporté (fig II-12) à des axes obliques quelconques ox_1, ox_2 , dont les vecteurs unité \vec{e}_1 et \vec{e}_2 peuvent avoir des longueurs e_1 et e_2 différentes.

A tout point M de la distribution, on peut associer un vecteur $\vec{x} = \overrightarrow{OM}$ dont les projections, selon un contour projectif parallèle aux axes, sont les vecteurs $\overrightarrow{OA_1}$ et $\overrightarrow{OA_2}$. Les composantes contravariantes du vecteur \overrightarrow{OM} ont donc pour valeurs :

$$\begin{cases} x_1 = \frac{\overrightarrow{OA_1}}{e_1} \\ x_2 = \frac{\overrightarrow{OA_2}}{e_2} \end{cases}$$

Choisissons arbitrairement dans le plan x_1, ox_2 un vecteur \vec{u} , défini par ses composantes covariantes u_1 et u_2 . Nous rappelons que, dans un espace métrique, les composantes covariantes du vecteur \vec{u} , proportionnelles à la longueur de ses projections orthogonales $\overrightarrow{OP_1}$ et $\overrightarrow{OP_2}$ sur les axes, ont pour valeurs :

$$\begin{cases} u_1 = \overrightarrow{OP_1} \times e_1 \\ u_2 = \overrightarrow{OP_2} \times e_2 \end{cases}$$

Dans ces conditions, on sait que le produit scalaire $\vec{u} \cdot \vec{x}$ des vecteurs \vec{u} et \vec{x} a pour expression :

$$\vec{u} \cdot \vec{x} = u_1 x_1 + u_2 x_2$$

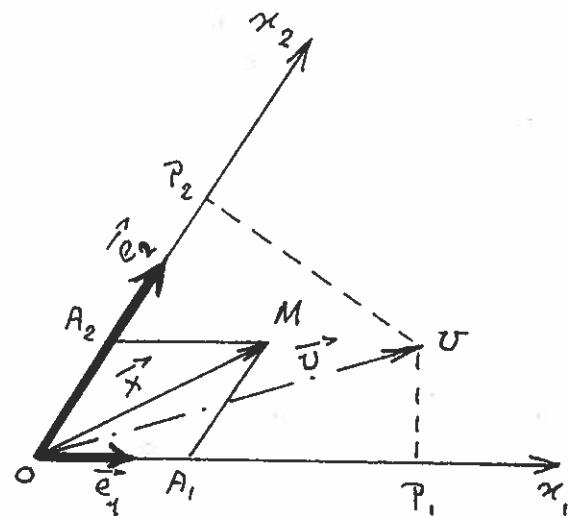


Fig II-12

Par définition la fonction caractéristique de la distribution du vecteur \vec{x} est une fonction du vecteur \vec{u} qui a pour expression:

$$\Phi_{\vec{x}}(u_1, u_2) = E[e^{i \vec{u} \cdot \vec{x}}] = E[e^{i(u_1 x_1 + u_2 x_2)}]$$

La fonction caractéristique étant la moyenne arithmétique des quantités scalaires $e^{i \vec{u} \cdot \vec{x}}$ est, elle-même, un scalaire dont la valeur est invariante pour tout changement de système de référence.

II-29 Fonction caractéristique des distributions marginales.

Si l'on choisit le vecteur \vec{u} perpendiculaire à l'axe ox_2 , ses composantes covariantes auront pour valeurs:

$$\begin{cases} u_1 = \overline{OP_1} \times e_1 \neq 0 \\ u_2 = 0 \end{cases}$$

et par suite le produit scalaire $\vec{u} \cdot \vec{x}$ se réduira à:

$$\vec{u} \cdot \vec{x} = u_1 x_1$$

Dans ces conditions la fonction caractéristique a pour valeur:

$$\Phi_{\vec{x}}(u_1, 0) = E[e^{iu_1 x_1}] = \iint_{-\infty}^{+\infty} e^{iu_1 x_1} f(x_1, x_2) dx_1 dx_2$$

formule que l'on peut encore écrire:

$$\Phi_{\vec{x}}(u_1, 0) = \int_{-\infty}^{+\infty} e^{iu_1 x_1} dx_1 \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 = \int_{-\infty}^{+\infty} e^{iu_1 x_1} f_1(x_1) dx_1$$

expression dans laquelle (cf. par. II-11) $f_1(x_1) dx_1$ représente la distribution marginale des x_1 .

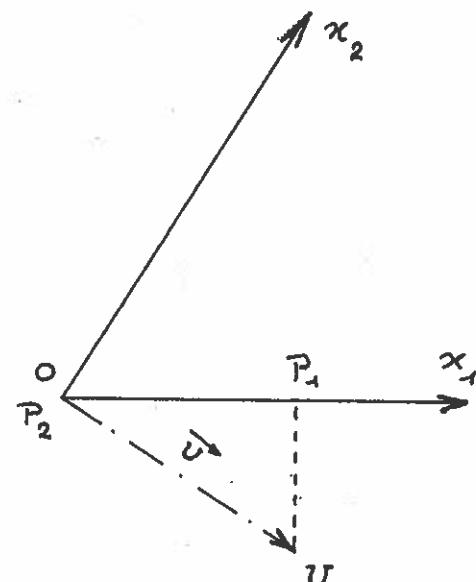


Fig II - 13

On démontre de même que

II - 30

$$\Phi_{\vec{x}}(0, u_2) = \int_{-\infty}^{+\infty} e^{iu_2 x_2} f_2(x_2) dx_2$$

expression dans laquelle $f_2(x_2) dx_2$ représente la distribution marginale des x_2 .

En résumé, si l'on désigne par $\Phi_{x_1}(u_1)$ et $\Phi_{x_2}(u_2)$ les fonctions caractéristiques des distributions marginales en x_1 et x_2 , on a :

$$\left\{ \begin{array}{l} \Phi_{x_1}(u_1) = \Phi_{\vec{x}}(u_1, 0) \\ \Phi_{x_2}(u_2) = \Phi_{\vec{x}}(0, u_2) \end{array} \right.$$

II.30 Expression de la fonction caractéristique lorsque les variables x_1 et x_2 sont indépendantes.

Lorsque les variables x_1 et x_2 sont indépendantes, on sait (cf. II - 14) que la densité de la distribution peut se mettre sous la forme :

$$f(x_1, x_2) = f_1(x_1) \times f_2(x_2)$$

$f_1(x_1)$ et $f_2(x_2)$ étant d'ailleurs les densités des distributions marginales de x_1 et x_2 . Dans ces conditions la fonction caractéristique a pour expression :

$$\begin{aligned} \Phi_{\vec{x}}(u_1, u_2) &= \iint_{-\infty}^{+\infty} e^{i(u_1 x_1 + u_2 x_2)} f_1(x_1) f_2(x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{+\infty} e^{iu_1 x_1} f_1(x_1) dx_1 \times \int_{-\infty}^{+\infty} e^{iu_2 x_2} f_2(x_2) dx_2 \end{aligned}$$

soit encore

$$\Phi_{\vec{x}}(u_1, u_2) = \Phi_{x_1}(u_1) \times \Phi_{x_2}(u_2)$$

Lorsque les variables sont indépendantes, la fonction caractéristique de la distribution globale est égale au produit des fonctions caractéristiques des distributions marginales et réciproquement.

II - 31. Changement de système de référence

Nous nous proposons de rechercher l'expression de la fonction caractéristique $\Phi_{\vec{x}}$ de la distribution globale

lorsque l'on rapporte les coordonnées du point M à un nouveau système de vecteurs de base \vec{e}_1' et \vec{e}_2' définis par les relations:

$$\left\{ \begin{array}{l} \vec{e}_1' = \alpha_{11} \vec{e}_1 + \alpha_{12} \vec{e}_2 \\ \vec{e}_2' = \alpha_{21} \vec{e}_1 + \alpha_{22} \vec{e}_2 \end{array} \right.$$

soit encore, en notation matricielle:

$$[\vec{e}'] = [\alpha] [\vec{e}]$$

on désignera par:

- $[\beta] = [\alpha]^{-1}$ la matrice inverse de la matrice $[\alpha]$
- $[\alpha_T]$ la matrice transposée de la matrice $[\alpha]$
- $[\beta_T]$ la matrice transposée de la matrice $[\beta]$

Dans ces conditions on sait que:

- 1) Les nouvelles composantes x_1' et x_2' du vecteur contravariant $\vec{x} = \vec{OM}$ sont données par la relation:

$$[\vec{x}'] = [\beta_T] [\vec{x}]$$

soit encore:

$$\left\{ \begin{array}{l} x_1' = \beta_{11} x_1 + \beta_{21} x_2 \\ x_2' = \beta_{12} x_1 + \beta_{22} x_2 \end{array} \right.$$

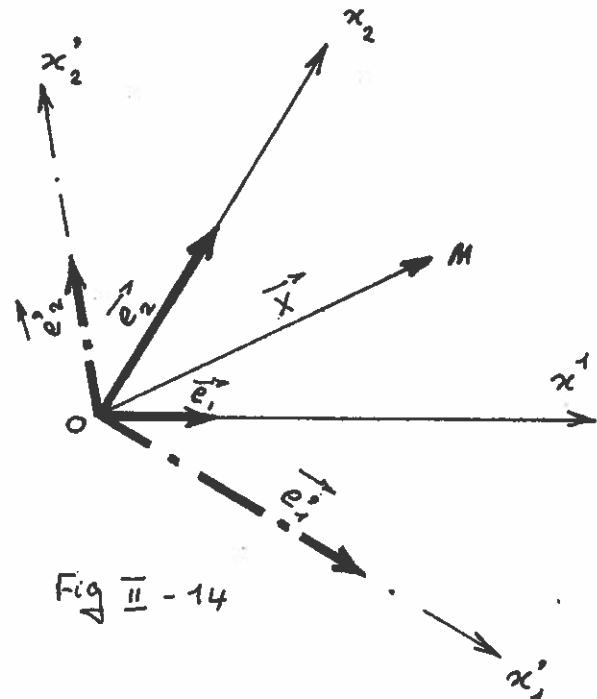


Fig II - 14

2°) Les anciennes composantes covariantes u_1 et u_2 du vecteur \vec{v} s'expriment, en fonction de ses nouvelles composantes u'_1 et u'_2 par la relation:

$$[u] = [\beta][u']$$

soit encore:

$$\begin{cases} u_1 = \beta_{11} u'_1 + \beta_{12} u'_2 \\ u_2 = \beta_{21} u'_1 + \beta_{22} u'_2 \end{cases}$$

La fonction caractéristique ayant une valeur invariante pour tout changement de coordonnées, on voit que l'expression $\Phi'_{\vec{x}}(u'_1, u'_2)$ de la fonction caractéristique dans le nouveau système d'axes se déduit de son expression $\Phi_{\vec{x}}(u_1, u_2)$ dans l'ancien système de référence par la relation:

$$\Phi'_{\vec{x}}(u'_1, u'_2) = \Phi_{\vec{x}}(\beta_{11} u'_1 + \beta_{12} u'_2, \beta_{21} u'_1 + \beta_{22} u'_2)$$

Il suffit de remplacer dans l'ancienne expression $\Phi_{\vec{x}}(u_1, u_2)$ les composantes u_1 et u_2 par leurs expressions en fonction des nouvelles composantes u'_1 et u'_2 du vecteur \vec{v} .

On montrerait, de même, que:

$$\Phi_{\vec{x}}(u_1, u_2) = \Phi'_{\vec{x}}(d_{11} u_1 + d_{12} u_2, d_{21} u_1 + d_{22} u_2)$$

Remarque. Si l'on choisit l'un des axes, par exemple l'axe ox' , de manière à ce qu'il soit perpendiculaire au vecteur \vec{v} , les composantes d_{21} et d_{22} du vecteur \vec{e}'_2 par rapport aux anciens axes vérifieront la relation:

$$d_{21} u_1 + d_{22} u_2 = 0$$

Dans ces conditions, les composantes u'_1 et u'_2 du vecteur \vec{U} dans le nouveau système de coordonnées auront pour valeurs :

$$\begin{cases} u'_1 = \alpha_{11} u_1 + \alpha_{12} u_2 \neq 0 \\ u'_2 = \alpha_{21} u_1 + \alpha_{22} u_2 = 0 \end{cases}$$

Dans le nouveau système de coordonnées, la fonction caractéristique du vecteur \vec{x}' aura donc pour expression

$$\Phi'_{\vec{x}'}(u'_1, 0) = \Phi'_{x'_1}(u'_1)$$

fonction caractéristique de la loi marginale en x'_1 .

La fonction caractéristique étant invariante pour tout changement d'axe, nous pouvons donc écrire :

$$\Phi_{\vec{x}}(u, v) = \Phi'_{x'_1}(u'_1)$$

On voit donc que la fonction caractéristique est égale à la fonction caractéristique marginale relative à un axe ox'_1 quelconque, la projection sur cet axe étant faite perpendiculairement au vecteur \vec{U} : la fonction caractéristique permet donc l'analyse de la distribution par tranches parallèles, perpendiculaires au vecteur \vec{U} .

II - 32 . Changement de vecteurs.

Donnons nous arbitrairement un vecteur contravariant \vec{B} fixe du plan $x_1 0 x_2$, ainsi qu'une matrice carré d'ordre [A] d'ordre 2.

A partir de ces éléments nous pouvons définir un nouveau vecteur contravariant \vec{Z} par la relation matricielle :

$$[\vec{Z}] = [A] [\vec{x}] + [\vec{B}]$$

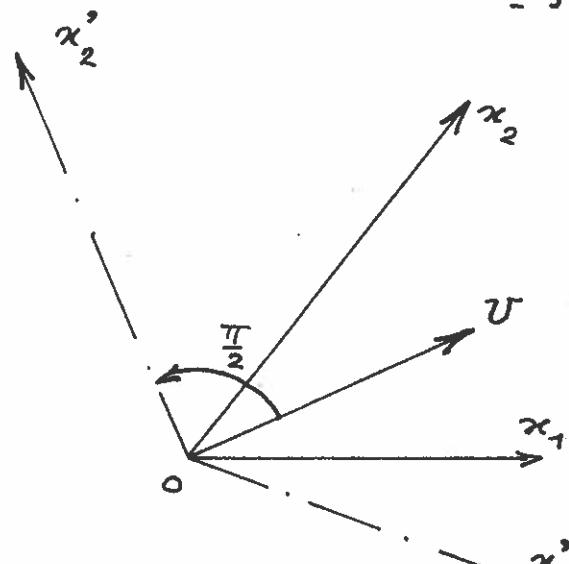


Fig II - 15

Les coordonnées contravariantes du vecteur
 \vec{z} ont pour valeurs:

$$\begin{cases} z_1 = a_{11}x_1 + a_{12}x_2 + b_1 \\ z_2 = a_{21}x_1 + a_{22}x_2 + b_2 \end{cases}$$

relations dans lesquelles b_1 et b_2 sont les coordonnées contravariantes du vecteur \vec{B} .

Si l'on désigne par $[U_T]$ la matrice transformée de la matrice $[U]$ des composantes covariantes du vecteur \vec{v} , le produit scalaire $\vec{v} \cdot \vec{z}$ a pour expression:

$$\vec{v} \cdot \vec{z} = [U_T][z] = [U_T][A][x] + [U_T][B]$$

La fonction caractéristique du vecteur \vec{z} a, par suite, pour valeur:

$$\Phi_{\vec{z}}(u_1, u_2) = E[e^{i\{[U_T][A][x] + [U_T][B]\}}] = e^{i[U_T][B]} \times E[e^{i[U_T][A][x]}].$$

soit, en définitive:

$$\Phi_{\vec{z}}(u_1, u_2) = e^{i[U_T][B]} \times \Phi_{\vec{x}}([U_T][A])$$

Cette relation généralise la formule identique obtenue au paragraphe I - 21 pour une distribution à une seule variable. Sous forme développée, la fonction $\Phi_{\vec{z}}$ s'écrit.

$$\Phi_{\vec{z}}(u_1, u_2) = e^{i(u_1 b_1 + u_2 b_2)} \times \Phi_{\vec{x}}\{a_{11}u_1 + a_{21}u_2, a_{12}u_1 + a_{22}u_2\}$$

II - 33. Développement en série de Mac-Laurin de la fonction caractéristique.

Lorsque ce développement est possible, la fonction caractéristique admet, au voisinage de l'origine, le développement en série de Mac-Laurin suivant:

$$\Phi(u_1, u_2) = 1 + \left[u_1 \frac{\partial^2 \Phi}{\partial u_1^2}(0,0) + u_2 \frac{\partial^2 \Phi}{\partial u_2^2}(0,0) + \dots \right] + \frac{\left(\left(u_1 \frac{\partial \Phi}{\partial u_1}(0,0) + u_2 \frac{\partial \Phi}{\partial u_2}(0,0) \right) \right)^r}{r!}$$

expression dans laquelle on doit remplacer les produits

$$\left(\frac{\partial \Phi}{\partial u_1} \right)^p \left(\frac{\partial \Phi}{\partial u_2} \right)^q \text{ par la dérivée } \left(\frac{\partial^{p+q} \Phi}{\partial u_1^p \partial u_2^q} \right)_{\substack{u_1=0 \\ u_2=0}}$$

Or, d'après la définition de la fonction caractéristique, cette dérivée a pour valeur :

$$\left(\frac{\partial^{p+q} \Phi}{\partial u_1^p \partial u_2^q} \right)_{\substack{u_1=0 \\ u_2=0}} = \left\{ \iint_{-\infty}^{+\infty} i^{p+q} x_1^p x_2^q e^{i(u_1 x_1 + u_2 x_2)} f(x_1, x_2) dx_1 dx_2 \right\}_{\substack{u_1=0 \\ u_2=0}}$$

$$= i^{p+q} m_{pq}$$

expression dans laquelle m_{pq} désigne le moment d'ordre p par rapport à x_1 et d'ordre q par rapport à x_2 .
En définitive la fonction caractéristique admet le développement suivant :

$$\begin{aligned} \Phi(u_1, u_2) &= 1 + i(u_1 m_{00} + u_2 m_{01}) - \frac{1}{2} (m_{20} u_1^2 + 2m_{11} u_1 u_2 + m_{02} u_2^2) + \dots \\ &\quad + \frac{i^r}{r!} \left((u_1 m_1 + u_2 m_2) \right)^r + \dots \end{aligned}$$

formule dans laquelle on doit remplacer les termes en $m_1^p m_2^q$ par le moment m_{pq} .

II - 34 Loi normale de Laplace-Gauss à deux variables

II - 35

Afin d'en simplifier l'écriture, nous supposons que la loi normale est rapportée à son centre de gravité, les variables x_1 et x_2 étant alors, par hypothèse, des variables centrées. Dans ces conditions la loi normale de Laplace-Gauss s'écrit :

$$\left\{ \begin{array}{l} dF = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}\chi^2} dx_1 dx_2 \\ \text{avec} \\ \chi^2 = \frac{1}{1-\rho^2} \left\{ \frac{x_1^2}{\sigma_1^2} - 2\rho \frac{x_1}{\sigma_1} \cdot \frac{x_2}{\sigma_2} + \frac{x_2^2}{\sigma_2^2} \right\} \end{array} \right.$$

Le calcul de la fonction caractéristique (par. II - 32) permettra de montrer que la matrice des variances s'écrit.

$$\tilde{V} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

et que, par suite, les courbes d'isodensité de la distribution (courbes $\chi^2 = \text{cste}$) sont homothétiques à son ellipse indicatrice.

II - 35 Tables numériques de la loi normale à deux variables.

Si l'on désigne par $1-\alpha$ la proportion de la population contenue à l'intérieur de l'ellipse d'isodensité définie par l'équation :

$$\chi^2 = \chi_{(2)}^2$$

dans laquelle $\chi_{(2)}^2$ est une constante, les valeurs de la constante $\chi_{(2)}^2$ sont données, en fonction des valeurs de α , sur la seconde ligne de la Table III. On trouve, en particulier, que :

20% de la population sont contenus à l'intérieur de l'ellipse $\chi^2 = 0,446$

50% " " " " $\chi^2 = 1,386$

80% " " " " $\chi^2 = 3,219$

95% " " " " $\chi^2 = 5,991$

99% " " " " $\chi^2 = 9,210$

999% " " " " $\chi^2 = 13,815$

II - 36. Distributions marginales et distributions liées de la loi normale.

On vérifie facilement que la loi normale à deux variables peut s'écrire sous l'une ou l'autre des deux formes suivantes:

$$(1) \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x_1^2}{\sigma_1^2}\right\} dx_1 \times \frac{1}{\sigma_2 \sqrt{1-\rho^2} \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{\left(x_2 - \rho \frac{\sigma_2}{\sigma_1} x_1\right)^2}{\sigma_2^2(1-\rho^2)}\right\} dx_2$$

$$(2) \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x_2^2}{\sigma_2^2}\right\} dx_2 \times \frac{1}{\sigma_1 \sqrt{1-\rho^2} \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{\left(x_1 - \rho \frac{\sigma_1}{\sigma_2} x_2\right)^2}{\sigma_1^2(1-\rho^2)}\right\} dx_1$$

Sous la première forme on voit que:

1-a) La distribution marginale de x_1 est la loi normale

$$x_1 | H = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x_1^2}{\sigma_1^2}\right\} dx_1, \quad \text{d'écart-type } \sigma_1.$$

1-b) La distribution de x_2 lié par x_1 est également une loi de Gauss qui a pour expression:

$$x_2 | x_1, H = \frac{1}{\sigma_2 \sqrt{1-\rho^2} \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{\left(x_2 - \rho \frac{\sigma_2}{\sigma_1} x_1\right)^2}{\sigma_2^2(1-\rho^2)}\right\} dx_2$$

L'espérance mathématique $E[x_2]$ de la variable x_2 pour une valeur donnée de x_1 (espérance mathématique liée), qui n'est autre que l'ordonnée de la corbe de régression, a pour valeur:

$$E[x_2]_{x_1} = \rho \frac{\sigma_2}{\sigma_1} x_1$$

(1) Nous adopterons l'écriture $\exp\{3\}$ pour e^3 lorsque l'exposant 3 sera très long.

ce qui prouve que la courbe de régression de x_2 en x_1 est rectiligne.

On voit, de plus, que l'écart-type de la loi liée, ou écart-type lié, a pour valeur:

$$\sigma_{x_1} = \sigma_2 \sqrt{1 - \rho^2}$$

Sous la deuxième forme on montrerait de même:

2-a) que la distribution marginale de x_2 est une loi normale d'écart-type σ_2

2-b) que la distribution de x_1 lié par x_2 est une loi normale dont l'espérance mathématique et l'écart-type liés ont pour valeurs:

$$\left\{ \begin{array}{l} E[x_1] = \rho \frac{\sigma_1}{\sigma_2} x_2 \\ \sigma_{x_1} = \sigma_1 \sqrt{1 - \rho^2} \end{array} \right. \quad (\text{régression linéaire})$$

II - 37. Calcul de la fonction caractéristique de la loi normale.

Pour calculer la fonction caractéristique de la loi normale à deux variables, nous utiliserons l'expression établie au paragraphe I-26, de la fonction caractéristique de la loi normale à une seule variable, que nous rappelons ci-dessous:

$$(3) \Phi(u) = \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{iux} \exp\left\{-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}\right\} dx = e^{ium} \cdot e^{-\frac{1}{2} u^2 \sigma^2}$$

La fonction caractéristique $\Phi_{\vec{x}}(u_1, u_2)$ de la loi

normale à deux variables se calcule par une intégrale double que l'on peut écrire de la manière suivante, en utilisant la forme (3) de la loi normale établie au paragraphe II-35 :

$$\Phi_{\vec{x}}(u_1, u_2) = \int_{-\infty}^{+\infty} \frac{e^{iu_1 x_1}}{\sigma_1 \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x_1^2}{\sigma_1^2}\right\} dx_1 \int_{-\infty}^{+\infty} \frac{e^{iu_2 x_2}}{\sigma_2 \sqrt{1-\rho^2} \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{(x_2 - \rho \frac{\sigma_2}{\sigma_1} x_1)^2}{\sigma_2^2 (1-\rho^2)}\right\} dx_2$$

D'après la formule (3), l'intégrale en dx_2 a pour valeur :

$$\int_{-\infty}^{+\infty} \frac{e^{iu_2 x_2}}{\sigma_2 \sqrt{1-\rho^2} \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{(x_2 - \rho \frac{\sigma_2}{\sigma_1} x_1)^2}{\sigma_2^2 (1-\rho^2)}\right\} dx_2 = e^{iu_2 \rho \frac{\sigma_2}{\sigma_1} x_1} \cdot \exp\left\{-\frac{1}{2} \frac{u_2^2 \sigma_2^2 (1-\rho^2)}{\sigma_2^2 (1-\rho^2)}\right\}$$

En reportant cette valeur dans la formule précédente il vient :

$$\Phi_{\vec{x}}(u_1, u_2) = \exp\left\{-\frac{1}{2} \frac{u_2^2 \sigma_2^2 (1-\rho^2)}{\sigma_2^2 (1-\rho^2)}\right\} \int_{-\infty}^{+\infty} \frac{e^{i(u_1 + \rho \frac{\sigma_2}{\sigma_1} u_2) x_1}}{\sigma_1 \sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x_1^2}{\sigma_1^2}\right\} dx_1$$

En appliquant une nouvelle fois la formule (3) pour évaluer maintenant l'intégrale en dx_1 , on obtient :

$$\Phi_{\vec{x}}(u_1, u_2) = \exp\left\{-\frac{1}{2} \frac{u_2^2 \sigma_2^2 (1-\rho^2)}{\sigma_2^2 (1-\rho^2)}\right\} \cdot \exp\left\{-\frac{1}{2} \left(u_1 + \rho \frac{\sigma_2}{\sigma_1} u_2\right)^2 \frac{\sigma_1^2}{\sigma_1^2}\right\}$$

Soit, finalement :

$$\boxed{\Phi_{\vec{x}}(u_1, u_2) = \exp\left\{-\frac{1}{2} \left(\sigma_1^2 u_1^2 + 2\rho \sigma_1 \sigma_2 u_1 u_2 + \sigma_2^2 u_2^2\right)\right\}}$$

Le développement en série de Mac. Laurin de la fonction caractéristique permet immédiatement de vérifier les résultats énoncés au paragraphe II - 34 à propos de la matrice des variances.

II - 38 Forme canonique de la loi normale. Couple de variables réduites indépendantes.

Par un changement de variables approprié, on peut faire disparaître le terme rectangle de l'exposant χ^2 qui figure dans l'expression de la loi normale réduite à deux variables. Le terme rectangle ayant disparu, les nouvelles variables sont alors des variables indépendantes. Pour obtenir un tel résultat, il suffit de choisir les directions des nouveaux axes de manière à ce qu'ils forment un couple de diamètres conjugués par rapport à l'ellipse indicatrice de la distribution.

On peut aller plus loin dans cette voie en s'imposant que les nouvelles variables soient non seulement des variables indépendantes mais encore des variables réduites. On démontre facilement qu'il existe une infinité simple de couples de variables réduites indépendantes t_1 et t_2 , correspondant à une loi normale donnée. Ces couples t_1 et t_2 sont définis par les formules de changement de coordonnées suivantes:

$$\left\{ \begin{array}{l} t_1 = \frac{1}{\sin \varphi} \left\{ -\frac{x_1}{\sigma_1} \sin(w - \frac{\varphi}{2}) + \frac{x_2}{\sigma_2} \sin(w + \frac{\varphi}{2}) \right\} \\ t_2 = \frac{1}{\sin \varphi} \left\{ -\frac{x_1}{\sigma_1} \cos(w - \frac{\varphi}{2}) + \frac{x_2}{\sigma_2} \cos(w + \frac{\varphi}{2}) \right\} \end{array} \right.$$

dans lesquelles $\left\{ \begin{array}{l} \text{l'angle } \varphi \text{ est défini par la relation } \cos \varphi = \rho \\ \text{l'angle } w \text{ est un paramètre arbitraire.} \end{array} \right.$

En appliquant un tel changement de variable on vérifie que la loi normale à deux variables s'écrit sous la forme suivante:

$$dF = \frac{1}{2\pi} e^{-\frac{1}{2}(t_1^2 + t_2^2)} dt_1 dt_2$$

appelée "forme canonique" de la loi normale à deux variables. On constate bien que les variables t_1 et t_2 sont des variables réduites et indépendantes.

La fonction caractéristique de la forme canonique a pour expression :

$$\underline{\Phi}_{\overline{T}}(u_1, u_2) = \exp\left\{-\frac{1}{2}(u_1^2 + u_2^2)\right\}$$

Le changement de variable précédent montre enfin que la quantité χ^2 qui figure dans l'expression générale d'une loi normale à deux variables (paragraphe II - 34) peut toujours se mettre sous la forme de la somme des carrés de deux variables normales réduites indépendantes :

$$\chi^2 = t_1^2 + t_2^2$$

Remarque.

Les formules de transformation inverses

$$\begin{cases} x_1 = \left\{ t_1 \cos(\omega + \frac{\varphi}{2}) - t_2 \sin(\omega + \frac{\varphi}{2}) \right\} \sigma_1 \\ x_2 = \left\{ t_1 \cos(\omega - \frac{\varphi}{2}) - t_2 \sin(\omega - \frac{\varphi}{2}) \right\} \sigma_2 \end{cases}$$

permettent, à partir de deux variables réduites indépendantes t_1 et t_2 , de définir une distribution de deux variables x_1 et x_2 , d'écart types σ_1 et σ_2 et de coefficient de corrélation $\rho = \cos \varphi$ données.

Ch III Fonctions Multivariantes

III-1 Esperance mathématique d'une Somme

Soient x_1, x_2, \dots, x_n les variables d'un même champ, distribuées ~~séparément~~ selon une loi de distribution commune:

$$dF = f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n$$

Considérons deux fonctions distinctes $Z_1(x_1, \dots, x_n)$ et $Z_2(x_1, \dots, x_n)$. Il nous est possible de calculer les espérances mathématiques des fonctions Z_1 et Z_2 par les relations:

$$E[Z_1] = \int Z_1 dF \quad E[Z_2] = \int Z_2 dF$$

Soit maintenant la somme

$$S = a_1 Z_1 + a_2 Z_2$$

dans laquelle, a_1 et a_2 sont des constantes. Il est évident que l'espérance mathématique de S est donnée par la formule:

$$E[S] = a_1 E[Z_1] + a_2 E[Z_2]$$

III-2 Esperance mathématique d'un produit

Considérons maintenant le produit $\Pi = Z_1 \times Z_2$. Pour calculer $E[\Pi]$ nous faisons apparaître les variables centrées \bar{z}_1 et \bar{z}_2 correspondant aux deux fonctions Z_1 et Z_2 . Nous posons

$$\begin{cases} Z_1 = m_1 + \bar{z}_1 & \text{avec } E[Z_1] = m_1 \text{ et par suite } E[\bar{z}_1] = 0 \\ Z_2 = m_2 + \bar{z}_2 & \quad " \quad E[\bar{z}_2] = 0 \end{cases}$$

$$\Pi = (m_1 + \bar{z}_1)(m_2 + \bar{z}_2) = m_1 m_2 + m_1 \bar{z}_2 + m_2 \bar{z}_1 + \bar{z}_1 \bar{z}_2$$

et par suite

$$E[\Pi] = m_1 m_2 + E[\bar{z}_1 \bar{z}_2] = E[Z_1] \times E[Z_2] + \text{cov}(Z_1, Z_2)$$

Si les fonctions Z_1 et Z_2 sont indépendantes,

III - 2

$\text{cov}(Z_1, Z_2) = 0$ et l'on a simplement.

$$E[\pi] = E[Z_1] \times E[Z_2]$$

III - 3 Ecart type d'une somme.

Considérons toujours les fonctions Z_1 et Z_2 définies au paragraphe I. Il est possible de calculer, par les équations suivantes, les écarts types et le coefficient de corrélation de Z_1 et Z_2 . On aura, avec des notations évidentes :

$$\left\{ \begin{array}{l} \sigma_1^2 = E[(Z_1 - E[Z_1])^2] = \int (Z_1 - E[Z_1])^2 dF \\ \sigma_2^2 = E[(Z_2 - E[Z_2])^2] = \int (Z_2 - E[Z_2])^2 dF \\ \rho \sigma_1 \sigma_2 = E[(Z_1 - E[Z_1])(Z_2 - E[Z_2])] = \int (Z_1 - E[Z_1])(Z_2 - E[Z_2]) dF \end{array} \right.$$

Proposons nous de calculer l'écart-type σ_S de la somme :

$$S = a_1 Z_1 + a_2 Z_2$$

$$\begin{aligned} \sigma_S^2 &= E[(S - E[S])^2] = E\left[\left\{a_1(Z_1 - E[Z_1]) + a_2(Z_2 - E[Z_2])\right\}^2\right] \\ &= a_1^2 E[(Z_1 - E[Z_1])^2] + a_2^2 E[(Z_2 - E[Z_2])^2] + 2a_1 a_2 E[(Z_1 - E[Z_1])(Z_2 - E[Z_2])] \end{aligned}$$

suit, finalement :

$$\sigma_S^2 = a_1^2 \sigma_1^2 + 2a_1 a_2 \rho \sigma_1 \sigma_2 + a_2^2 \sigma_2^2$$

En particulier si les variables x_1, x_2, \dots, x_n sont indépendantes, l'écart type de la somme

$$S = \sum a_i x_i$$

a pour valeur

$$\sigma_S^2 = \sum a_i^2 \sigma_i^2$$

(variables indépendantes)

III-4 Ecart-type d'une moyenne arithmétique.

III - 3

Soient x_1, x_2, \dots, x_n n valeurs indépendantes d'une variable x , d'écart type σ_x et soit \bar{x}_n la moyenne arithmétique de ces n valeurs :

$$\bar{x}_n = \frac{1}{n} \sum x_i$$

On voit que \bar{x}_n peut être considérée comme la somme

$$\bar{x}_n = \sum_i \frac{1}{n} x_i$$

et que, par suite :

$$\sigma_{\bar{x}_n}^2 = \sum_i \frac{1}{n^2} \sigma_x^2 = \frac{1}{n^2} \sigma_x^2 \sum_i^n (1) = \frac{n \sigma_x^2}{n^2} = \frac{\sigma_x^2}{n}$$

$$\sigma_{\bar{x}_n} = \frac{\sigma_x}{\sqrt{n}} \quad (\text{variables indépendantes})$$

III-5 Fonction caractéristique d'une somme de variables indépendantes

Soient x_1, x_2, \dots, x_n n variables indépendantes dont les fonctions caractéristiques ont pour expression :

$$\Phi_i(u) = E[e^{iu x_i}]$$

Considérons la somme

$$S = \sum a_i x_i$$

La fonction caractéristique de cette somme s'écrit :

$$\Phi_S(u) = E[e^{iu \sum a_i x_i}] = E\left[\prod_i^n (e^{iu a_i x_i})\right]$$

Les variables étant indépendantes, on sait que l'espérance mathématique du produit est égale au produit des espérances mathématiques de ses facteurs d'où :

$$\Phi_S(u) = \prod_i^n E[e^{iu a_i x_i}] = \prod_i^n \{\Phi_i(a_i u)\}$$

III-6 Composition de variables normales

III-4

A titre d'application du théorème précédent, nous allons démontrer que si les variables x_1, x_2, \dots, x_n sont des variables indépendantes distribuées suivant des lois normales d'écart-type $\sigma_1, \sigma_2, \dots, \sigma_n$, la somme

$$S = \sum a_i x_i$$

est elle-même, distribuée suivant une loi de Gauss.

Nous désignons par $m_1, m_2, \dots, m_i, \dots, m_n$ les espérances mathématiques des variables x_1, x_2, \dots, x_n . En se reportant au paragraphe I-26, on voit que la fonction caractéristique de la variable x_i a pour expression:

$$\Phi(u) = e^{iu m_i} e^{-\frac{1}{2} u^2 \sigma_i^2}$$

En appliquant la formule établie au paragraphe III-5, on voit que:

$$\Phi(u) = \underset{s}{e}^{iu \sum a_i m_i} e^{-\frac{1}{2} u^2 \sum a_i^2 \sigma_i^2}$$

qui prouve que la variable S est distribuée suivant une loi de Gauss d'espérance mathématique et d'écart-type :

$$\left\{ \begin{array}{l} E[S] = \sum a_i m_i \\ \end{array} \right.$$

$$\sigma_S = \sqrt{\sum a_i^2 \sigma_i^2}$$

En particulier, la moyenne arithmétique d'une série de n valeurs indépendantes d'une variable normale x , d'écart-type σ_x , est elle-même une variable normale d'écart-type et d'espérance mathématique:

$$\left\{ \begin{array}{l} E[\bar{x}_n] = E[x] \\ \sigma_{\bar{x}_n} = \frac{\sigma_x}{\sqrt{n}} \end{array} \right.$$

III-7 Ecart type des petites fluctuations d'une fonction quelconque, autour d'une valeur fixe

III-5

Soit x_1, x_2, \dots, x_n les variables d'un même champ. On posera

$$\begin{cases} E(x_i) = m_i & \xi_i = x_i - m_i \\ E(x_j) = m_j & \xi_j = x_j - m_j \\ \text{var}(x_i) = \sigma_i^2 \\ \text{var}(x_j) = \sigma_j^2 \\ \text{cov}(x_i, x_j) = \rho_{ij} \sigma_i \sigma_j \end{cases}$$

Les variables ξ_i et ξ_j sont des fluctuations aléatoires, centrées sur les valeurs m_i et m_j . On suppose ces fluctuations ξ_i faibles devant les espérances mathématiques correspondantes m_i .

Soient, dans ce champ, une fonction $Z(x_1, x_2, \dots, x_n)$ des variables x_i . Nous appellerons Z_0 la valeur de Z lorsque tous les ξ_i sont nuls.

$$Z_0 = Z(m_1, m_2, \dots, m_i, \dots, m_n)$$

Pour de petites fluctuations ξ_i autour des valeurs m_i , on peut écrire :

$$Z = Z_0 + \sum_i^n \left(\frac{\partial Z}{\partial x_i} \right) \xi_i$$

$x_1 = m_1$
 $x_2 = m_2$

et par suite

$$\frac{\sigma^2}{(\partial Z)^2} = \sum_i \left(\frac{\partial Z}{\partial x_i} \right)^2 \sigma_i^2 + 2 \sum_{i \neq j} \rho_{ij} \sigma_i \sigma_j \left(\frac{\partial Z}{\partial x_i} \right) \left(\frac{\partial Z}{\partial x_j} \right)$$

avec $i \neq j$ dans le terme rectangle.

III-8 Tolérances d'ajustement

III-6

Soient x_1 et x_2 les diamètres d'un arbre et d'un alésage devant s'emboîter l'un dans l'autre. Le jeu J de l'assemblage est défini par :

$$J = x_2 - x_1$$

Supposons les diamètres x_1 et x_2 distribués suivant deux lois normales, de même écart type σ , mais d'espérance mathématique différentes

$$E[x_1] = m_1, \quad E[x_2] = m_2$$

99,8% des arbres auront leurs diamètres compris dans l'intervalle.

$$m_1 \pm 3,09 \sigma$$

99,8% des alésages dans l'intervalle $m_2 \pm 3,09 \sigma$

Théorie classique

$$\begin{cases} \text{Jeu maxi} & m_2 + 3\sigma - (m_1 - 3\sigma) = m_2 - m_1 + 6\sigma \\ \text{Jeu mini} & m_2 - 3\sigma - (m_1 + 3\sigma) = m_2 - m_1 - 6\sigma \end{cases}$$

Soit un intervalle total de variation

$$\text{Jeu} = m_2 - m_1 \pm 6\sigma$$

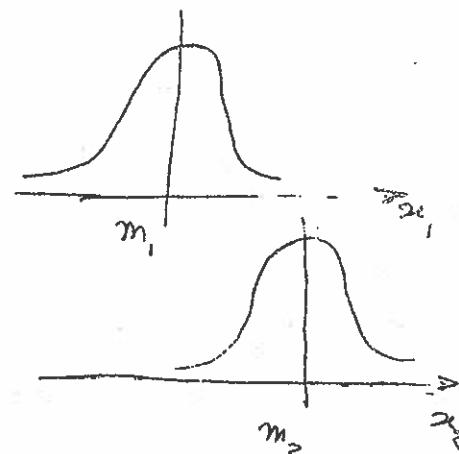
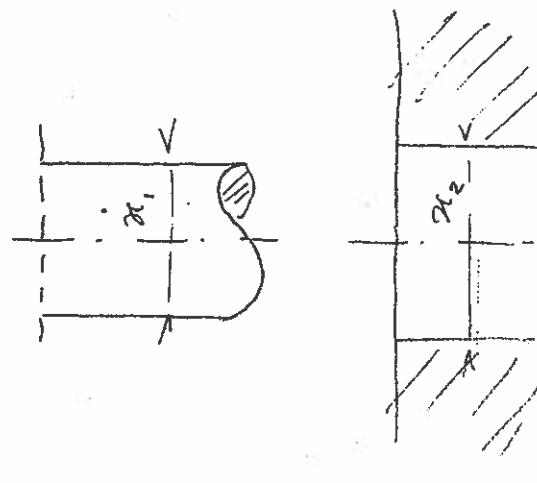
Théorie statistique

$$\frac{\sigma}{J}^2 = \frac{\sigma_1^2 + \sigma_2^2}{J^2} = 2\sigma^2 \quad \sigma_J = \sigma\sqrt{2}$$

99,8% des ~~jeux~~ assemblages auront des jeux compris dans l'intervalle

$$m_2 - m_1 \pm 3,09 \sigma_J = m_2 - m_1 \pm 3,09\sqrt{2}\sigma_J = m_2 - m_1 \pm 4,36\sigma$$

Résultat de plus de 27% inférieur à celui de la théorie classique



Etude statistique des diviseurs de tension.

1 Objectif de l'étude

On se propose,
à l'aide de deux
résistances R_1 et R_2
montées en série, de
réaliser un diviseur
de tension de rapport

$$(1) \quad \lambda = \frac{R_1}{R_1 + R_2}$$

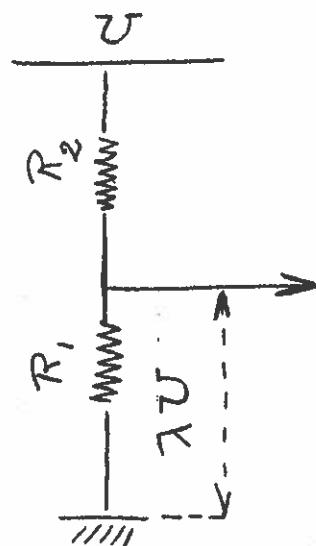


Fig. 1. Diviseur de tension

On utilise, pour cette application, des
résistances de valeurs nominales
effectives ρ_1 et ρ_2 , mais dont les
résistances réelles R_1 et R_2 sont
suivant l'ordre de présente, par rapport à
leurs valeurs nominales, de fluctuation
aléatoires ξ_1 et ξ_2 . On posera donc :

$$(2) \left\{ \begin{array}{l} R_1 = p_1 + \xi_1 \\ R_2 = p_2 + \xi_2 \end{array} \right.$$

Connaissons la distribution des fluctuations ξ_1 et ξ_2 , on se propose d'étudier les fluctuations correspondantes du rapport λ

II. Expression de λ pour de faibles fluctuations ξ_1 et ξ_2

En différentiant l'expression (1) on trouve immédiatement la relation

$$d\lambda = \frac{R_2 dR_1 - R_1 dR_2}{(R_1 + R_2)^2} = \frac{R_1 R_2}{(R_1 + R_2)^2} \left\{ \frac{dR_1}{R_1} - \frac{dR_2}{R_2} \right\}$$

Soit, encore :

$$(3) \quad d\lambda = \lambda(1-\lambda) \left\{ \frac{dR_1}{R_1} - \frac{dR_2}{R_2} \right\}$$

Si l'on désigne par λ_0 la valeur normale du rapport de division :

$$(4) \quad \lambda_0 = \frac{p_1}{p_1 + p_2}$$

et si l'on suppose que les fluctuations ξ_1 et ξ_2 sont assez petites pour que le rapport λ soit proche de λ_0 .

à un développement du premier degré, on trouve, d'après la formule (3) ci-dessus :

$$d\lambda = \lambda_0(1-\lambda_0) \left\{ \frac{\xi_1}{P_1} - \frac{\xi_2}{P_2} \right\}$$

Soit

$$(5) \quad \lambda = \lambda_0 + \lambda_0(1-\lambda_0) \left\{ \frac{\xi_1}{P_1} - \frac{\xi_2}{P_2} \right\}$$

III Fluctuations du rapport λ Parce que les résistances utilisées sont distribuées suivant des lois normales d'écart-type σ_1 et σ_2 données

La formule (5) permet le calcul immédiat de l'écart-type σ_λ caractérisant les fluctuations du rapport λ .

$$(6) \quad \sigma_\lambda = \lambda_0(1-\lambda_0) \sqrt{\frac{\sigma_1^2}{P_1^2} + \frac{\sigma_2^2}{P_2^2}}$$

Si l'on définit la classe d'une résistance par la valeur ρ_s de l'erreur relative maximum des résistances de celle classe, soit

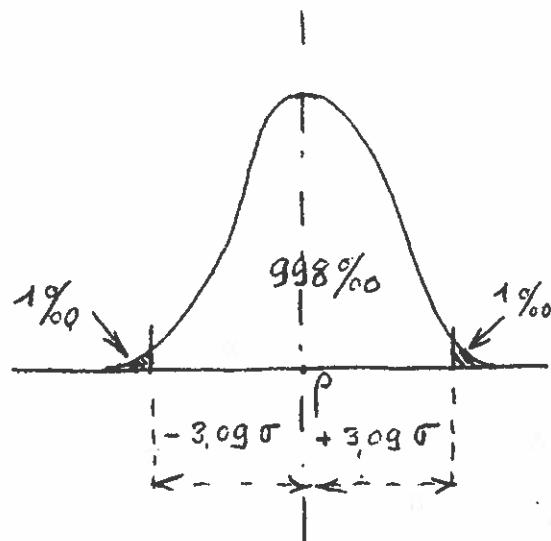


Fig 2. Loi normale

$$\rho = \frac{|\bar{\rho}_{\max}|}{\rho}$$

On sait que pour une distribution normale on peut admettre (Fig 2) que :

$$|\bar{\rho}_{\max}| = 3,09 \sigma$$

Dans ces conditions

$$\rho = \frac{3,09 \sigma}{\rho}$$

d'où l'on tire la valeur du rapport

$$\frac{\sigma}{\rho} :$$

$$\frac{\sigma}{\rho} = \frac{\rho}{3,09}$$

En reportant cette valeur dans la
formule 6, il vient

$$(7) \quad \sigma_{\lambda} = \frac{\lambda_0(1-\lambda_0)}{3,09} \sqrt{p_1^2 + p_2^2}$$

La fluctuation $\delta\lambda$ étant elle-même
distribuée suivant une loi de Gauss,
on sait, en particulier :

a) que 95% des valeurs λ seront
comprises dans l'intervalle $\lambda_0 \pm 1,96 \sigma_{\lambda}$
soit, ici

$$(8) \quad (95\%) \quad \lambda_0 \pm 0,635 \lambda_0(1-\lambda_0) \sqrt{p_1^2 + p_2^2}$$

b) que 99,8% des valeurs de λ
seront comprises dans l'intervalle
 $\lambda_0 \pm 3,09 \sigma_{\lambda}$, soit ici

$$(9) \quad 99,8\% \quad \lambda_0 \pm \lambda_0(1-\lambda_0) \sqrt{p_1^2 + p_2^2}$$

Ex. numérique

$$\begin{cases} p_1 = p_2 \text{ d'où } \lambda_0 = \frac{1}{2} \\ p_1 = p_2 = 10\% \end{cases}$$

95% des valeurs de λ seront comprises dans l'intervalle

$$\frac{1}{2} \pm 0,635 \times \frac{1}{2} \times \frac{1}{2} \times \sqrt{\frac{2}{100}} = \frac{1}{2} \pm 0,0224$$

soit $(0,478 ; 0,522)$

IV Fluctuations du rapport λ Parce que les résistances utilisées sont distribuées uniformément dans leurs intervalles de tolérance d'après les études relatives p_1 et p_2 .

Les distributions des résistances R_1 et R_2 sont reportées sur la fig. 3, conformément aux hypothèses adoptées dans ce paragraphe.

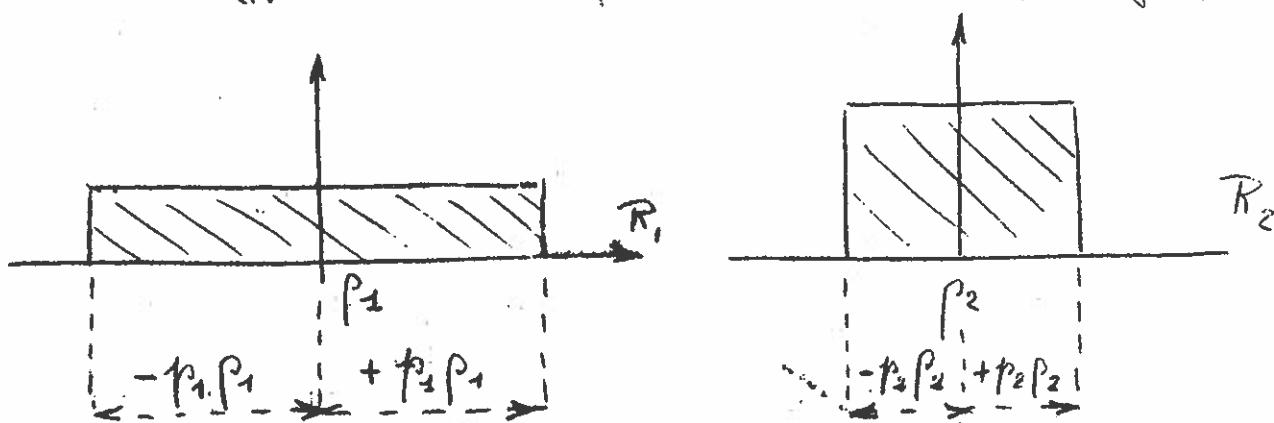


Fig. 3 : Distributions uniformes des résistances R_1 et R_2

Dans ces conditions les variables $\frac{\xi_1}{P_1}$ et $\frac{\xi_2}{P_2}$, figurant dans la formule (5) sont, elles mêmes, distribuées uniformément dans des intervalles $\pm P_1$ et $\pm P_2$ autour de la valeur 0.

Un point choisi, formé par l'association de deux résistances R_1 et R_2 choisies au hasard, constitue, du point de vue statistique, un objet possédant deux caractères (les valeurs des deux résistances R_1 et R_2) dont la distribution est représentée sur la fig. 4, dans le plan des écart-relatifs $x_1 = \frac{\xi_1}{R_1}$ et $x_2 = \frac{\xi_2}{R_2}$.

En dehors du contour rectangle ABCD, la surface de fréquence est confondue avec le plan de référence x_1, x_2 , alors qu'à l'intérieur du contour ABCD la surface de fréquence est représentée par une fraction de plan horizontal de côté non nulle (analogue avec une boîte rectangulaire posée sur le plan horizontal (x_1, x_2)).

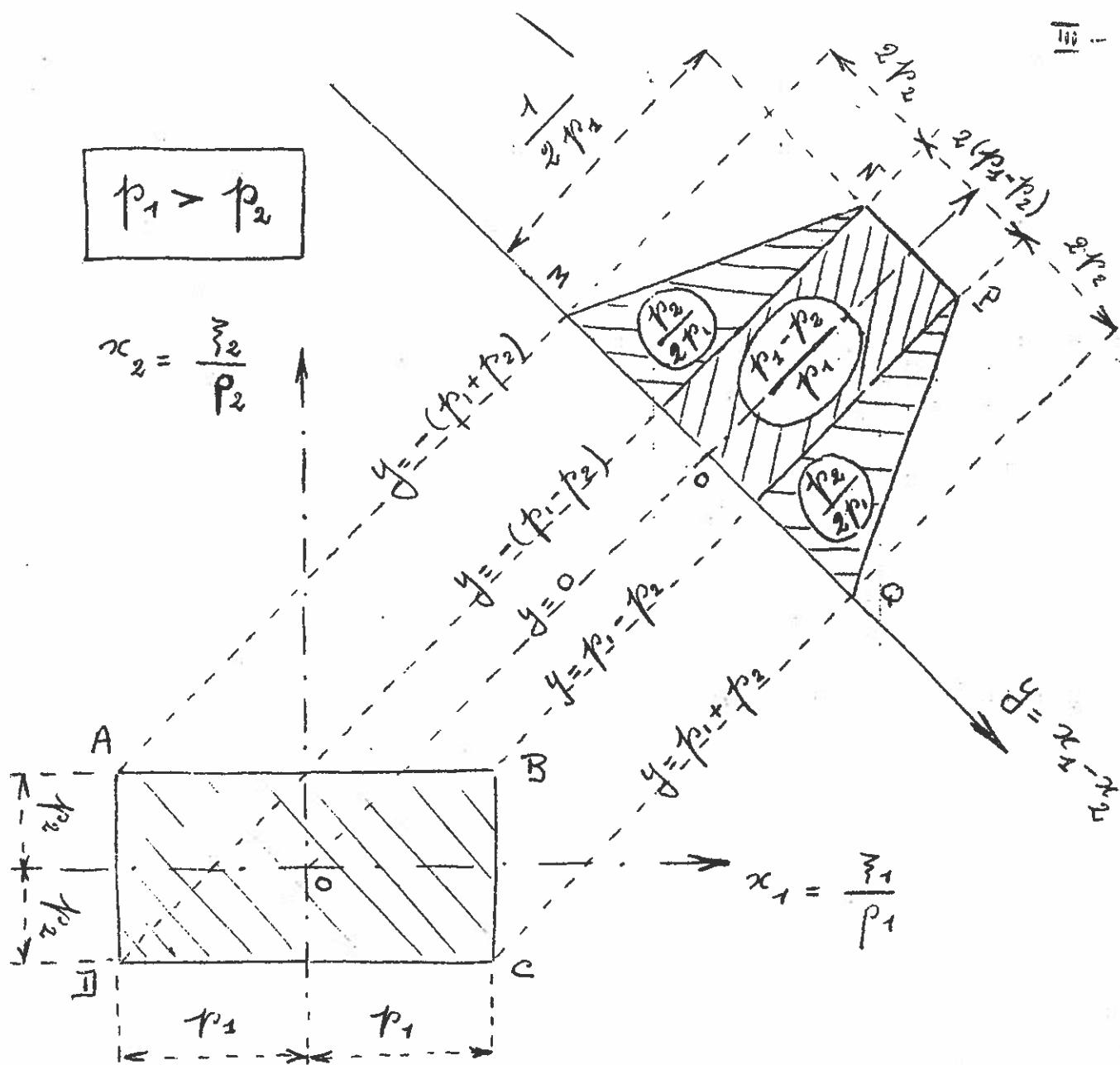


Fig 4 Joint Gaussian simulation
des écarts relatifs x_1 et x_2 et distribution
de la différence $y = x_1 - x_2$

D'après la formule (5), la valeur du rapport λ de division des tensions peut s'écrire.

$$(10) \quad \lambda = \lambda_0 + \lambda_0(1-\lambda_0)(x_1 - x_2)$$

A des constantes près, le rapport λ est donc distribué comme la différence

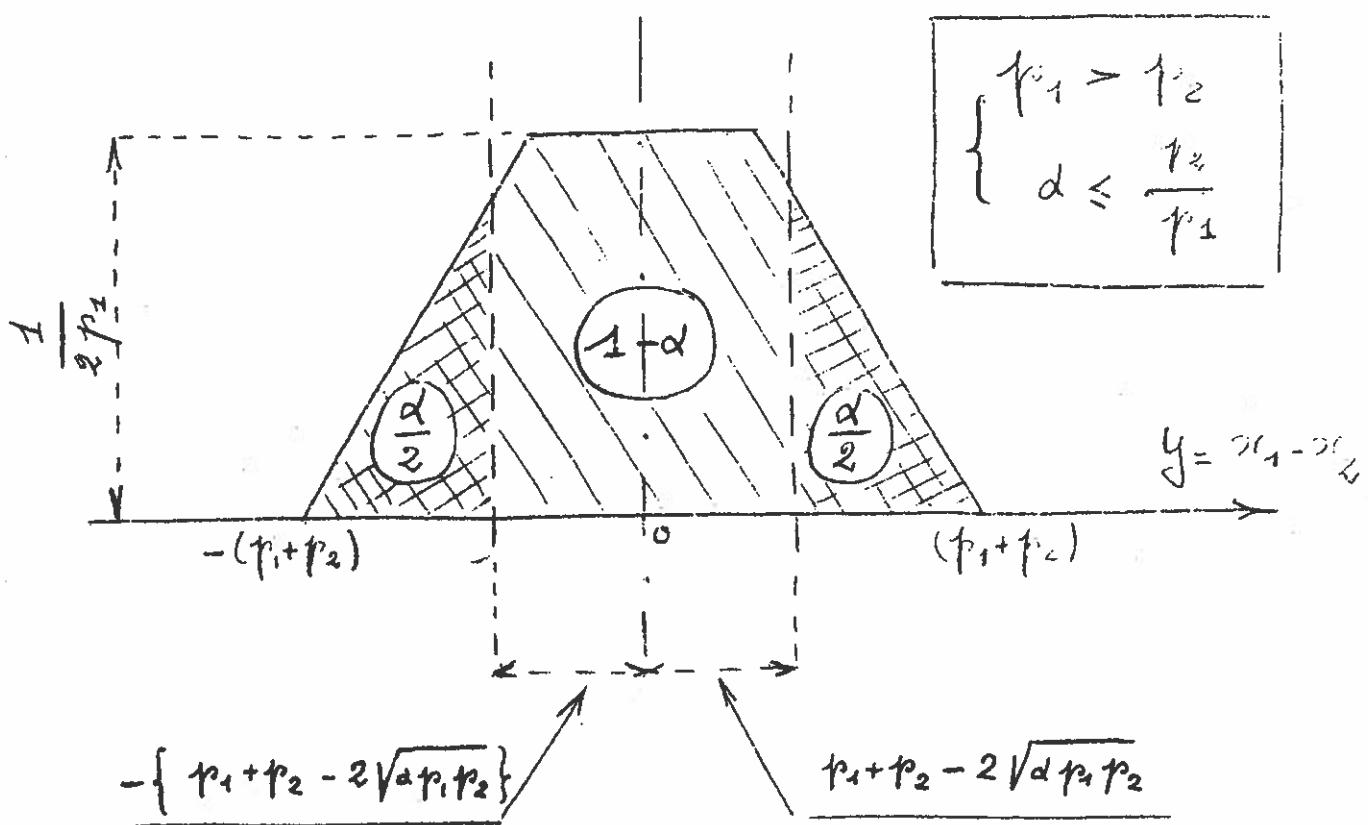
$$y = x_1 - x_2$$

La distribution de la différence y , figurée dans l'angle supérieur droit de la fig 4, s'obtient facilement en couvrant la surface de fréquence ABCD par des plans $y = \text{cste}$. On voit immédiatement sur la figure 4 (trace en suffisant $p_1 > p_2$) que le diagramme de distribution de y a la forme d'un trapèze dont on a déterminé la hauteur de manière à ce que la surface de ce trapèze soit égale à l'unité.

La distribution de la différence y est reportée sur la fig 5. On peut montrer facilement sur cette figure que $1-d$ des valeurs de y

Fig 5 Distribution ...

$$y = x_1 - x_2$$



se trouvent situées dans l'intervalle

$$\pm \left\{ \dots \right\}$$

En se reportant à la formule
(10) on peut donc conclure que :

1°) Lorsque $p_1 > p_2$ et

$$\alpha \leq \frac{p_2}{p_1}$$

1-d des valeurs du rapport λ
sont comprises dans l'intervalle

(11)

$$p_1 \geq p_2 \quad d \leq \frac{p_2}{p_1}$$

$$\lambda_0 \pm \lambda_0(1-\lambda_0) \left\{ p_1 + p_2 - 2\sqrt{dp_1p_2} \right\}$$

Lorsque $p_2 > p_1$ et
 $d \leq \frac{p_1}{p_2}$, on montrera de même
que 1-d des valeurs du rapport λ
sont comprises dans l'intervalle

(12)

$$p_1 \leq p_2 \quad d \leq \frac{p_1}{p_2}$$

$$\lambda_0 \pm \lambda_0(1-\lambda_0) \left\{ p_1 + p_2 - 2\sqrt{dp_1p_2} \right\}$$

Application numérique

$$\begin{cases} p_1 = p_2 \\ p_1 = p_2 = 10\% \end{cases} \text{ d'où } \lambda_0 = \frac{1}{2}$$

$$d = 0,05$$

95% des valeurs du rapport λ sont :
Comprises dans l'intervalle

$$\frac{1}{2} \pm \frac{1}{2} \times \frac{1}{2} \left\{ 0,20 - 2\sqrt{0,05 \times 0,1 \times 0,1} \right\} = \frac{1}{2} \pm 0,0388$$

Soit $(0,4612 ; 0,538)$

Ces valeurs sont à rapprocher des résultats obtenus à la page III-12, pour des résistances distribuées suivant deux normatifs.

II Distribution du rapport λ lorsque les résistances sont distribuées suivant deux Pois en parallèle

Lorsqu'un même fabricant produit simultanément des résistances de qualité différentes, par exemple des résistances à 10% et des résistances à 5%, il lui est possible d'obtenir les résistances à 5% parmi les résistances à 10%. Il résulte d'une telle pratique que les résistances à 5% sont distribuées à peu près uniformément dans leur intervalle de tolérances, alors

que les résistances rendues à 10% sont approximativement distribuées selon une distribution en creneau, analogue à celle représentée sur la fig 6

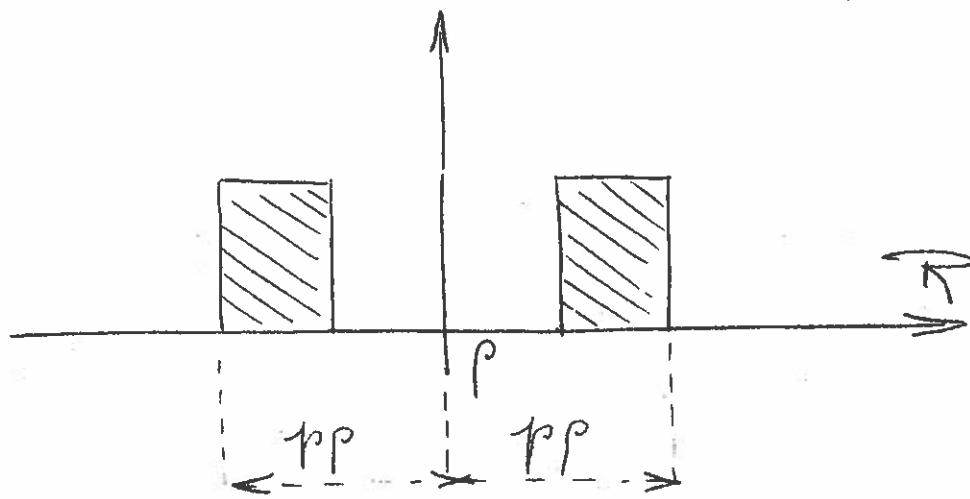


Fig 6 Distribution en creneau

En supposant que les résistances utilisées pour établir le niveau de tension sont celles suivant des distributions en creneau identiques en valeurs relatives, la fig 7, résultant d'une manière analogue à la fig 4, montre que la distribution des différences y se compose de 3 triangles isocèles de hauteurs inégales et dont les surfaces représentent, respectivement, 25% 50% et 25% de la population totale.

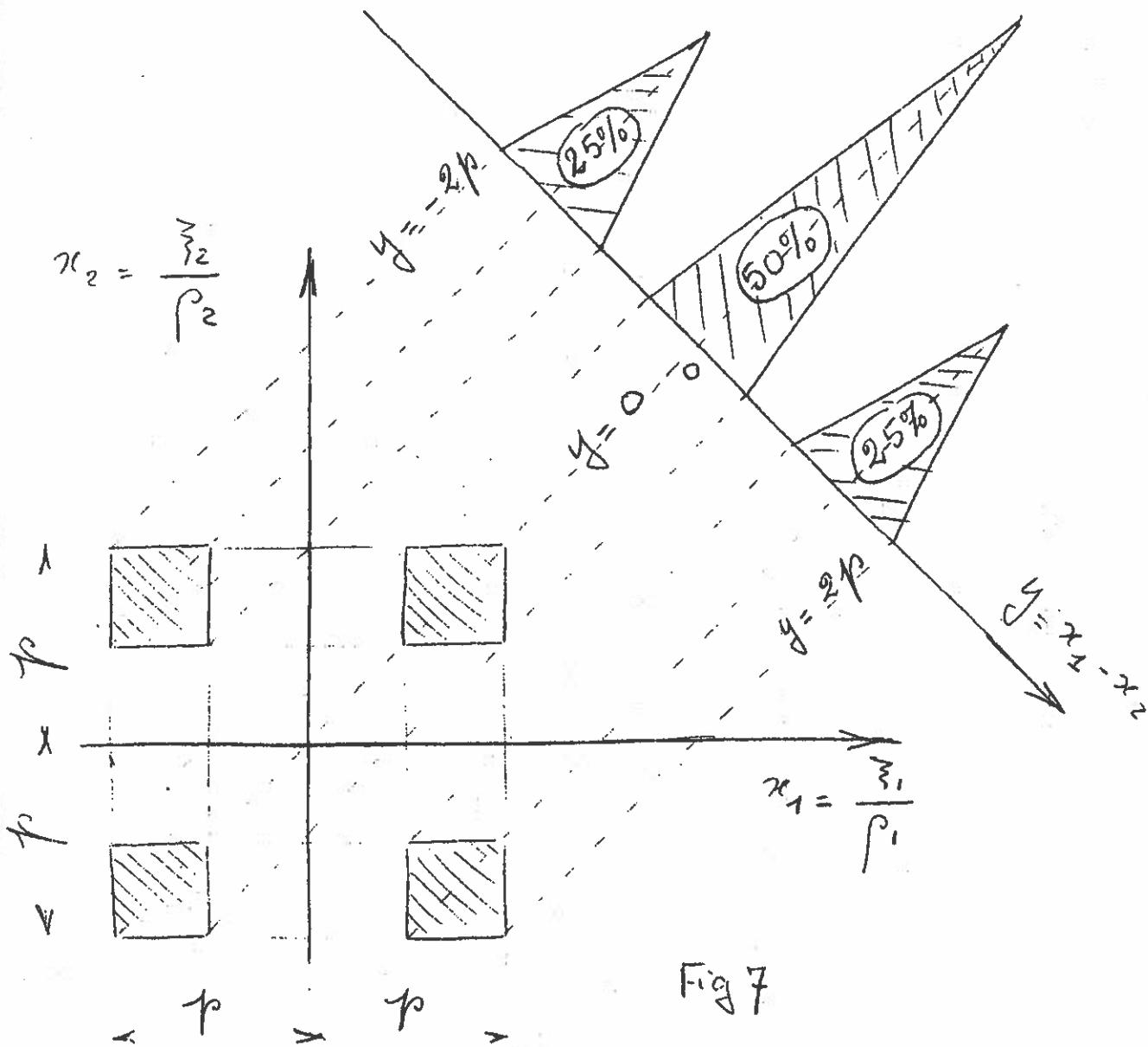


Fig 7

Au facteur $\lambda_0(\lambda - \lambda_0)$ près, la distribution du rapport λ est identique à celle de la différence y . On constate donc que l'emploi de résistances distribuées suivant des lois en creux peut entraîner l'apparition de groupements moléculaires dans la distribution du rapport λ .

II. Partie. Calcul des Probabilités

ch. IV. Événements aléatoires et probabilités

II.-1 Événement aléatoire

On dit qu'un événement est aléatoire lorsque, nous plongant dans des circonstances dans lesquelles nous savons que cet événement peut se produire, nous sommes incapables de prévoir s'il se réalisera ou non. Ex.

partie de pile ou face

Prélever un hasard d'une boîte et dans un sac contenant des billets de différentes couleurs.

Portée exacte d'un coup de canon

Trajectoire d'un électron particulier rencontrant un réseau cristallin

II.-2 Probabilité d'un événement aléatoire

Tous certaines circonstances on peut observer que la fréquence de réalisation d'un événement aléatoire semble tendre vers une limite fixe lorsque le nombre des essais augmente indefinitely.

Cette fréquence limite, si on peut en admettre l'existence, est appelée "Probabilité de l'événement considéré".

Cette notion de fréquence limite suppose essentiellement une stabilité des circonstances dans lesquelles peut prendre naissance l'événement considéré, afin que l'on puisse raisonnablement considérer que ce qui s'est produit dans le passé se reproduira dans l'avenir.

On ne peut donc parler de probabilité sans définir ou concevoir le champ de circonstances dans lequel l'événement pourra se produire. Aussi

La probabilité de tirer le roi de pique dans un jeu de carte n'est pas la même dans un jeu de 32 cartes ou dans un jeu de 52.

IV - 3. Restriction de von Mises

Pour définir la probabilité d'un événement aléatoire nous avons été amené à envisager la suite chronologique des résultats d'une série indéfinie d'essai, par exemple d'une série de parties de pile ou face dont nous notons les résultats sous la forme suivante :

P F F P P F P F P P P F P ..

Si l'expérience démontre que la fréquence d'apparition de pile tend vers une limite fixe p (que cette limite soit ou non égale à $1/2$), nous dirons que cette limite p constitue la probabilité de l'événement considéré.

Considérons alors la suite de résultats suivante

P F P F P F P F P F P F ..,

suite alternée dans laquelle la fréquence limite d'apparition de pile tend bien vers une limite fixe, ici égale à $1/2$. Dans de telles circonstances nous ne pourrions cependant pas parler de probabilité car l'événement n'est plus aléatoire : sa réalisation peut-être prévue si l'on connaît le résultat de la partie précédente.

Le concept d'événement aléatoire suppose que la suite chronologique des résultats d'une série d'essais est dépourvue de toute propriété systématique, ce que von Mises analyse comme suit :

Si par un procédé systématique quelconque on ~~peut~~ extrait de la suite chronologique des résultats une suite partielle (par exemple tous les résultats dont le rang est pair, ou dont le rang est un nombre premier etc...), la fréquence des succès dans la suite partielle ainsi isolée doit, à la limite, être égale à la fréquence limite des succès dans la suite complète des résultats d'essais.

L'expérience des jeux de hasard ou le succès des théories probabilistes dans les sciences physiques démontre qu'il est possible d'engendrer des suites de résultats qui soient effectivement dépourvus de propriétés systématiques.

IV - 4 Signification de l'échelle de probabilité

IV - 3

Tout comme une fréquence, la probabilité d'un événement aléatoire ne peut varier que de 0 à 1.

À une probabilité $p=0$ correspond un événement impossible.

À une probabilité $p=1$ correspond un événement certain.

A toute valeur intermédiaire correspond un événement que l'on peut observer plus ou moins fréquemment, selon la valeur de la probabilité correspondante.

Dans le domaine industriel, où le nombre des étais à enjolager dans une période de temps relativement courte, ne dépasse pas quelques milliers, nous considérons:

- a) comme normal tout événement dont la probabilité est inférieure à 0,05. L'apparition d'un tel événement sera le plus souvent considérée comme l'avertissement d'une modification éventuelle des conditions de fabrication.
- b) comme impossible tout événement dont la probabilité est inférieure à 0,001.

Ces règles ne sont évidemment pas valables lorsque l'on doit prendre en considération un nombre très grand (supérieur à 10^6 par exemple) d'occurrences dans lesquelles l'événement peut se produire. Ainsi est-il possible d'observer, à l'aide d'un scintillateur, la désintégration individuelle d'atomes d'une préparation radioactive très diluée placée sous le champ d'un microscope, bien que la probabilité de désintégration dans une minute d'un atome particulier soit extrêmement petite. Mais dans ce cas le nombre d'occurrences possibles correspond aux nombres d'atomes situés dans le champ d'observation et ce nombre dépasse certainement 10^{10} .

Notons cependant que ceux qui jouent régulièrement à la Loterie Nationale ne désespèrent pas (bien à tort et à leurs dépens!) de gagner au moins une fois dans le courant de leur vie, l'un des gros lots de la Loterie! À la chance de 52 tirages par an pendant 60 ans, le nombre d'occurrences ne dépasse pas 3.120, ce qui est très faible pour des probabilités de gain de gros lots bien inférieures à 10^{-4} .

IV - 5. Estimation de la probabilité de réalisation d'un événement aléatoire.

IV - 4

En règle générale, il est impossible de déterminer, à priori, la probabilité de réalisation d'un événement aléatoire. On doit en déterminer expérimentalement une valeur approchée, en procédant à une série d'essais permettant de calculer la fréquence des succès observés. Cette fréquence sera d'autant plus voisine de la probabilité cherchée que le nombre des essais effectués sera plus élevé.

Dans quelques cas exceptionnels, on peut cependant déterminer à priori la valeur de la probabilité de réalisation de l'événement aléatoire considéré. Nous citerons :

a) Le cas où l'événement aléatoire résulte d'un choix "au hasard" dans un ensemble ou population dont la structure est connue, étant admis que l'on entend par "choix au hasard" l'existence d'un procédé de choix qui, répété indéfiniment, conduirait à une population de même structure que celle dans laquelle le prélevement est effectué. Ainsi nous disons qu'il y a une chance sur 52 de tirer l'as de coeur dans un jeu de 52 cartes parce que nous admettons implicitement que si l'on nous présente 520.000 jeux de 52 cartes et si nous tirons une carte au hasard dans chaque jeu, nous aurons finalement tiré 520.000 cartes nous permettant de reconstituer très approximativement 10.000 jeux complets de 52 cartes. Nous disposons accessoirement de 10.000 as de coeur sur 520.000 cartes tirées, soit une fréquence de succès de $1/52$.

b) Le cas où la répétition indéfinie des essais conduirait à une population dont la structure peut être déterminée à priori par l'étude des phénomènes qui lui donnent naissance. Ainsi en jouant à pile ou face avec une pièce parfaitement symétrique, on peut admettre, a priori, que la probabilité de succès sera égale à $1/2$. Dans de telles circonstances, la valeur numérique de la probabilité d'un événement aléatoire résulte du dénominateur direct des cas favorables et des cas possibles, supposés tous également probables.

$$P = \frac{\text{nombre des cas favorables}}{\text{nombre des cas possibles}}$$

A la réalisation de chacun des événements aléatoires d'un ensemble d'événements possibles, on peut toujours associer une variable x dont la valeur dépendra de l'événement réalisé.

Dans le cas où deux alternatives sont seules possibles (par exemple prélevement d'une pièce bonne ou d'une pièce mauvaise dans un lot de pièces usinées présentées à un contrôle) la variable x peut être une variable binaire prenant la valeur 0 pour une pièce bonne et la valeur 1 pour une pièce mauvaise.

Lorsque l'événement aléatoire peut être identifié par un caractère mesurable x (par exemple le poids ou le diamètre d'une pièce, le rendement d'une réaction chimique...), on est naturellement conduit à choisir cette même x comme variable aléatoire.

Population-mère Loi de probabilité.

Pour définir les probabilités des divers événements d'un ensemble d'événements aléatoires possibles, nous avons été conduits à admettre que la répétition indéfinie des essais engendre une population limite de structure bien déterminée, que nous appellerons désormais "Population Mère".

Lorsque les divers événements aléatoires possibles peuvent être caractérisés par les diverses valeurs d'une ou plusieurs variables aléatoires $x, y, z\dots$, la distribution des variables $x, y, z\dots$ dans la population Mère constitue, par définition, la loi de probabilité des variables $x, y, z\dots$. Si cette loi de distribution admet, par exemple, une densité $f(x, y, z\dots)$, la probabilité dP d'observer un événement aléatoire dont les mesures soient comprises dans les intervalles $x, x+dx$; $y, y+dy$; $z, z+dz\dots$ a pour valeur

$$dP = f(x, y, z\dots) dx dy dz \dots$$

Cette loi de probabilité, qui n'est, en fait, que la loi de distribution d'une population limite, permet, en particulier, de définir l'espérance mathématique et les moments de la variable aléatoire ainsi que la fonction caractéristique de la loi de probabilité correspondante.

Dans un jeu de hasard on voit donc que

l'espérance mathématique du gain de l'un des joueurs n'est autre que son "espérance de gain" ou gain théorique moyen par partie, dont il se rapprochera d'autant plus qu'il aura joué un plus grand nombre de parties.⁽¹⁾ Cette observation justifie le terme d'espérance mathématique adopté pour désigner la valeur moyenne calculée pour la distribution limite à laquelle correspond la loi de probabilité.

IV - 7 Événements Composés.

Lorsque la réalisation d'un événement aléatoire R résulte de la réalisation ou de la non réalisation de deux événements aléatoires A et B , on dit que l'événement R est un événement composé (ou résultant) et que les événements A et B en sont les composants.

Nous nous proposons, connaîtissant les probabilités de réalisation des événements composants A et B , de calculer les probabilités de réalisation des deux types suivants d'événements composés :

1^o) L'événement composé R ne se produit qu'à la condition que l'un au moins des événements composants se soit, lui-même, réalisé. On dit alors que l'événement R résulte de la réunion des deux événements A et B , ce que l'on écrit symboliquement sous la forme

$$R = A + B \quad (\text{Réunion})$$

(Certains auteurs désignent la réunion par le symbole $A \cup B$)

2^o) L'événement composé R ne se produit qu'à la condition que les deux événements composants se soient produits simultanément. On dit alors que l'événement R résulte de l'intersection des événements A et B , ce que l'on écrit sous la forme :

$$R = A \cdot B \quad (\text{Intersection})$$

(Certains auteurs désignent l'intersection par le symbole $A \cap B$)

a) Il faut noter que dans tous les jeux publics de hasard, l'espérance de gain du joueur est toujours négative, ce qui se traduira toujours par la ferte irrémédiable du joueur persévérant !

Nous appellerons "complément de l'événement A" ou "événement complémentaire de l'événement A", l'ensemble des situations que l'on peut observer en cas de non réalisation de l'événement A considéré. Cet événement complémentaire sera désigné par le symbole \bar{A} . S'il s'agit des événements composés $A \cup B$ ou $A + B$, nous noterons leurs compléments par une barre s'étendant sur toute la longueur du symbole, soit respectivement par $\bar{A} \bar{B}$ et $\bar{A} + \bar{B}$.

Rappelons que la complémentation des événements composés est régie par les deux théorèmes de Morgan qui s'expriment par les relations d'identité:

$$\left\{ \begin{array}{l} \overline{A + B} = \bar{A} \bar{B} \\ \overline{AB} = \bar{A} + \bar{B} \end{array} \right.$$

La démonstration de ces relations est évidente si l'on se reporte au tableau (fig IV-1) sur lequel on a disposé toutes les intersections qu'il est possible de construire à partir des événements A et B et de leurs compléments. On voit, par exemple, que le complément $\bar{A} \bar{B}$ de l'événement AB (carreau inférieur droit) est formé des événements de la première ligne et de la première colonne du tableau, c'est à dire de la réunion $\bar{A} + \bar{B}$ des événements \bar{A} et \bar{B} .

	\bar{A}	A
\bar{B}	$\bar{A} \bar{B}$	$A \bar{B}$
B	$B \bar{A}$	AB

fig IV-1

IV - 8. Principe des probabilités totales.

IV - 8

Soient deux événements aléatoires A et B dont les probabilités de réalisation dans un champ de circonstances bien défini sont supposées connues.

Nous désignerons par H le champ de circonstances dans lequel peuvent se produire les événements A et B et nous noterons leurs probabilités de réalisation par les symboles :

$A|H$

et

$B|H$

Nous nous proposons, dans ces conditions, de calculer la probabilité $(A+B)|H$ de la réunion des événements A et B.

Pour concrétiser ce calcul, nous supposerons que nous disposons d'un sac contenant des boules en matière plastique de même diamètre, mais de couleur blanche ou noire. L'événement B se produira si l'on tire, au hasard, une boule blanche du sac.

Nous supposerons, en outre, que l'on a inclus au centre de chaque boule, au moment de son mouillage, un petit barreau métallique qui peut être soit magnétique, soit non magnétique. L'approche d'un aimant permet, dans ces conditions, de reconnaître la nature du barreau inclus dans chaque boule. Nous dirons que l'événement A s'est produit lorsque nous aurons tiré du sac une boule magnétique.

Les divers événements aléatoires qui peuvent se produire correspondent aux quatre intersections AB , $\bar{A}B$, $A\bar{B}$ et $\bar{A}\bar{B}$ qui symbolisent les quatre combinaisons possibles de couleur ou de propriété magnétique.

Le champ de circonstances H, dans lequel peuvent se produire l'un ou l'autre des 4 événements aléatoires possibles, est entièrement spécifié par le tableau de la fig IV-2, sur lequel on a reporté la composition du sac, contenant, au total, n boules. La désignation du nombre de boules contenues dans chaque case s'effectue logiquement en application des règles suivantes :

	non magnétiques Ā	Magnétiques A	Somme
B (noires)	$n - (a + b - c)$	$a - c$	$n - b$
B (Blanches)	$b - c$	c	b
Sommes	$n - a$	a	n

Fig. IV-2

- 1°) On désignera par \bar{b} le nombre total de boules blanches (composition marginale figurant à droite de la deuxième ligne)
- 2°) Il en résulte que le nombre de boules noires (composition marginale figurant à droite de la première ligne) sera égal à $n - \bar{b}$
- 3°) de la même manière on désignera par a le nombre total de boules magnétiques et par $n - a$ le nombre total de boules non magnétiques. (compositions marginales figurant à la dernière ligne)
- 4°) On désignera enfin par c le nombre total de boules blanches et magnétiques (caractère AB correspondant au Carré inférieur droit)
- 5°) Les nombres des boules des autres cases sont alors complètement déterminés si l'on remarque que les sommes des boules contenues dans chaque ligne ou chaque colonne doivent être égales aux compositions marginales correspondantes

Dans ces conditions on voit immédiatement que la réunion $A + B$ des événements A et B , qui correspond aux cases situées sur la deuxième ligne et sur la deuxième colonne du corps du tableau de la fig. IV-2, a pour probabilité :

$$(A + B) | \mathbb{H} = \frac{b - c + c + a - c}{n} = \frac{a}{n} + \frac{b}{n} - \frac{c}{n}$$

soit, en définitive :

$$(A + B) | \mathbb{H} = A | \mathbb{H} + B | \mathbb{H} - AB | \mathbb{H}$$

Résultat qui peut s'énoncer sous la forme suivante (principe des probabilités totales) :

"Lorsqu'un événement aléatoire $A + B$ peut se produire, soit par la réalisation d'un événement A , soit par celle d'un événement B , la probabilité de réalisation de l'événement considéré est égale à la somme des probabilités de réalisations de chacun des événements A et B , diminuée de la probabilité de réalisation simultanée des événements A et B ".

La fig. IV-3, dans laquelle on suppose que les domaines de réalisation des événements A et B correspondent aux régions situées à l'intérieur de deux courbes fermées, permet une interprétation visuelle immédiate de la relation précédente.

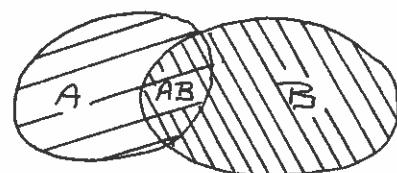


Fig. IV - 3

IV-9 Événements exclusifs.

Deux événements A et B sont dits exclusifs quand ils ne peuvent se réaliser simultanément. Dans ce cas la probabilité $AB | \mathbb{H}$ est nulle et l'on peut écrire

$$(A + B) | \mathbb{H} = A | \mathbb{H} + B | \mathbb{H} \quad \begin{matrix} \text{(Événements)} \\ \text{(exclusifs)} \end{matrix}$$

En particulier un événement A et l'événement complémentaire \bar{A} forment deux événements

et l'on peut écrire la relation :

$$(A + \bar{A})|H = A|H + \bar{A}|H = 1$$

IV - 10 Généralisation du principe des probabilités totales

Le principe des probabilités totales peut être étendu à un nombre quelconque d'événements aléatoires, en vertu du fait que l'addition logique ou réunion est une opération associative et commutative, et que la multiplication logique est distributive par rapport à l'addition.

Soit, en effet, à calculer la probabilité $(A+B+C)|H$ de trois événements aléatoires A , B et C . On peut écrire :

$$(A+B+C)|H = \{(A+B)+C\}|H = (A+B)|H + C|H - (A+B)C|H$$

Soit encore en développant le premier et le dernier terme :

$$\begin{aligned} (A+B+C)|H &= A|H + B|H - AB|H + C|H - (AC + BC)|H \\ &= A|H + B|H - AB|H + C|H - \{AC|H + BC|H - (AC)(BC)|H\} \end{aligned}$$

En tenant finalement compte de la propriété d'identité $a.a = a$ il vient :

$$(A+B+C)|H = A|H + B|H + C|H - (AB|H + BC|H + CA|H) + ABC|H$$

relation schématisée
par la fig IV-4

Dans le cas de plus de trois événements, les combinaisons des événements un à un, deux par deux, trois par trois, quatre par quatre etc... s'ajoutent avec des signes alternativement positifs et négatifs.

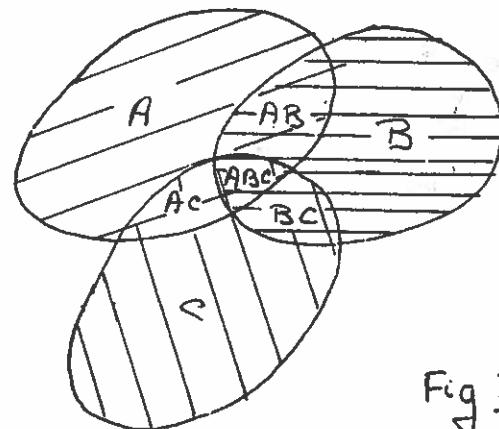


Fig IV-4

IV - 11 Principe des probabilités composées

Considérons à nouveau le sac rempli de boules noires ou blanches, magnétiques ou non magnétiques, dont la composition est définie par le tableau IV - 2.

Supposons que, ne disposant pas d'aimant, nous ayons tiré une boule blanche du sac. Nous ignorons si cette boule est magnétique ou non. Nous pouvons donc considérer comme un événement aléatoire le résultat qui donnera l'identification à l'aide d'un aimant du caractère magnétique ou non de cette boule. En se reportant au tableau IV - 2, on voit immédiatement que la probabilité $A|BH$ d'identifier comme magnétique une boule précédemment reconnue blanche est égale au quotient c/b du nombre c de boules blanches et magnétiques par le nombre total b de boules blanches.

$$A|BH = \frac{c}{b}$$

Cette probabilité liée se rattache aux probabilités $AB|H$ et $B|H$ si l'on remarque que l'égalité:

$$\frac{c}{n} = \frac{b}{n} \cdot \frac{c}{b}$$

s'interprète par la relation:

$$AB|H = B|H \times A|B|H$$

on démontrerait de même que :

$$AB|H = A|H \times B|AH$$

Ces résultats peuvent s'énoncer sous la forme suivante (principe des probabilités composées)

"Lorsqu'un événement aléatoire AB résulte de la réalisation de deux événements aléatoires A et B (intersection) La probabilité de réalisation de l'événement AB considéré est égale à la probabilité de réalisation de l'un des événements composants multipliée par la probabilité de réalisation de l'autre, sachant que le premier s'est effectivement réalisé"

Deux événements aléatoires A et B sont indépendants lorsque la probabilité de réalisation de l'un d'eux est indépendante de la réalisation ou de la non réalisation de l'autre.

Cette définition se traduit par les deux conditions suivantes :

$$\left\{ \begin{array}{l} A|BH = A|\bar{B}H \\ B|AH = B|\bar{A}H \end{array} \right.$$

En se reportant au tableau IV-2, on voit que ces deux conditions sont compatibles.
La première s'écrit, en effet :

$$\frac{c}{b} = \frac{a-c}{n-b} \quad \text{soit} \quad nc = ab$$

et la seconde

$$\frac{c}{a} = \frac{b-c}{n-a} \quad \text{soit encore} \quad nc = ab$$

La relation $nc = ab$ peut se mettre sous l'une des deux formes suivantes :

$$\frac{c}{a} = \frac{b}{n} \quad \text{soit} \quad B|AH = B|H$$

$$\frac{c}{b} = \frac{a}{n} \quad \text{soit} \quad A|BH = A|H$$

En reportant ces valeurs dans l'une ou l'autre des formules établies au paragraphe précédent, on obtient la relation fondamentale suivante :

$$AB|H = A|H \times B|H \quad (\text{variables indépendantes})$$

Résultat qui peut s'énoncer : "La probabilité de réalisation simultanée de deux événements aléatoires indépendants est égale au produit de leurs probabilités de réalisation respectives."

On remarquera que si l'on attribue à la réalisations des événements aléatoires A et B deux variables binaires x et y , les formules précédentes ne font que traduire l'indépendance de ces deux variables, telle que nous l'avons définie au paragraphe II-9 du chapitre II.

Soit un événement aléatoire A qui ne puise se produire qu'en un certain nombre de circonstances C_1, C_2, \dots, C_n . Nous appellerons "causes de l'événement A" les diverses circonstances C_i dont l'une au moins doit l'être probablement réalisée pour que il soit possible d'observer l'événement A considéré.

Nous désignerons par le symbole $A|C_i H$ la probabilité de réalisation de l'événement A, après réalisation effective de la cause C_i .

Nous supposerons en outre que les diverses causes possibles sont, elles-mêmes, des événements aléatoires dont les probabilités de réalisation $C_i | H$ sont toutes connues. Ces probabilités, connues antérieurement à toute tentative d'observation du phénomène A, sont appelées "Probabilités à priori" des causes C.

Nous nous proposons de calculer la probabilité "a posteriori" $C_i | AH$ pour que la cause C_i se soit effectivement réalisée, dans le cas où nous aurions physiquement constaté la réalisation de l'événement A. Cette probabilité "a posteriori" s'exprime en fonction des probabilités $A|C_i H$ de réalisation de l'événement A dans les diverses circonstances (ou causes) dans lesquelles il est possible de l'observer.

Considérons, en effet, les deux manières différentes d'exprimer la probabilité de réalisation simultanée des événements A et C_i :

$$\begin{cases} AC_i | H = A | H \times C_i | AH \\ AC_i | H = C_i | H \times A | C_i H \end{cases}$$

En comparant les deux seconds membres on en déduit la relation:

$$C_i | AH = C_i | H \times \frac{A | C_i H}{A | H}$$

qui constitue l'expression la plus générale du "Théorème de Bayes", reliant les probabilités "a priori" et "a posteriori" d'un même système de causes

Si l'on admet, de plus, que les diverses causes C_i sont exclusives les unes des autres, on peut alors obtenir la valeur de la probabilité $A|H$ par application du principe des probabilités totales. En se souvenant, en effet, que l'événement A ne peut être observé qu'après la réalisation de l'une des causes C_i , on en conclut que la probabilité $A|H$ a pour expression :

$$A|H = \sum_i A C_i | H = \sum_i \{ C_i | H \times A | C_i H \}$$

Dans ce cas particulier le théorème de Bayes s'exprime par la relation :

$$C_i | AH = C_i | H \times \frac{A | C_i H}{\sum_i \{ C_i | H \times A | C_i H \}} \quad (\text{causes exclusives})$$

Les applications pratiques du théorème de Bayes sont très limitées car il est, en général, impossible de déterminer (et même souvent de définir) les probabilités "a priori" $C_i | H$ des diverses causes possibles de l'événement A observé.

Lorsque l'on ignore les valeurs exactes des probabilités "a priori", Bayes a lui-même proposé, faut-il mieux, de les considérer comme toutes égales entre elles. Cette hypothèse, comme nous le nommons "Principe de Bayes", n'est, en général, pas justifiée. Son adoption entraîne donc toute valeur aux conclusions que l'on peut tirer du théorème de Bayes appliquée dans de telles conditions.

Exemple d'application du théorème de Bayes.

On dispose de n_1 jeux de 52 cartes et de n_2 jeux de 32 cartes. On choisit au hasard un de ces jeux et l'on tire dans ce jeu un roi de pique. On demande de calculer la probabilité "a posteriori" pour que le jeu choisi soit un jeu de 52 cartes.

Ici les deux cartes possibles sont le choix d'un jeu de 52 cartes ou d'un jeu de 32 cartes. Les probabilités "a priori" ont pour valeurs :

$$C_1 | H = \frac{n_1}{n_1 + n_2} \quad C_2 | H = \frac{n_2}{n_1 + n_2}$$

Les probabilités de tirer un roi de pique ont pour valeurs respectives.

$$A | C_1 H = \frac{1}{52} \quad A | C_2 H = \frac{1}{32}$$

Les deux causes possibles étant exclusives l'une de l'autre, on a :

$$A | H = C_1 | H \cdot A | C_1 H + C_2 | H \cdot A | C_2 H$$

Soit :

$$A | H = \frac{n_1}{n_1 + n_2} \times \frac{1}{52} + \frac{n_2}{n_1 + n_2} \times \frac{1}{32} = \frac{32n_1 + 52n_2}{32 \times 52 \times (n_1 + n_2)}$$

La probabilité cherchée a, par suite, pour valeur :

$$C_1 | AH = \frac{n_1}{n_1 + n_2} \times \frac{\frac{1}{52}}{\frac{32n_1 + 52n_2}{32 \times 52 \times (n_1 + n_2)}} = \frac{32n_1}{32n_1 + 52n_2}$$

$$C_1 | AH = \frac{\frac{32}{52} \times \frac{n_1}{n_2}}{1 + \frac{32}{52} \times \frac{n_1}{n_2}}$$

On voit que la valeur de la probabilité "a posteriori" varie de 0 à 1 lorsque le rapport $\frac{n_1}{n_2}$ varie de 0 à $+\infty$.

Cette probabilité "a posteriori" dépend donc essentiellement des informations (n_1 et n_2) que l'on a pu recueillir "a priori" sur les probabilités des causes.

On voudra :

Ensemble fondamental de parties \mathcal{F} et on définit une

s-Algèbre ou famille de Boole: \mathcal{F}

on considère $P(\mathcal{F})$ et $\bigcup_{\bar{A} \in \mathcal{F}} \bar{A} \in P(\mathcal{F})$ tel que :

$$\cdot \forall A \in \mathcal{F} \quad \bar{\bar{A}} = \bar{A} \in \mathcal{F}$$

• L'ensemble \mathcal{F} est fermé par rapport à la réunion d'un fini ou infini dénombrable d'éléments de \mathcal{F} . : $\bigcup_{i=1}^{\infty} (A_i \in \mathcal{F}) \in \mathcal{F}$

Propriétés de la s-Algèbre:

$$1) \forall A \in \mathcal{F}: \quad A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F} \Rightarrow A \cup \bar{A} \in \mathcal{F} \Rightarrow \emptyset \in \mathcal{F}$$

$$2) \emptyset \in \mathcal{F} \quad \emptyset = \bigcup_{\bar{A}} \bar{A} \in \mathcal{F}$$

$$3) \bigcup_{i=1}^{\infty} (A_i \in \mathcal{F}) \in \mathcal{F} \quad (\forall A_i \in \mathcal{F}) \quad \forall \bar{A}_i \in \mathcal{F} \quad \left[(\bigcup \bar{A}_i \in \mathcal{F}) = \bigcap_{i=1}^{\infty} A_i \in \mathcal{F} \right]$$

$(\mathcal{F}, \cup, \cap)$ = anneau commutatif unitaire non intégren

Ω possible. σ -Algèbre \mathcal{F} . Événement = $\forall A \in \mathcal{F}$

$\omega \in \Omega$: événement A réalisé.

$B|A$: B implique A. = B réalisé ne peut l'être que si A l'est. On considère $\mathcal{F} \xrightarrow{\begin{array}{l} A \mapsto \mathbb{R}_+ \\ A \mapsto p(A) \end{array}} \mathbb{R}_+$: $(\mathcal{F}, \mathcal{F}, p)$ définit un ensemble probabilité \mathcal{P} ,

$$1) p(\emptyset) = 0$$

$$2) \forall A \in \mathcal{F}, \quad \forall B \in \mathcal{F}, \quad A \cap B = \emptyset \Rightarrow p(A \cup B) = p(A) + p(B)$$

Propriétés:

$$1) \forall A \in \mathcal{F} \quad p(A) + p(\bar{A}) = 1 \quad (= p(A \cup \bar{A}) = 1)$$

$$2) p(\emptyset) = 0 \quad \emptyset = \bigcup_{\bar{A}} \bar{A} \Rightarrow p(\emptyset) + p(\bar{\emptyset}) = 1$$

$$3) \forall A \in \mathcal{F}, \quad p(A) + p(\bar{A}) = 1 \Rightarrow p(A) = 1 - p(\bar{A}) \leq 1.$$

 $\Rightarrow \text{Im } p = [0, 1]$

$$4) \forall A \in \mathcal{F} \quad B|A \Rightarrow p(B) \leq p(A) \quad \text{car } (B|A) \Leftrightarrow (A \cap B) / B = \emptyset \Rightarrow$$
$$p((A \cap B) / B) = p(A \cap B) = p(A) - p(A \cap B) = p(A)$$
$$\Rightarrow p(B) = p(A) - p(A \cap B) \leq p(A)$$

$$5) \text{ Si } A, B, C, D \dots \text{ lâts incompatibles, } p(A \cup B \cup C \cup \dots) = p(A) + p(B) + \dots$$

Principe de probabilité totale:

$\forall A \in \mathcal{F}$ et $\forall B \in \mathcal{F}$

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

$$A \cup B = (A \cap \bar{B}) \cup (\bar{A} \cap B) \cup (A \cap B)$$

$$\begin{aligned} P(A \cup B) &= P(A \cap \bar{B}) + P(\bar{A} \cap B) + P(A \cap B) - P(A \cap B) \\ &\quad P(A \cap (B \cap \bar{B})) + P(B \cap (\bar{A} \cup \bar{B})) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

on appelle système d'événement complet toute partition de l'algèbre \mathcal{F} .

$$\sum \text{prob. d'évén. complet} = 1$$

Probabilités conditionnelles:

(Ω, \mathcal{F}, P)

$$A / P(A) \neq 0$$

$$\mathcal{F} \quad P_A : \Omega \rightarrow \mathbb{R}^+ \quad \forall B \in \mathcal{F} \quad B \xrightarrow{P_A} P_A(B) \quad P_A(B) = \frac{P(A \cap B)}{P(A)} \quad \begin{matrix} \text{probab de } \\ \text{conditionnée} \\ \text{par } A. \end{matrix}$$

$$1) P_A(B) \in \mathbb{R}^+$$

$$2) P_A(\Omega) = \frac{P(A \cap \Omega)}{P(A)} = 1$$

$$3) \forall B \in \mathcal{F} \text{ et } \forall C \in \mathcal{F} \quad B \cap C = \emptyset \quad P_A(B) + P_A(C) = P_A(B \cup C) \quad \left. \Rightarrow (\Omega, \mathcal{F}, P_A) \right\} \text{ ensemble probabilité.}$$

• (A, \mathcal{F}_A, P_A) est un ensemble probabilisé; avec $\mathcal{F}_A = \{A \cap B / B \in \mathcal{F}\}$

$$\text{Notation: } P_A(B) = \frac{P(A \cap B)}{P(A)} = \text{prob. } (B/A) = \text{probabilité de } B \text{ si } A.$$

Principe des probabilités composées:

$$P(A) \neq 0 \quad \text{prob}(A \cap B) = \text{prob}(A) \cdot \text{prob}(B/A)$$

$$\begin{cases} P(A) \neq 0 \\ P(B) \neq 0 \end{cases}$$

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

Événements indépendants

$$P(A) \neq 0 \quad P(B|A) = P(B) \quad \Rightarrow \quad P(AB) = P(A) \times P(B)$$

généralement A et B indép. L'a 2.

Théorème de Bayes.

$$(\Omega, \mathcal{F}, P) \quad A \in \mathcal{F}$$

Supposons que nous savons x probabilités que n'ont pas de sens simultanément avec \wedge et \vee :
C₁, C₂, ..., C_n $\in \mathcal{F}$.

i.e.: $P(A) \neq 0$

$$\exists C_i \subset \{C_1, C_2, \dots, C_n\} \quad (P(C_i) \neq 0)$$

$$A = \bigcup_{i=1}^n C_i$$

Probabilité de priori $P(C_i)$

Probabilité a posteriori $P(C_i|A)$

$$P(AC_i) = P(A) P(C_i|A) = P(C_i) \times P(A|C_i)$$

$$\Rightarrow P(C_i|A) = \frac{P(C_i) \times P(A|C_i)}{P(A)}$$

$$\text{ou } P(A) = P\left(\bigcup_{i=1}^n (AC_i)\right)$$

Si tous C_i sont exclusifs d'après:

$$P(C_i|A) = \frac{P(C_i) \times P(A|C_i)}{P(A) \times P\left(\bigcup_{i=1}^n C_i\right)} \quad ? \text{ si } \bigcup_{i=1}^n C_i = \Omega \quad P(\Omega) = 1$$

$$P(C_i|A) = \frac{P(C_i) \times P(A|C_i)}{\sum_i [P(C_i) \times P(A|C_i)]}$$

Variable aléatoire:

(Ω, \mathcal{F}, P) Variable aléatoire sur \mathbb{R}^n = toute application de Ω vers \mathbb{R}^n :

$$\Omega \rightarrow \mathbb{R}^n \quad \omega \mapsto V(\omega)$$

telle que:

$$X_1(\omega) < x_1 \\ X_2(\omega) < x_2$$

$$X_n(\omega) < x_n$$

l'ensemble des ω ainsi définis
est nommé v est la loi de répartition de la variable aléatoire

ch V. Les épreuves multiples

Au cours de ce chapitre nous nous proposons de calculer, dans un certain nombre de cas particuliers, la probabilité pour qu'un ou plusieurs événements aléatoires se réalisent un nombre imposé de fois au cours d'une série d'épreuves multiples. Nous supposons comme les probabilités, éventuellement différentes d'une épreuve à l'autre, de réalisation de chacun des événements envisagés.

Nous examinerons successivement les problèmes d'échantillonnage dans une population d'effectif limité, puis les séries d'épreuves rejettées conduisant à la loi de Bernoulli. Nous rechercherons enfin les formes asymptotiques de la loi des épreuves rejettées, soit lorsque le nombre des épreuves augmente indefinitely (loi des Grands Nombres), soit lorsque la probabilité de l'événement attendu est très faible (loi de Poisson).

A. Echantillonnages exhaustifs.

V - 0.1 Echantillonnage par tirage exhaustif dans une population contenant des éléments de deux catégories différentes.

Disposant d'un lot de N pièces, dont on sait qu'il contient I pièces défectueuses et, par suite, $N-I$ pièces bonnes, on convient de prélever au hasard n pièces dans ce lot. On demande de calculer la probabilité pour que cet échantillon de n pièces contienne exactement d pièces défectueuses.

Ce mode d'échantillonnage est appelé "exhaustif" en ce sens que l'échantillon prélevé est susceptible d'épuiser le lot, lorsque l'effectif de l'échantillon est pris égal à celui du lot.

Remarquons tout d'abord que si l'on préleve l'une après l'autre les n pièces de l'échantillon, ces prélevements ne constituent pas n épreuves statistiquement indépendantes. En effet la probabilité d'extraire une pièce défectueuse dépend non seulement du rang du prélevement mais encore des prélevements antérieurs. On ne pourra donc pas utiliser les règles de calcul des probabilités se rapportant à un ensemble d'événements indépendants.

V-02

La méthode la plus simple pour calculer la probabilité P cherchée consiste, ici, à évaluer directement le nombre des cas possibles et le nombre des cas favorables.

$$P = \frac{\text{nombre des cas favorables}}{\text{nombre des cas possibles}}$$

a) nombre des cas possibles. Le nombre des cas possibles correspond au nombre d'échantillons différents de n pièces qu'il est possible de prélever dans le lot de N pièces. Ce nombre est égal au nombre de combinaisons de N individus différents pris n à n , soit

$$C_N^n = \frac{N!}{n!(N-n)!}$$

b) nombre des cas favorables. Le nombre des cas favorables correspond au nombre d'échantillons différents de n pièces contenant d pièces défectueuses qu'il est possible de prélever dans le lot de N pièces. On peut construire un tel échantillon:

1) en prélevant d'abord les d pièces défectueuses parmi les \bar{D} pièces défectueuses du lot. Le nombre de manières différentes d'effectuer ce choix est égal au nombre de combinaisons de \bar{D} individus pris d à d , soit:

$$C_{\bar{D}}^d = \frac{\bar{D}!}{d!(\bar{D}-d)!}$$

2) en prélevant ensuite les $n-d$ pièces bonnes parmi les $N-\bar{D}$ pièces bonnes du lot. Le nombre de manières différentes d'effectuer ce choix est égal au nombre de combinaisons de $N-\bar{D}$ individus pris $n-d$ à $n-d$, soit:

$$C_{N-\bar{D}}^{n-d} = \frac{(N-\bar{D})!}{(n-d)!\{N-\bar{D}-(n-d)\}!}$$

Le nombre total d'échantillons différents de n pièces contenant d pièces de lecteur est donc égal au produit.

$$C_{\bar{N}}^d \times C_{N-\bar{N}}^{n-d}$$

La probabilité cherchée a, par suite, pour expression:

$$(V-01) \quad p = \frac{C_{\bar{N}}^d \times C_{N-\bar{N}}^{n-d}}{C_N^n}$$

Application numérique

Dans un jeu de $N = 52$ cartes, on préleve au hasard une "main" de $n = 13$ cartes. Calculer la probabilité pour que cette main contienne au moins 5 honneurs d'une même couleur en appelant "honneurs" les 5 cartes dix, valet, dame, roi et as).

L'événement R auquel nous nous intéressons se réalise lorsque l'un ou l'autre des 4 événements élémentaires suivants se trouve lui-même réalisé:

- { T : la main contient 5 honneurs à trèfle
- Q : " " " " " carreau
- C : " " " " " coeur
- P : " " " " " pique

On peut donc écrire que l'événement R résulte de la réunion des quatre événements élémentaires T , Q , C et P .

Si l'on remarque que les intersections triples, du type TQC , et l'intersection quadruple $TQCP$ sont impossibles, comme correspondant à des nombres de cartes (-5 et 20) supérieurs à l'effectif de la main, on voit que la probabilité cherchée a pour valeur:

$$R|H = T|H + Q|H + C|H + P|H - \{TQ|H + TC|H + TP|H + QC|H + CP|H + CP|H\}$$

Les probabilités des 4 événements élémentaires du type T sont toutes égales. Il en est de même des probabilités des 6 intersecteurs doublés du type TQ . On pourra donc, par exemple, écrire la probabilité $R|H$ sous la forme :

$$R|H = 4 \times (T|H) - 6 \times (TQ|H)$$

Il nous reste à calculer les probabilités $T|H$ et $TQ|H$.

1) Calcul de $T|H$. Dans le jeu de $N=52$ cartes, nous pouvons considérer comme "défectueuses" les 5 cartes correspondant aux 5 hommes à trèfle, soit $I=5$.

Dans la main de $n=13$ cartes nous considérons également comme cartes défectueuses les 5 hommes à trèfle, soit $d=5$

Dans ces conditions.

$$T|H = \frac{\binom{5}{8} \times \binom{47}{5}}{\binom{52}{13}} = \frac{47!}{8! 39!} \times \frac{13! \times 39!}{52!} = \frac{13! \times 47!}{8! \times 52!}$$

2) Calcul de $TQ|H$. Dans le jeu de 52 cartes nous considérons maintenant comme cartes défectueuses les 5 hommes à trèfle et les 5 hommes à carreau, soit $I=10$

Dans la main de $n=13$ cartes nous considérons comme cartes défectueuses ces mêmes 10 hommes à trèfle et à carreau, soit $d=10$

Dans ces conditions :

$$TQ|H = \frac{\binom{10}{3} \times \binom{42}{10}}{\binom{52}{13}} = \frac{42!}{3! 39!} \times \frac{13! \times 39!}{52!} = \frac{13! \times 42!}{3! \times 52!}$$

On en déduit :

$$R|H = 4 \times \frac{13! \times 47!}{8! \times 52!} - 6 \times \frac{13! \times 42!}{3! \times 52!}$$

Soit :

$$R|H = 4 \times \frac{13! \times 47!}{8! \times 52!} \left\{ 1 - \frac{5}{4} \times \frac{4 \times 5 \times 6 \times 7 \times 8}{43 \times 44 \times 45 \times 46 \times 47} \right\}$$

ou :

$$R|H = \frac{33}{49 \times 340} \left\{ 1 - \frac{28}{11 \times 23 \times 43 \times 47} \right\}$$

soit

$$R|H = 1,98 \times 10^{-3}$$

Il existe donc un peu moins de 2 chances sur 1000 d'avoir au moins 5 honneurs d'une même couleur dans une main de 13 cartes.

V-02 Tirage exhaustif dans une urne contenant des boules de q catégories différentes.

On peut généraliser le problème résolu au paragraphe précédent en considérant une urne contenant des boules de plus de 2 catégories différentes. La composition de cette urne sera définie par les nombres D_1, D_2, \dots, D_q de boules de chaque catégorie. Son effectif total aura, par suite, pour valeur:

$$N = \sum D_i$$

La probabilité P de prélever dans cette urne un échantillon de n pièces contenant d_1 boules de la catégorie 1, d_2 boules de la catégorie 2, ..., \dots, d_q boules de la catégorie q , avec

$$n = \sum d_i$$

sera donnée par la formule :

$$(V-02) \quad P = \frac{C_{D_1}^{d_1} \times C_{D_2}^{d_2} \times \dots \times C_{D_q}^{d_q}}{C_N^n}$$

V - 03 Tirage exhaustif de plusieurs échantillons dans une urne contenant des boules de q catégories différentes.

Considérons à nouveau une urne contenant N boules de q catégories différentes et désignons comme précédemment par D_i le nombre total de boules de la catégorie i .

Nous nous proposons de calculer la probabilité P de prélever, simultanément ou successivement, dans cette urne k échantillons d'effectifs n_1, n_2, \dots, n_k (avec $\sum n_i = N$), de compositions fixées à l'avance. La composition de l'échantillon d'un rang i que l'on peut sera définie par les nombres (dont certains peuvent être éventuellement nuls)

$$d_{i1}, d_{i2}, \dots, d_{iq} \quad (\text{avec } \sum d_{ij} = n_i)$$

de boules de chaque catégorie qu'il devra contenir.

On calculera ici encore la probabilité cherchée en suivant directement le nombre des cas possibles et le nombre des cas favorables.

a) Nombre des cas possibles Le nombre des cas possibles correspond au nombre de manières différentes de prélever k échantillons d'effectifs n_1, n_2, \dots, n_k dans une urne d'effectif N . Ce nombre est égal au nombre de combinaisons de N individus différents, répartis en k groupes d'effectifs n_1, n_2, \dots, n_k , soit:

$$N!$$

$$n_1! n_2! \dots n_k!$$

b) Nombre de cas favorables Le nombre des cas favorables correspond au nombre de manières différentes qu'il existe de prélever dans l'urne k échantillons dont les compositions correspondent aux compositions spécifiées à l'avance par l'ensemble des valeurs d_{ij} . On pourra construire l'ensemble de ces échantillons de la manière suivante :

1) On répartira de toutes les manières possibles les D_p pièces de catégorie p en k groupes d'effectifs $d_{1p}, d_{2p}, \dots, d_{kp}$, correspondant aux k échantillons à constituer.

Il y a :

$$\frac{D_e!}{d_{1e}! d_{2e}! \dots d_{ke}!}$$

manières d'effectuer cette répartition

2°) Ce processus devra être répété pour les q catégories de boules. Le nombre des cas favorables est, par suite, égal au produit :

$$\prod_e \left\{ \frac{D_e!}{d_{1e}! d_{2e}! \dots d_{ke}!} \right\}$$

La probabilité cherchée a donc pour expression :

$$P = \frac{\prod_e \left\{ \frac{D_e!}{d_{1e}! d_{2e}! \dots d_{ke}!} \right\}}{N! n_1! n_2! \dots n_k!}$$

ou, sous une forme plus condensée :

$$(V-03) P = \frac{1}{N!} \times \frac{\prod_i (n_i!) \cdot \prod_j (D_j!) }{\prod_{ij} (d_{ij}!)}$$

qui fait apparaître les rôles symétriques joués par les catégories de boules et par les échantillons.

Loi des Épreuves rejétées

La probabilité de réalisation p d'un événement à l'atome A demeure constante au cours d'une série d'épreuves successives, on dit que cette série d'épreuves constitue, vis-à-vis de l'événement A, une suite d'épreuves rejettées. Nous nous proposons, dans ces conditions, de calculer la probabilité

$$x \mid p^n$$

pour que l'événement A se réalise exactement x fois au cours de n épreuves successives.

V-04 Expression de la loi de Bernoulli, ou loi des épreuves rejettées

l'événement qui nous intéresse correspond à la réalisation de l'une quelconque des suites chronologiques dans lesquelles l'événement A, de probabilité p , s'est réalisé x fois et l'événement complémentaire \bar{A} , de probabilité $q = 1 - p$, s'est réalisé $n-x$ fois. Dans ces conditions on peut remarquer :

- 1) que les résultats d'une série d'épreuves rejettées constituent des événements aléatoires indépendants, la probabilité de réalisation de l'une quelconque de ces suites a pour valeur :

$$p^x q^{n-x}$$

- 2) que ces suites, qui diffèrent entre elles par l'ordre de réalisation des événements A et \bar{A} qu'elles comportent, forment donc un ensemble d'événements aléatoires exclusifs.

- 3) que le nombre de suites chronologiquement distinctes qui répondent au problème posé correspond au nombre C_n^x de manières d'attribuer un rang aux x événements A parmi les n rangs possibles que comporte la suite des épreuves effectuées.

Par application du principe des probabilités totales à l'ensemble des suites exclusives ainsi dénombrées, on en déduit que la probabilité cherchée a pour expression :

$$(V-04) \quad x \mid p^n = C_n^x p^x q^{n-x} \quad \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

On peut remarquer que la probabilité $x | p^n$ constitue le terme général du développement du binôme:

$$(p+q)^n = p^n + np^{n-1}q + \dots + C_n^x p^x q^{n-x} + \dots + npq^{n-1} + q^n = 1$$

ce qui justifie le nom de "Loi Binomiale" souvent donné à la loi de Bernoulli.

Comme $p+q=1$ on constate, d'après le développement précédent, que la somme des probabilités correspondant à toutes les valeurs possibles de la variable x est bien égale à l'unité.

Application numérique.

En supposant qu'il naîsse autant de garçons que de filles, calculer la probabilité de rencontrer, dans une même famille, soit 3 garçons et 3 filles, soit 6 enfants du même sexe.

La probabilité de rencontrer, dans une même famille, 3 garçons et 3 filles a pour valeur:

$$C_6^3 \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^3 = \frac{5}{16} = 0,3125$$

La probabilité d'avoir 6 enfants du même sexe correspond, soit à avoir 6 garçons, soit à avoir 6 filles. Elle a donc pour valeur:

$$2 \times \left(\frac{1}{2}\right)^6 = \frac{1}{32} = 0,03125$$

On voit que, parmi les familles de 6 enfants, celles qui possèdent 6 enfants du même sexe sont rares. Malgré cela on doit cependant admettre, dans le cadre de notre hypothèse initiale (égalité de probabilité de naissance des garçons et des filles) qu'une mère de 5 filles a encore une chance sur deux de donner naissance à une sixième fille, sous réserve que les naissances successives puissent être considérées comme indépendantes.

V-05. Utilisation de la loi de Bernoulli comme expression approchée d'un tirage exhaustif.

Nous avons établi au paragraphe V-01 l'expression rigoureuse de la probabilité correspondant au tirage

d'un échantillon de composition donnée dans une population de structure connue. La formule rigoureuse à laquelle nous étions arrivés, qui peut s'écrire

$$x \left| \begin{matrix} n \\ \frac{n}{N} \end{matrix} \right. = \frac{\binom{n}{x} \binom{n-x}{N-x}}{\binom{n}{N}} \quad (\text{formule exacte})$$

se prête assez mal aux calculs numériques.

A titre d'approximation on peut lui substituer le terme correspondant de la loi de Bernoulli, soit

$$x \left| \begin{matrix} n \\ p = \frac{x}{N} \end{matrix} \right. = \binom{n}{x} p^x (1-p)^{n-x} \quad (\text{formule approchée})$$

expression dans laquelle on a posé $p = \frac{x}{N}$.

L'approximation obtenue est excellente tant que l'effectif n de l'échantillon prélevé ne dépasse pas l'xième de l'effectif du lot ($n \leq 0,10 N$). A titre de justification on peut remarquer que, dans ce cas, la proportion de pièces défectueuses dans le lot demeure sensiblement constante au cours de l'échantillonnage. On peut alors sans grande erreur, assimiler à une série d'épreuves répétées les prélevements successifs de chacune des pièces de l'échantillon.

v - 06 Fonction caractéristique de la loi de Bernoulli

La première fonction caractéristique de la loi de Bernoulli a pour expression :

$$\Phi_x(u) = E[e^{iux}] = \sum_{x=0}^{x=n} e^{iux} \binom{n}{x} p^x q^{n-x}$$

qui peut s'écrire :

$$\Phi_x(u) = \sum_{x=0}^{x=n} \binom{n}{x} (pe^{iu})^x q^{n-x} = (pe^{iu} + q)^n$$

En remplaçant q par $1-p$ il vient, en définitive:

V-11

$$(\underline{V}-05) \quad \Phi_x(u) = \left\{ 1 + p(e^{iu} - 1) \right\}^n$$

Remarque. Il est possible de calculer directement la fonction caractéristique de la loi de Bernoulli sans utiliser l'expression de la probabilité x/n que nous avons établie au paragraphe V-04.

Considérons, en effet, une variable binaire x susceptible, à chaque épreuve, de prendre la valeur 1 lorsque l'événement A est réalisé et la valeur 0 dans le cas contraire. La variable y est une variable aléatoire dont la loi de probabilité est donnée par le tableau suivant :

Événement réalisé	A	\bar{A}
valeur de la variable y	1	0
Probabilité de réalisation	p	q

La fonction caractéristique de la variable y a, par suite, pour expression :

$$\Phi_y(u) = p e^{iu \cdot 1} + q e^{iu \cdot 0} = p e^{iu} + q$$

Si l'on remarque :

1) que l'on doit considérer comme des variables indépendantes les n variables y_1, y_2, \dots, y_n correspondant à chacune des épreuves d'une série de n épreuves rejétées.

2) que le nombre de fois où l'événement A s'est réalisé au cours de ces n épreuves est égal à la somme des valeurs y_1, y_2, \dots, y_n prises par la variable y à chaque épreuve :

$$x = \sum_{i=1}^{i=n} y_i$$

on en conclut que la fonction caractéristique

de la variable x est égal au produit des n fonctions caractéristiques des variables y_i , soit ici :

$$\Phi_x(u) = (pe^{iu} + q)^n$$

V-07 Premiers moments de la loi de Bernoulli.

Pour calculer les valeurs des premiers moments de la loi de Bernoulli nous utiliserons le développement en série de la seconde fonction caractéristique de cette loi :

$$\psi_x(u) = \log_e \Phi_x(u)$$

Il y a après le résultat obtenu au paragraphe précédent, on peut écrire :

$$\psi_x(u) = \log_e \{ 1 + p(e^{iu} - 1) \}^n = n \log_e \{ 1 + p(e^{iu} - 1) \}$$

soit encore

$$(V-06) \quad \psi_x(u) = n \left\{ p(e^{iu} - 1) - \frac{1}{2} p^2 (e^{iu} - 1)^2 + \frac{1}{3} p^3 (e^{iu} - 1)^3 + \dots \right\}$$

En développant en série les termes $e^{iu} - 1$, $(e^{iu} - 1)^2$, ..., on obtient le résultat définitif suivant :

$$(V-07) \quad \psi_x(u) = npiu + \frac{1}{2} npq(iu)^2 + \frac{1}{6} npq(q-p)(iu)^3 + \frac{1}{24} npq(1-6pq)(iu)^4 + \dots$$

Si l'on compare ce résultat au développement général établi au paragraphe I-20, soit :

$$\psi(u) = m_1 iu + \frac{1}{2} \mu_2 (iu)^2 + \frac{1}{3!} \mu_3 (iu)^3 + \dots$$

on obtient, en identifiant terme à terme les coefficients des puissances égales de u dans les deux développements, les expressions suivantes :

$$(V-08) \quad \begin{cases} m_1 = E[x] = np \\ \mu_2 = \sigma_x^2 = npq \\ \mu_3 = npq(q-p) \end{cases}$$

On peut en déduire les valeurs des moments de la fréquence $f_n = \frac{x}{n}$ de réalisation de l'événement A au cours d'une série de n épreuves répétées. Il est utile de retenir les deux résultats suivants :

Variable x:	Variable f_n
x nombre de succès	$f_n = \frac{x}{n}$ fréquence des succès
$E[x] = np$	$E[f_n] = p$
$\sigma_x = \sqrt{npq}$	$\sigma_{f_n} = \sqrt{\frac{pq}{n}}$

V-08 Fonction de répartition de la loi de Bernoulli.

La fonction de répartition de la loi de Bernoulli correspond à la somme des $n+1$ premiers termes du développement du binôme $(q+p)^n$, soit :

$$(V-09) \quad F(x) = q^n + nq^{n-1}p + \frac{n(n-1)}{2}q^{n-2}p^2 + \dots + \sum_n^x q^{n-x}p^x$$

La fonction $F(x)$ représente donc la probabilité totale pour que l'événement A (de probabilité p) se réalise au maximum x fois au cours de n épreuves.

Nous nous proposons de démontrer que la probabilité totale $F(x)$ a pour valeur :

$$(V-10) \quad F(x) = 1 - \frac{B_p\{x+1, n-x\}}{B_1\{x+1, n-x\}}$$

formule dans laquelle $B_p(a, b)$ désigne la fonction

En l'espèce incomplète de première espèce, ou
fonction Béta incomplète et $B_1(a, b)$ désigne
la fonction En l'espèce complète de 1^{re} espèce,
dont les définitions sont rappelées ci-dessous:

$$(V-11) \quad B_p(a, b) = \int_0^p z^{a-1} (1-z)^{b-1} dz$$

$$(V-12) \quad B_1(a, b) = \int_0^1 z^{a-1} (1-z)^{b-1} dz$$

Des tables de la fonction Béta incomplète
pourront, en particulier, être consultées dans les
publications suivantes:

PEARSON, K. : Tables of the Incomplete Beta Function,
Biometrika Office, University College,
London 1934

THOMPSON, C. M. : Tables of percentage points of the
incomplete Beta Function, Biome-
trika, vol. xxxii, (1941) p.p. 151-181

D'après (V-11) la fonction $B_p\{x+1, n-x\}$ a pour
expression:

$$B_p\{x+1, n-x\} = \int_0^p z^x (1-z)^{n-x-1} dz$$

En posant $u = (1-z)^{n-x-1}$ et $dv = z^x dz$,
on obtient, en intégrant par parties:

$$(V-13) \quad B_p\{x+1, n-x\} = \frac{x}{n-x} \left\{ B_p\{x, n-(x-1)\} - \frac{p^x q^{n-x}}{x} \right\}$$

on a formule dans laquelle, comme précédemment,
on a posé $q = 1 - p$.

Par calcul direct on trouve pour $x=0$:

$$B_p\{0+1, n-0\} = \int_0^p (1-z)^{n-1} dz = \frac{1}{n} \{1 - q^n\}$$

En appliquant successivement la formule de
référence (V-13) pour $x=1, x=2 \dots$ on trouve :

Pour $x=1$ $B_p\{1, n-1\} = \frac{1}{n-1} \left\{ \frac{1}{m} [1 - q^n] - pq^{n-1} \right\}$

$$B_p\{2, n-1\} = \frac{1}{n(n-1)} \left\{ 1 - (q^n + nq^{n-1}p) \right\}$$

Pour $x=2$ $B_p\{3, n-2\} = \frac{2}{n-2} \left\{ \frac{1}{n(n-1)} [1 - (q^n + nq^{n-1}p)] - \frac{p^2 q^{n-2}}{2} \right\}$

$$B_p\{3, n-2\} = \frac{2}{n(n-1)(n-2)} \left\{ 1 - (q^n + nq^{n-1}p + \frac{n(n-1)}{2} q^{n-2} p^2) \right\}$$

formule que l'on peut écrire :

$$B_p\{3, n-2\} = \frac{2! (n-3)!}{n!} \left\{ 1 - (q^n + nq^{n-1}p + \frac{n!}{2!(n-2)!} q^{n-2} p^2) \right\}$$

Supposons la formule exacte pour $x-1$ et démontrons
qu'elle est encore vraie pour x . Nous partirons donc
de la relation

Pour $x-1$: $B_p\{x, n-x-1\} = \frac{(x-1)! (n-x)!}{n!} \left\{ 1 - (q^n + \dots + \frac{n! q^{n-x-1} p^{x-1}}{(x-1)! (n-(x-1))!}) \right\}$

En appliquant à nouveau la formule de récurrence (V-13)
il vient :

$$\underline{B}_p \{x+1, n-x\} = \frac{x}{n-x} \left\{ \frac{(x-1)!(n-x)!}{n!} \left[1 - \left(q^n + \dots + \frac{n! q^{n-x-1} p^{x-1}}{(x-1)![n-x+1]!} \right) \right] - \frac{q^{n-x} p^x}{x} \right\}$$

formule que l'on peut écrire :

$$(V-14) \quad \underline{B}_p \{x+1, n-x\} = \frac{x!(n-x-1)!}{n!} \left\{ 1 - \left(q^n + \dots + \frac{n!}{x!(n-x)} q^{n-x} p^x \right) \right\}$$

qui démontre que l'expression précédente est générale et valable pour toute valeur entière de x comprise entre 0 et $n-1$, bornes comprises.

En faisant $p=1$ (et donc $q=0$) dans la formule (V-14), on obtiendra la valeur de la fonction Beta complète :

$$(V-15) \quad \underline{B}_1 \{x+1, n-x\} = \frac{x!(n-x-1)!}{n!}$$

La formule (V-14) peut alors s'écrire :

$$\underline{B}_p \{x+1, n-x\} = \underline{B}_1 \{x+1, n-x\} \left\{ 1 - \left(q^n + nq^{n-1} p + \dots + {}^n C_m q^{n-x} p^x \right) \right\}$$

soit encore :

$$\underline{B}_p \{x+1, n-x\} = \underline{B}_1 \{x+1, n-x\} \{ 1 - F(x) \}$$

relation qui n'est autre que la formule (V-10) que nous voulions démontrer.

V - 09. Généralisation de la Loi de Bernoulli.

On peut généraliser la loi des épreuves répétées en supposant qu'à chaque épreuve doit obligatoirement se réaliser l'un quelconque des événements aléatoires :

$$A_1, A_2, \dots, A_k$$

dont l'ensemble constitue la totalité des cas possibles. Nous supposons, de plus, que les événements considérés sont exclusifs les uns des autres, de sorte que seul l'un d'eux puisse se réaliser à chaque épreuve. Dans ces conditions, les probabilités p_1, p_2, \dots, p_k de réalisation des événements A_i vérifient la relation

$$(V-16) \quad p_1 + p_2 + \dots + p_k = 1$$

Lorsque les probabilités p demeurent constantes au cours d'une série d'épreuves successives, on dit que cette série d'épreuves constitue une suite d'épreuves répétées.

Nous nous proposons, dans ces conditions, de calculer la probabilité P pour que :

$$\left. \begin{array}{l} \text{l'événement } A_e \text{ se réalise exactement } x_e = \alpha \text{ fois} \\ \text{l'événement } A_f \quad " \quad " \quad x_f = \beta \text{ fois} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \text{l'événement } A_h \quad " \quad " \quad x_h = \lambda \text{ fois} \end{array} \right\}$$

au cours d'une série de n épreuves successives.

Nous remarquons :

1) que l'ensemble des événements A_1, A_2, \dots, A_k constituant la totalité des cas possibles, les nombres $\alpha, \beta, \dots, \lambda$ des réalisations des événements observés au cours de n épreuves vérifient la relation :

$$\alpha + \beta + \dots + \lambda = n$$

2) que l'événement qui nous intéresse correspond à la réalisation de l'une quelconque des suites chronologiques d'épreuves comportant la réalisation

un nombre imposé de fois (respectivement $\alpha, \beta, \dots, \lambda$ fois), des événements A_e, A_f, \dots, A_h , choisis parmi les n événements possibles.

3) que la probabilité de réalisation de l'une quelconque de ces suites a pour valeur:

$$p_e^\alpha p_f^\beta \cdots p_h^\lambda$$

4) que ces suites, qui diffèrent entre elles par l'ordre de réalisation des événements qui les composent, forment un ensemble d'événements exclusifs.

5) que le nombre des suites chronologiques distinctes qui répondent au problème posé correspond au nombre de manière de ranger n événements en catégories comportant respectivement $\alpha, \beta, \dots, \lambda$ d'entre eux. On sait que ce nombre est égal à

$$\frac{n!}{\alpha! \beta! \cdots \lambda!}$$

On en déduit que la probabilité P cherchée a pour valeur:

$$(V-17) \quad P = \frac{n!}{\alpha! \beta! \cdots \lambda!} p_e^\alpha p_f^\beta \cdots p_h^\lambda$$

expression qui constitue l'un des termes du développement du polynôme:

$$(V-18) \quad (p_1 + p_2 + \cdots + p_n)^n = \sum_i p_i^n + n \sum_{fg} p_f^{\alpha-1} p_g^{\beta-1} + \cdots + \frac{n!}{\alpha! \beta! \cdots \lambda!} \sum_{e f \cdots h} p_e^\alpha p_f^\beta \cdots p_h^\lambda$$

V-10. Fonction caractéristique de la loi généralisée de Bernoulli

Les fréquences absolues x_1, x_2, \dots, x_k de réalisation des événements A_1, A_2, \dots, A_k au cours d'une suite de n épreuves répétées sont des nombres entiers nuls ou positifs qui doivent, de plus, vérifier la relation:

$$(V-19) \quad x_1 + x_2 + \dots + x_k = n$$

La fonction caractéristique

$$\Phi_{\vec{x}}(u_1, u_2, \dots, u_k) = E \left[e^{i(u_1 x_1 + u_2 x_2 + \dots + u_k x_k)} \right]$$

correspondant à la distribution de Bernoulli s'exprimera donc par une somme de termes correspondant à toutes les combinaisons de valeurs des variables x_i vérifiant la relation (V-19). En conservant les notations du paragraphe précédent, l'un de ces termes aura pour expression:

$$\frac{n!}{\alpha! \beta! \dots \lambda!} p_e^\alpha p_f^\beta \dots p_n^\lambda e^{i(\alpha u_e + \beta u_f + \dots + \lambda u_n)}$$

que l'on peut encore écrire sous la forme:

$$\frac{n!}{\alpha! \beta! \dots \lambda!} (p_e e^{iu_e})^\alpha (p_f e^{iu_f})^\beta \dots (p_n e^{iu_n})^\lambda$$

sous laquelle on reconnaît l'un des termes du développement du polynôme:

$$(V-20) \quad \Phi_{\vec{x}}(u_1, u_2, \dots, u_k) = (p_1 e^{iu_1} + p_2 e^{iu_2} + \dots + p_k e^{iu_k})^n$$

qui est l'expression de la fonction caractéristique cherchée.

Remarque Si l'on fait $u_1 = u_2 = \dots = u_k = u$ dans l'expression (V-20) on trouve:

$$\Phi_{\vec{x}}(u) = \{(p_1 + p_2 + \dots + p_k) e^{iu}\}^n = e^{iun}$$

qui est bien la fonction caractéristique de la constante $x_1 + x_2 + \dots + x_k = n$

I-21 Matrice des variances de la loi de Bernoulli.

En tenant compte de la relation :

$$p_1 + p_2 + \dots + p_k = 1$$

la fonction caractéristique (I-20) de la loi de Bernoulli peut se mettre sous la forme :

$$(I-21) \quad \Phi_{\vec{X}}(u_1, u_2, \dots, u_k) = \left\{ 1 + \sum_{g=1}^{g=k} p_g (e^{iu_g} - 1) \right\}^n$$

qui correspond, pour la deuxième fonction caractéristique, à l'expression :

$$(I-22) \quad \Psi_{\vec{X}}(u_1, u_2, \dots, u_k) = n \log_e \left\{ 1 + \sum_{g=1}^{g=k} p_g (e^{iu_g} - 1) \right\}$$

La seconde fonction caractéristique admet un développement en série dont les premiers termes ont pour valeurs :

$$(I-23) \quad \Psi_{\vec{X}}(u_1, \dots, u_k) = \left\{ \begin{array}{l} i \sum_f n p_f u_f \\ - \frac{1}{2} \left\{ \sum_f n p_f (1-p_f) u_f^2 - 2 \sum_{fg} n p_f p_g u_f u_g \right\} \\ + \frac{i^3}{6} \left\{ \sum_f n p_f (1-p_f) (1-2p_f) u_f^3 - 3 \sum_{fg} n p_f (1-2p_f) u_f^2 u_g \right. \\ \left. + 12 \sum_{fgh} p_f p_g p_h u_f u_g u_h \right\} \\ + \dots \end{array} \right\}$$

On en déduit immédiatement les expressions suivantes :

$$(V-24) \quad \left\{ \begin{array}{l} E[x_f] = n p_f \\ \sigma_f^2 = n p_f (1-p_f) \\ \text{cov}(x_f x_g) = -n p_f p_g \\ P_{fg} = -\sqrt{\frac{p_f p_g}{(1-p_f)(1-p_g)}} \end{array} \right.$$

La matrice des variances s'écrit :

$$(V-25) \quad [V] = \begin{bmatrix} np_1(1-p_1) & -np_1p_2 & -np_1p_k \\ -np_2p_1 & np_2(1-p_2) & -np_2p_k \\ -np_kp_1 & -np_kp_2 & np_k(1-p_k) \end{bmatrix}$$

Le déterminant correspondant a pour expression :

$$(V-26) \quad |V| = n^k p_1 p_2 \cdots p_k \begin{vmatrix} 1-p_1 & -p_2 & -p_k \\ -p_1 & 1-p_2 & -p_k \\ -p_1 & -p_2 & 1-p_k \end{vmatrix}$$

Ce déterminant est nul, comme il est facile de le montrer en ajoutant à la première colonne toutes les autres colonnes et en tenant compte du fait que $\sum p_i = 1$:

$$|V| = n^{k \choose 2} p_1 p_2 \dots p_k \begin{vmatrix} 1-p_1 & -p_2 & -p_k \\ -p_1 & 1-p_2 & -p_k \\ -p_1 & -p_2 & 1-p_k \end{vmatrix} = n^{k \choose 2} p_1 p_2 \dots p_k \begin{vmatrix} 0 & -p_2 & -p_k \\ 0 & 1-p_2 & \\ 0 & -p_2 & 1-p_k \end{vmatrix}$$

La matrice V , d'ordre n est, par suite, une matrice irrégulière, dont le rang est, au maximum égal à $n-1$.

Ceci provient du fait que le point de coordonnées $x_1 x_2 \dots x_k$ est astreint à demeurer dans un hyperplan à $n-1$ dimensions car ses coordonnées doivent vérifier la relation (V-19) qui est d'ailleurs l'équation de cet hyperplan.

Le rang de la matrice des variances est effectivement égal à $n-1$ lorsque les probabilités p_1, p_2, \dots, p_k sont toutes différentes de zéro.

C. Loi des Grandes Mâmes.

Nous nous proposons de démontrer qu'il existe une forme limite de la Loi de Bernoulli lorsque le nombre des épreuves répétées devient suffisamment grand. Nous examinerons tout d'abord qualitativement la manière dont se modifie dans ce cas le diagramme de fréquence de la Loi de Bernoulli. Nous déduirons de cette étude le changement de variable qu'il faut effectuer pour trouver la forme limite de la distribution et nous étendrons enfin les résultats obtenus au cas où chaque épreuve peut donner lieu à plusieurs événements aléatoires possibles.

V - 12. Diagramme de fréquence de la loi de Bernoulli.

Le diagramme de fréquence de la loi de Bernoulli est un diagramme en bâtons égaux dont les hauteurs croissent (Fig V-01) jusqu'au moment où la variable x atteint le nombre entier immédiatement inférieur ou égal à son espérance mathématique $E[x] = np$.

La variable x peut prendre toutes les valeurs entières comprises entre 0 et n inclusivement et la somme des hauteurs de tous les bâtons correspondants est égale à l'unité.

Lorsque le nombre n des épreuves augmente, le nombre des valeurs possibles de la variable x s'accroît d'une manière identique. Il en résulte que la hauteur de chacun des bâtons doit tendre vers zéro, puisque leur nombre augmente indéfiniment alors que leur somme doit toujours demeurer égale à l'unité.

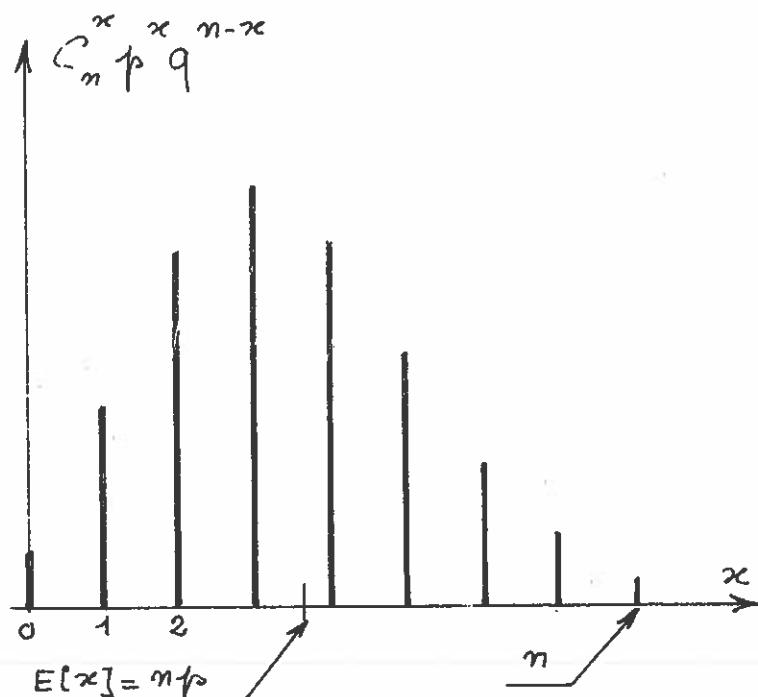


Fig. V - 01

La distribution s'aplatit et s'étale sur l'axe des x , sans tendre vers aucune limite fixe.

Pour trouver une distribution limite il faut, par le choix d'une variable appropriée, s'efforcer à cet étalement tout en maintenant la partie centrale de la distribution au voisinage de l'origine. On peut obtenir ce double résultat en adoptant, comme nouvelle variable, la variable réduite :

$$(V-27) \quad t = \frac{x - E[x]}{\sigma_x} = \frac{x - np}{\sqrt{npq}}$$

dont l'espérance mathématique et l'écart-type demeurent finis (respectivement 0 et 1) lorsque le nombre n des épreuves augmente indéfiniment.

V-13. Fonction caractéristique de la variable réduite de Bernoulli.

La variable réduite de Bernoulli pouvant s'écrire sous la forme :

$$t = -\sqrt{\frac{np}{q}} + \frac{x}{\sqrt{npq}}$$

La seconde fonction caractéristique de la variable t se déduit de la deuxième fonction caractéristique de la variable x par application de la formule (paragraphe I-21) :

$$\psi_t(u) = -iu\sqrt{\frac{np}{q}} + \psi_x\left(\frac{u}{\sqrt{npq}}\right)$$

soit, en se reportant à la formule (V-07) :

$$\psi_t(u) = -iu\sqrt{\frac{np}{q}} + \frac{npiu}{\sqrt{npq}} + \frac{npq(iu)^2}{2npq} + \frac{npq(q-p)(iu)^3}{6npq\sqrt{npq}} + \frac{npq(1-6pq)(iu)^4}{24(npq)^2} + \dots$$

ou, en définitive :

$$(V-28) \quad \psi_t(u) = -\frac{1}{2}u^2 + \frac{1}{3!}\frac{q-p}{\sqrt{n}pq}(iu)^3 + \frac{1}{4!}\frac{1-6pq}{n^2pq}(iu)^4 + \dots$$

V - 14. Forme limite de la loi réduite de Bernoulli, lorsque le nombre des épreuves augmente indéfiniment.

Lorsque le nombre n des épreuves augmente indéfiniment, la formule (V-28) montre que la seconde fonction caractéristique tend vers $-\frac{1}{2}u^2$:

$$\psi_E(u) \xrightarrow{n \rightarrow \infty} -\frac{1}{2}u^2$$

c'est à dire vers la seconde fonction caractéristique de la loi normale réduite. On peut énoncer ce résultat sous la forme suivante, comme sous le nom de "Loi des Grands Nombres":

Lorsque le nombre des épreuves augmente indéfiniment, la loi de répartition de la variable réduite de Bernoulli tend vers la loi de répartition de la variable normale réduite ou loi de Gauss réduite.

Lorsque le nombre des épreuves n'est pas très grand, la loi de Gauss ne constitue qu'une approximation de la loi de Bernoulli. Il résulte de la formule (V-28) que cette approximation est d'autant plus valable que la quantité $\sqrt{n}pq$ est, elle-même, plus grande. Le nombre d'épreuves doit donc être d'autant plus élevé que l'une des probabilités p ou q est plus voisine de 0.

En pratique, on peut considérer l'approximation normale comme valable dès que $\sqrt{n}pq \geq 5$. Dans ces conditions les écarts entre la fonction de répartition de la loi de Bernoulli et celle de la loi normale restent toujours inférieurs à 0,02.

V - 15 Exemple numérique.

La loi des Grands Nombres apporte une simplification considérable à l'exécution des calculs numériques relatifs aux séries d'épreuves séparées, comme le montre l'exemple suivant.

Proposons-nous de calculer la probabilité d'obtenir exactement 50 succès au cours d'une partie de "pile ou face" comportant 100 coups.

En admettant que la pièce est parfaitement symétrique, la valeur exacte de cette probabilité a pour expression :

$$P\{x=50 \mid n=100, p=\frac{1}{2}\} = C_{100}^{50} \left(\frac{1}{2}\right)^{50} \times \left(\frac{1}{2}\right)^{50} = \frac{100!}{2^{100} (50!)^2}$$

On ne peut effectuer ce calcul d'une manière précise que si l'on dispose d'une table des logarithmes des factorielles. On trouve ainsi :

$$\begin{array}{rcl} \log(100!) & = & \dots + 157,970\,0037 \\ -\log(50!)^2 & = & -2 \times 64,4830749 = \dots - 128,966\,1498 \\ \hline \text{Différence :} & & + 29,003\,8539 \\ \text{colog } 2^{100} & & \overline{31,897\,00} \\ \text{Somme :} & & \overline{\overline{2,900\,85}} \end{array}$$

soit, pour la probabilité cherchée :

$$\underline{\text{Valeur exacte :}} \quad C_{100}^{50} \times \left(\frac{1}{2}\right)^{100} = 0,079589$$

On peut obtenir plus facilement une valeur approchée de cette probabilité en utilisant la Loi des Grandes Nombres, dont les paramètres ont ici pour valeurs :

$$\begin{cases} E[x] = np = 100 \times \frac{1}{2} = 50 \\ \sigma_x = \sqrt{npq} = \sqrt{100 \times \frac{1}{4}} = 5 \end{cases}$$

Il subvient toutefois une difficulté à l'application immédiate de la loi des Grands Nombres car on substitue une variable normale continue à une variable de Bernoulli qui, par nature, est essentiellement disccontinue.

D'après le paragraphe précédent, on sait que les fonctions de répartition de ces deux variables sont proches l'une de l'autre.

Comme l'indique la fig. V-02, tracée schématiquement au voisinage de $x = 50$ pour l'une et l'autre des deux distributions, la probabilité cherchée est rigoureusement égale au restant AB correspondant à la valeur $x = 50$ de la combe de répartition disccontinue de la loi de Bernoulli.

On voit que l'on obtiendra une valeur approchée de cette probabilité en substituant au restant AB la différence :

$$\Delta F = F(50,5) - F(49,5)$$

correspondant aux valeurs de la fonction de répartition continue de la variable de Gauss pour les valeurs $x = 49,5$ et $x = 50,5$ de la variable x .

Cette méthode revient, sur le diagramme de distribution de la variable continue x (combe en cloche de Gauss, fig V-03)

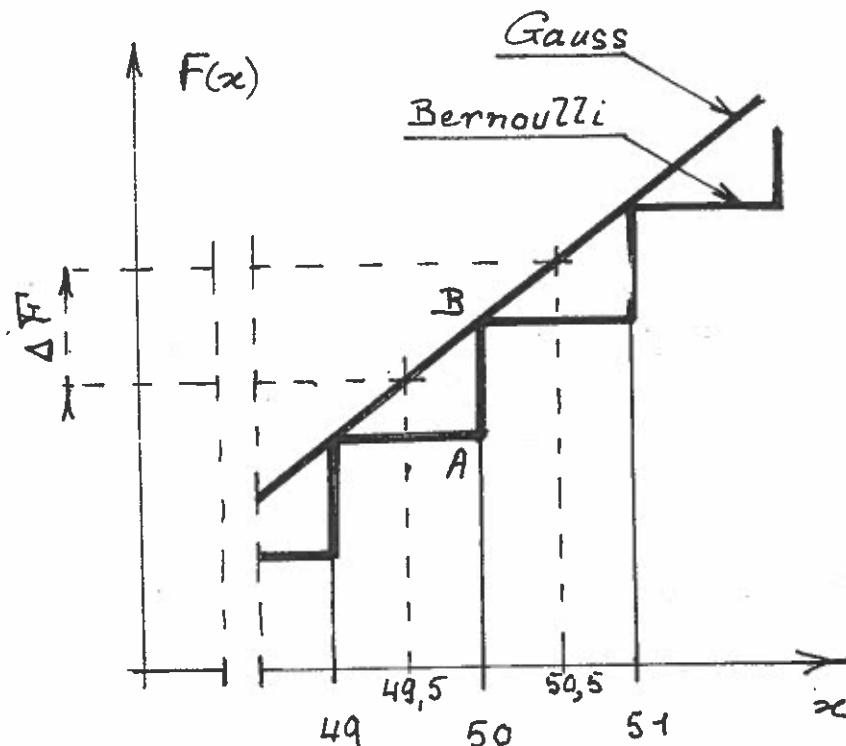


Fig. V-02

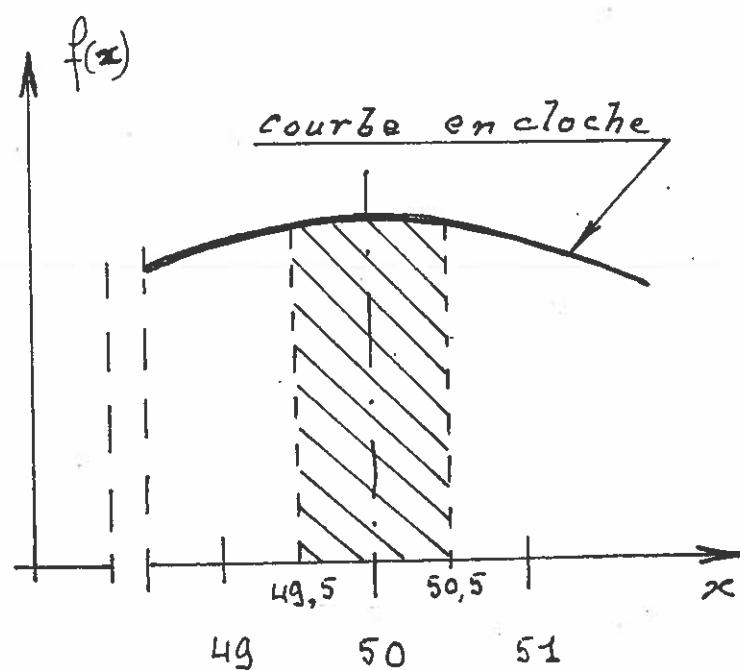


Fig V-03

V - 28

à attribuer à la valeur $x = 50$ une probabilité mesurée par l'aire du trapèze curviligne de base $(49,5, 50,5)$.

En tenant compte des valeurs $E[x] = 50$ et $\sigma_x = 5$ calculées précédemment, on détermine les valeurs t_1 et t_2 de la variable réduite qui correspondent aux valeurs $x = 49,5$ et $x = 50,5$, Soient :

$$\left\{ \begin{array}{l} t_1 = \frac{49,5 - 50}{5} = -0,1 \\ t_2 = \frac{50,5 - 50}{5} = +0,1 \end{array} \right.$$

et, par suite :

$$\Delta F = F(+0,1) - F(-0,1) = 2 \times F(+0,1) - 1$$

relation dans laquelle F représente la fonction de répartition de la variable de Gauss réduite. En utilisant une table plus précise que la table II, on obtient la valeur :

$$F(+0,1) = 0,5398278$$

et, par suite :

$$\Delta F = 1,0796556 - 1 = 0,0796556$$

valeur approchée très voisine de la valeur exacte trouvée précédemment.⁽¹⁾

On mesurera encore mieux l'intérêt que présente la loi des Grands Nombres, en cherchant à répondre à la question complémentaire suivante :

⁽¹⁾ Designant toujours par x le nombre de succès

(1) Rem. On peut également adopter, avec une approximation encore très satisfaisante, la valeur

$$\frac{1}{5\sqrt{2\pi}} = 0,079788$$

correspondant à l'aire du rectangle de base $(49,5, 50,5)$ et dont la hauteur est égale à l'ordonnée au sommet $1/\sigma\sqrt{2\pi}$ de la courbe de Gauss.

- au cours d'une partie de 100 coups à pile ou face on demande de déterminer dans quel intervalle doit se situer ce nombre x dans 998 pour mielle des parties jouées."

La réponse à cette question est, en effet, immédiate, lorsque l'on peut valablement utiliser la Loi des Grands Nombres. On sait, dans ces conditions, qu'il y a 998 chances sur 1000 pour que la variable x , d'espérance mathématique $E[x] = 50$ et d'écart-type $\sigma_x = 5$, soit située dans l'intervalle:

$$50 \pm 3,09 \times 5 = (34,55, 65,45)$$

Il y a donc 998 chances sur 1000 pour que le nombre de succès d'une partie de 100 coups à pile ou face soit compris dans l'intervalle $(35, 65)$, bornes comprises.

A titre de confirmation, on peut retrouver ce résultat d'une manière directe, par application de la formule de Bernoulli

$$P\{x \mid \begin{matrix} n=100 \\ p=\frac{1}{2} \end{matrix}\} = C_{100}^x \times \left(\frac{1}{2}\right)^{100}$$

aux valeurs successives de x comprises entre 0 et 35. On trouve ainsi:

Pour $x \leq 22$

$$P(x \mid \frac{1}{2}) \leq 2 \times 10^{-8}$$

Pour $x = 23$

$$P(x) = 2 \times 10^{-8}$$

" " 24

" " 25

" " 26

" " 27

" " 28

" " 29

" " 30

" " 31

" " 32

" " 33

" " 34

" " 35

$$P(x) = 6 \times 10^{-8}$$

$$P(x) = 19 \times 10^{-8}$$

$$P(x) = 55 \times 10^{-8}$$

$$P(x) = 151 \times 10^{-8}$$

$$P(x) = 394 \times 10^{-8}$$

$$P(x) = 979 \times 10^{-8}$$

$$P(x) = 2317 \times 10^{-8}$$

$$P(x) = 5232 \times 10^{-8}$$

$$P(x) = 11282 \times 10^{-8}$$

$$P(x) = 23247 \times 10^{-8}$$

$$P(x) = 45815 \times 10^{-8}$$

$$P(x) = 86385 \times 10^{-8}$$

On en déduit :

$$\left\{ \begin{array}{l} x = 65 \\ \sum_{x=35}^{x=65} P(x) = 0,99821 \end{array} \right.$$

$$\left\{ \begin{array}{l} x = 64 \\ \sum_{x=36}^{x=64} P(x) = 0,99648 \end{array} \right.$$

Résultats qui sont en parfait accord avec l'intervalle calculé d'une manière beaucoup plus simple par application de la Loi des Grands Nombres.

V-16. Généralisation de la Loi des Grands Nombres

Nous avons étudié, au paragraphe V-09, une forme généralisée de la Loi de Bernoulli, s'appliquant au cas où l'événement qui peut se produire à chaque épreuve appartient obligatoirement à un ensemble de k événements exclusifs, A_1, A_2, \dots, A_k , de probabilités p_1, p_2, \dots, p_k .

Si l'on désigne, comme précédemment, par x_1, x_2, \dots, x_k les fréquences absolues de réalisation des événements A_1, A_2, \dots, A_k au cours d'une suite de n épreuves, nous nous proposons de démontrer que, lorsque n augmente indéfiniment, la distribution du vecteur \vec{x} , de coordonnées :

$$(V-29) \quad z_f = \frac{x_f - np_f}{\sqrt{n}}$$

Tend vers une distribution normale de Laplace-Gauss, définie dans l'hyperplan à $k-1$ dimensions d'équation

$$(V-30) \quad z_1 + z_2 + \dots + z_k = 0$$

perpendiculaire à la 1^{re} bissectrice.

Mais tout d'abord, que la contrainte (V-30) imposée aux composantes du vecteur \vec{Z} résulte immédiatement des équations (V-16) et (V-19) établies précédemment.

Pour déterminer ensuite la forme limite de la distribution du vecteur \vec{Z} , nous allons montrer que la fonction caractéristique de ce vecteur tend vers la fonction caractéristique d'une distribution normale de Laplace-Gauß.

Le vecteur \vec{Z} se déduit du vecteur \vec{x} par le changement de variable (V-29) que l'on peut écrire sous la forme:

$$\beta_f = \frac{1}{\sqrt{n}} x_f - \sqrt{n} p_f$$

Il en résulte que la fonction caractéristique du vecteur \vec{Z} se déduit de la fonction caractéristique du vecteur \vec{x} par la relation:

$$\Phi_{\vec{Z}}(u_1, u_2, \dots, u_k) = e^{-i\sqrt{n} \sum u_f p_f} \Phi_{\vec{x}}\left(\frac{u_1}{\sqrt{n}}, \frac{u_2}{\sqrt{n}}, \dots, \frac{u_k}{\sqrt{n}}\right)$$

soit encore, pour la seconde fonction caractéristique:

$$\Psi_{\vec{Z}}(u_1, u_2, \dots, u_k) = -i\sqrt{n} \sum u_f p_f + \Psi\left(\frac{u_1}{\sqrt{n}}, \frac{u_2}{\sqrt{n}}, \dots, \frac{u_k}{\sqrt{n}}\right)$$

En tenant compte de la formule (V-23) on obtient ainsi l'expression de la seconde fonction caractéristique du vecteur \vec{Z} sous la forme suivante

(V-31)

$$\psi_{\vec{z}}(u_1, u_2, \dots, u_n) = \left\{ \begin{array}{l} -\frac{1}{2} \left\{ \sum_f p_f (1-p_f) u_f^2 - 2 \sum_{fg} p_f p_g u_f u_g \right\} \\ + \frac{i^3}{6\sqrt{n}} \left\{ \sum_f p_f (1-p_f) (1-2p_f) u_f^3 - 3 \sum_{fg} p_f (1-2p_f) u_f^2 u_g \right. \\ \quad \left. + 12 \sum_{fgh} p_f p_g p_h u_f u_g u_h \right\} \\ + \frac{1}{n} \left\{ \dots \right\} \end{array} \right\}$$

On voit donc que, lorsque $\vec{z}^n \rightarrow \infty$, la seconde fonction caractéristique du vecteur \vec{z} tend vers:

$$(V-32) \quad \psi_{\vec{z}}(u_1, u_2, \dots, u_n) \rightarrow -\frac{1}{2} \left\{ \sum_f p_f (1-p_f) u_f^2 - 2 \sum_{fg} p_f p_g u_f u_g \right\}$$

qui est bien la seconde fonction caractéristique d'une loi normale de Laplace-Gauss, comme nous voulions le démontrer.

V-17. Expression de La Loi des Grands Nombres généralisée.

Nous venons de montrer que la distribution du vecteur \vec{z} tend vers une distribution de Gauss localisée dans un hyperplan perpendiculaire à la première bissectrice. DPT l'en suit que le vecteur \vec{z} ne comporte que $k-1$ composantes indépendantes, que l'on pourrait, par exemple, choisir dans un système de $k-1$ axes perpendiculaires à la première bissectrice.

En lieu d'effectuer un tel changement de variables, nous nous proposons d'utiliser comme

Composantes indépendantes des $k-1$ premières composantes z_1, z_2, \dots, z_{k-1} du vecteur \bar{z} , que l'on peut d'ailleurs considérer comme les composantes d'un vecteur \bar{z}' résultant de la projection, parallèlement à l'axe ox_k , du vecteur \bar{z} sur le sous-espace $x_k = 0$ défini par les $k-1$ premiers axes.

Dans ces conditions, la distribution du vecteur \bar{z}' est une distribution marginale dont la fonction caractéristique se déduit de la fonction caractéristique du vecteur \bar{z} en attribuant une valeur nulle à la composante u_k du vecteur \bar{u} . Lorsque n tendra vers l'infini, la seconde fonction caractéristique du vecteur \bar{z}' tendra donc vers :

$$\Psi_{\bar{z}'}(u_1, u_2, \dots, u_{k-1}) = \Psi_{\bar{z}}(u_1, u_2, \dots, u_{k-1}, 0)$$

soit, d'après (V-32) vers :

$$(V-33) \quad \Psi_{\bar{z}'}(u_1, u_2, \dots, u_{k-1}) = -\frac{1}{2} \sum_{f=1}^{k-1} \left\{ p_f (1-p_f) u_f^2 - 2 \sum_{g=f+1}^{k-1} p_f p_g u_f u_g \right\}$$

On en déduit que les $k-1$ composantes z_1, z_2, \dots, z_{k-1} du vecteur \bar{z}' sont, à la limite, distribuées suivant une loi normale dont la matrice des variances a pour expression :

$$(V-34) \quad [V'] = \begin{bmatrix} p_1(1-p_1) & -p_1 p_2 & \cdots & -p_1 p_{k-1} \\ -p_2 p_1 & p_2(1-p_2) & \cdots & -p_2 p_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{k-1} p_1 & -p_{k-1} p_2 & \cdots & p_{k-1}(1-p_{k-1}) \end{bmatrix}$$

Le calcul du déterminant $|V'|$ correspondant à cette matrice s'effectue comme ceci :

- en mettant tout d'abord en facteur les éléments communs à chacun des termes d'une même ligne.
- en ajoutant ensuite à la première colonne toutes les autres colonnes du déterminant ainsi simplifié.

Il vient :

$$\begin{vmatrix} V' \end{vmatrix} = p_1, p_2, \dots, p_{k-1} \begin{vmatrix} 1-p_1 & -p_2 & -p_{k-1} \\ -p_1 & 1-p_2 & -p_{k-1} \\ -p_1 & -p_2 & 1-p_{k-1} \end{vmatrix}$$

$$\begin{vmatrix} V' \end{vmatrix} = p_1, p_2, \dots, p_{k-1} \begin{vmatrix} 1-\sum p_f & -p_2 & -p_{k-1} \\ 1-\sum p_f & 1-p_2 & -p_{k-1} \\ 1-\sum p_f & -p_2 & 1-p_{k-1} \end{vmatrix}$$

$$\begin{vmatrix} V' \end{vmatrix} = p_1, p_2, \dots, p_{k-1} \left(1 - \sum_{f=1}^{k-1} p_f \right) \begin{vmatrix} 1 & -p_2 & -p_{k-1} \\ 1 & 1-p_2 & -p_{k-1} \\ 1 & -p_2 & 1-p_{k-1} \end{vmatrix}$$

Le déterminant qui figure dans la dernière formule est la somme de déterminants qui sont tous nuls (comme ayant deux colonnes proportionnelles) et d'un déterminant dont la valeur est égale à l'unité.

$$\begin{vmatrix} 1 & -p_2 & -p_{k-1} \\ 1 & 1-p_2 & -p_{k-1} \\ 1 & -p_2 & 1-p_{k-1} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 1$$

En définitive si l'on tient en outre compte de la relation (V-16) qui peut s'écrire:

$$p_k = p_{k-1}$$

$$1 - \sum_{f=1}^{k-1} p_f = p_k$$

le déterminant $|V'|$ a pour valeur:

$$(V-35) \quad |V'| = p_1 p_2 p_3 \dots p_k$$

Enfin on vérifie facilement que la matrice inverse de la matrice $[V']$ a pour expression:

$$(V-36) \quad [V']^{-1} = \begin{bmatrix} \frac{1}{p_k} + \frac{1}{p_1} & \frac{1}{p_k} & \frac{1}{p_k} \\ \frac{1}{p_k} & \frac{1}{p_k} + \frac{1}{p_2} & \frac{1}{p_k} \\ \frac{1}{p_k} & \frac{1}{p_k} & \frac{1}{p_k} + \frac{1}{p_{k-1}} \end{bmatrix}$$

V-36

On en déduit immédiatement l'expression de la forme quadratique $\chi^2(k-1)$ qui figure dans l'exponentielle d'une loi de Gauss dans un espace à $k-1$ dimensions, soit :

$$\chi^2(k-1) = \sum_{f=1}^{f=k-1} \left\{ \left(\frac{1}{p_k} + \frac{1}{p_f} \right) z_f^2 + \frac{2}{p_k} \sum_{g=f+1}^{g=k-1} z_f z_g \right\}$$

ou

$$\chi^2(k-1) = \sum_{f=1}^{f=k-1} \frac{z_f^2}{p_f} + \frac{1}{p_k} \sum_{f=1}^{f=k-1} \left\{ z_f^2 + 2 \sum_{g=f+1}^{g=k-1} z_f z_g \right\}$$

$$\chi^2(k-1) = \sum_{f=1}^{f=k-1} \frac{z_f^2}{p_f} + \frac{1}{p_k} \left(z_1 + z_2 + \dots + z_{k-1} \right)^2$$

Or, d'après (V-30) :

$$z_1 + z_2 + \dots + z_{k-1} = z_k$$

et, par suite :

$$(V-37) \quad \chi^2(k-1) = \frac{z_1^2}{p_1} + \frac{z_2^2}{p_2} + \dots + \frac{z_k^2}{p_k}$$

On sait enfin que, dans un espace à $k-1$ dimensions, la loi normale a pour expression générale :

$$dF = \frac{\exp \left\{ -\frac{1}{2} \chi^2(k-1) \right\}}{(2\pi)^{\frac{k-1}{2}} \sqrt{|V|}} dy_1 dy_2 \dots dy_{k-1}$$

formule dans laquelle $|V|$ désigne le déterminant de la matrice des variances des variables y_f .

Tenant compte des valeurs (V-35) et (V-37) du déterminant de la matrice des variances et de la forme quadratique $\chi^2(k-1)$ relative aux variables, on obtient, sous la forme suivante, l'expression définitive de la loi des Grands Nombres généralisée:

$$(V-38) \quad dF = \frac{\exp\left\{-\frac{1}{2}\left(\frac{\bar{z}_1^2}{p_1} + \frac{\bar{z}_2^2}{p_2} + \dots + \frac{\bar{z}_{k-1}^2}{p_{k-1}}\right)\right\}}{(2\pi)^{\frac{k-1}{2}} \sqrt{p_1 p_2 \dots p_{k-1}}} dz_1 dz_2 \dots dz_{k-1}$$

avec

$$(V-29) \quad \bar{z}_f = \frac{x_f - n p_f}{\sqrt{n}}$$

II. Loi des fewes probabilités ou loi de Poisson

Lorsque la probabilité p d'obtenir un succès au cours d'une série de n épreuves répétées est faible, on se propose de montrer qu'il est possible, avec une approximation d'autant meilleure que p est plus faible, de remplacer la distribution de Bernoulli par une distribution de Poisson.

L'expression mathématique de la loi de Poisson ne dépend que du produit $\lambda = np$, alors que l'expression de la loi de Bernoulli fait intervenir les valeurs individuelles de chacun des paramètres n et p . Il est, en conséquence, beaucoup plus facile de calculer les termes d'une loi de Poisson ou d'en publier une table numérique, que d'effectuer les mêmes opérations pour la loi de Bernoulli. Cette facilité d'emploi justifie l'usage de la loi de Poisson au lieu et place de la loi de Bernoulli.

Indépendamment de l'usage précédent, la loi de Poisson représente d'une manière rigoureuse la distribution de nombreux phénomènes aléatoires que l'on peut rencontrer dans les sciences physiques (par exemple distribution spatiale de particules isolées très diluées dans un milieu continu), dans l'art de l'ingénieur (par exemple distribution du nombre d'accidents observés par heure de fonctionnement d'un grand ensemble électrique) ou dans les sciences économiques (par exemple théorie des files d'attente).

V-18. La loi de Poisson comme approximation de la loi de Bernoulli.

La manière la plus simple d'étudier l'évolution de la loi de Bernoulli, lorsque la probabilité p de succès diminue, consiste à rechercher l'évolution correspondante de la seconde fonction caractéristique de la loi de Bernoulli, dont le développement en série (voir Par. V-07) s'écrit :

$$\psi_x(u) = n \left\{ p(e^{iu} - 1) - \frac{1}{2} p^2 (e^{iu} - 1)^2 + \frac{1}{3} p^3 (e^{iu} - 1)^3 + \dots \right.$$

soit encore

$$\psi_x(u) = np(e^{iu} - 1) \left\{ 1 - \frac{1}{2} p(e^{iu} - 1) + \frac{1}{3} p^2 (e^{iu} - 1)^2 + \dots \right.$$

Lorsque p est suffisamment petit, on peut limiter le développement en série de $\psi_x(u)$ à son terme principal et admettre l'approximation

$$(V-39) \quad \psi_x(u) \approx np(e^{iu} - 1)$$

Mous allons montrer que cette fonction caractéristique correspond à une distribution de Poisson (voir Par. I-9) de paramètre $\nu = np$, distribution pour laquelle la probabilité d'une observation x a pour valeur :

$$(V-40) \quad \forall x \in \mathbb{N} : \quad p(x) = e^{-\nu} \cdot \frac{\nu^x}{x!}$$

- (1) L'emploi des formules de réciprocité de Fourier (voir Par. I-22) permettrait de passer directement de la fonction caractéristique (V-39) à la distribution correspondante mais ce calcul direct est plus délicat que le calcul inverse.

En effet la fonction caractéristique de la loi de Poisson à $\nu = 40$ a pour expression :

$$\Phi(u) = E[e^{iuX}] = \sum_{x=0}^{\infty} e^{iuX} \cdot e^{-\nu} \frac{\nu^x}{x!} = e^{-\nu} \sum_{x=0}^{\infty} \frac{(\nu e^{iu})^x}{x!}$$

Soit :

$$\Phi(u) = e^{-\nu} \cdot e^{\nu e^{iu}} = e^{\nu(e^{iu}-1)}$$

d'où l'on déduit la valeur de la seconde fonction caractéristique de la loi de Poisson :

$$(V-41) \quad \psi(u) = \log_e(\Phi(u)) = \nu(e^{iu}-1)$$

Cette expression étant de la même forme que la fonction (V-39), on en déduit qu'il est possible, pour p suffisamment petit, de calculer la probabilité $p(x)$ d'observer x succès au cours de n épreuves en substituant à l'expression exacte⁽¹⁾:

$$\forall (x \leq n) \in N : p(x) = C_n^x p^x q^{n-x}$$

l'expression approchée

$$(V-42) \quad \boxed{\forall (x \leq n) \in N : \left\{ \begin{array}{l} p(x) \neq e^{-\nu} \cdot \frac{\nu^x}{x!} \\ \text{avec} \\ \nu = np \end{array} \right.}$$

⁽¹⁾ N désigne, comme d'habitude, l'ensemble des entiers naturels $\{0, 1, 2, 3, \dots\}$

Il faut bien noter que la loi de Poisson ne constitue qu'une approximation des termes de la loi de Bernoulli qui ne peut lui être substituée que pour toute valeur de x inférieure ou égale au nombre total n d'épreuves de la série étudiée (le nombre x des succès ne peut évidemment pas dépasser le nombre n des épreuves). Lorsque p est petit, on constate bien (voir Abridge IX du recueil de tables numériques) que la somme :

$$\sum_{x=0}^{x=n} e^{-\nu} \frac{\nu^x}{x!} = \sum_{x=0}^{x=\nu/p} e^{-\nu} \frac{\nu^x}{x!}$$

est très voisine de l'unité et, qu'en conséquence, les termes de la loi de Poisson n'ont plus que des valeurs absolument négligeables pour tout x supérieur à n .

En pratique la loi de Poisson constitue une bonne approximation de la loi de Bernoulli lorsque la probabilité p est inférieure ou égale à 0,90.

V - 19. Mode de la distribution de Poisson.

Le diagramme de distribution de la loi de Poisson est un diagramme en bâton rompus s'étendant à l'ensemble des entiers naturels. Le mode de la distribution est égal à la valeur de x correspondant au bâton le plus long. On peut calculer le mode en étudiant les variations du rapport :

$$\frac{p(x+1)}{p(x)} = \frac{e^{-\nu} \frac{\nu^{x+1}}{(x+1)!}}{e^{-\nu} \frac{\nu^x}{x!}} = \frac{\nu}{x+1}$$

Le mode correspond donc à l'entier immédiatement supérieur à la racine de l'équation

$$\frac{\nu}{x+1} = 1$$

soit :

(V-43) Mode = valeur entière (ν)

Lorsque ν est entier il existe deux points de longueur maximum correspondant aux valeurs :

$$x = \nu - 1 \quad \text{et} \quad x = \nu$$

V - 20 Propriétés de la distribution de Poisson.

Le développement en série de Mac-Laurin de la seconde fonction caractéristique (V-41) de la loi de Poisson s'écrit :

$$(V-44) \quad \Psi(u) = \nu \cdot iu + \nu \frac{(iu)^2}{2!} + \nu \frac{(iu)^3}{3!} + \dots + \nu \frac{(iu)^k}{k!} + \dots$$

Par identification on en déduit immédiatement les valeurs suivantes :

$$(V-45) \quad \left\{ \begin{array}{l} E[x] = \nu \\ \sigma_x = \sqrt{\nu} \\ \mu_3 = \nu \end{array} \right.$$

qu'il est utile de retenir.

On constate également que tous les "semi-invariants" (voir paragraphe I - 20) de la distribution de Poisson sont égaux à ν et, par suite, égaux entre eux. Cette propriété est d'ailleurs caractéristique de la loi de Poisson.

La table IX du recueil de tables numériques donne les valeurs des probabilités $P(x)$ pour les valeurs de ν comprises entre 0,1 et 0,9, puis incrément de 0,1 puis pour les valeurs entières de ν inférieures ou égale à 11.

V-21. Fonction de répartition de la loi de Poisson.

On se propose de démontrer que la fonction de répartition de la loi de Poisson, définie par l'expression :

$$(V-46) \quad \forall x, c \in \mathbb{N} : F(c) = \sum_{x=0}^{x=c} e^{-\nu} \frac{\nu^x}{x!}$$

peut encore se mettre sous la forme :

$$(V-47) \quad F(c) = 1 - \frac{1}{c!} \Gamma(c+1) = 1 - \frac{\Gamma(c+1)}{\Gamma(c+1)}$$

expression dans laquelle $\Gamma(c)$ et $\Gamma(c)$ désignent les fonctions élémentaires complètes et incomplètes de deuxième espèce.

On rappelle que la fonction élémentaire complète de 2^e espèce ou "fonction factorielle" a pour expression :

$$(V-48) \quad \Gamma(c) = \int_0^{\infty} e^{-z} z^{c-1} dz$$

pour tout nombre réel ou complexe dont la partie réelle est positive. En outre, lorsque c est un entier positif,

$$(V-49) \quad \forall c \in \mathbb{N} : \Gamma(c+1) = c!$$

De même la fonction élémentaire incomplète de 2^e espèce a pour expression :

$$(V-50) \quad \Gamma(c) = \int_0^{\nu} e^{-z} z^{c-1} dz$$

Pour démontrer la formule (V-47) il suffit de calculer par récurrence la valeur de l'intégrale :

$$\Gamma(c) = \frac{1}{(c-1)!} \cdot \Gamma(c-1) = \frac{1}{(c-1)!} \cdot \int_0^{\nu} e^{-z} \cdot z^{c-1} dz$$

On intégrera par parties en posant :

$$\left\{ \begin{array}{l} u = \frac{1}{(c-1)!} \cdot 3^{c-1} \Rightarrow du = \frac{1}{(c-2)!} 3^{c-2} \cdot d_3 \\ dv = e^{-3} \cdot d_3 \Rightarrow v = -e^{-3} \end{array} \right.$$

Il vient :

$$I_v(c) = \left[-\frac{e^{-3} 3^{c-1}}{(c-1)!} \right]_0^v + \frac{1}{(c-2)!} \int_0^v e^{-3} \cdot 3^{c-2} \cdot d_3$$

soit :

$$(V-51) \quad I_v(c) = I_v(c-1) - \frac{e^{-v} v^{c-1}}{(c-1)!}$$

Par calcul direct on trouve :

$$I_1(1) = \int_0^1 e^{-z} dz = 1 - e^{-1}$$

En appliquant successivement la formule de récurrence
(V-51) il vient :

$$I_1(2) = 1 - e^{-1} - e^{-1} \cdot \frac{1}{1!}$$

$$I_1(3) = 1 - e^{-1} - e^{-1} \cdot \frac{1}{1!} - e^{-1} \cdot \frac{1}{2!}$$

$$I_1(c+1) = 1 - e^{-v} \left(1 + \frac{v}{1!} + \frac{v^2}{2!} + \cdots + \frac{v^c}{c!} \right)$$

d'où l'on déduit que :

$$e^{-v} \left(1 + \frac{v}{1!} + \frac{v^2}{2!} + \cdots + \frac{v^c}{c!} \right) = 1 - I_v(c+1) = 1 - \frac{1}{c!} \Gamma_v(c+1)$$

(V-47) expression qui n'est autre que la formule
que nous voulions démontrer.

Des tables de la fonction "Gamma incomplète"
ont été publiées par K. Pearson à Londres
en 1922.

L'abaque \bar{I}_X donne les valeurs de la fonction de répartition de la loi de Poisson

$$F(c) = \sum_{x=0}^{x=c} e^{-\nu} \cdot \frac{\nu^x}{x!}$$

Pour les valeurs de ν comprises entre 0,1 et 30. Pour les valeurs de ν supérieures à 25, il est possible de substituer la loi des Grands Nombres à la loi de Poisson. On admettra alors que la variable

$$t = \frac{x - \nu}{\sqrt{\nu}}$$

peut être assimilée à une variable normale réduite.

V - 22 Ajustement d'une loi de Poisson sur une série statistique d'observations.

A l'issue d'une série d'expériences ou d'enquêtes au cours desquelles font se réaliser un événement aléatoire, on dresse une statistique indiquant pour chaque période d'observation le nombre x d'apparitions de l'événement aléatoire considéré. On se propose de représenter cette statistique par une distribution de Poisson, quitte à rejeter ce mode de représentation si les écarts entre les fréquences prédictes par la loi de Poisson et les fréquences observées sont trop importantes. Ce problème d'ajustement se réduit, dans ce cas, à trouver une estimation du paramètre ν de la loi de Poisson proposée.

Sachant que le paramètre ν est égal à l'espérance mathématique de la variable de Poisson, il est normal⁽¹⁾ d'adopter comme valeur du paramètre ν la moyenne arithmétique du nombre d'apparitions de l'événement aléatoire par expérience ou période d'observation.

$$\nu \# \bar{x}$$

(1) Cette méthode d'estimation sera justifiée d'une manière plus rigoureuse ultérieurement.

L'exemple suivant illustrera la pratique d'un tel ajustement.

V-22-1 Exemple d'ajustement.

"Dans un atelier de $n = 500$ ouvriers on note pendant $J = 200$ jours les nombres x_i d'accidents du travail journaliers. Les deux premières colonnes du tableau V-22-I procéderont à la statistiques des nombres x_i de journées où l'on a observé un nombre x_i d'accidents du travail. On demande si cette statistique peut être valablement représentée par une distribution de Poisson"

Tableau V-22-I

1	2	3	4	5
Nombre d'accidents journaliers	Nombre de journées où l'on a constaté un nombre x_i d'accidents	Produit des colonnes 1 et 2	$e^{-0,6} \cdot \frac{(0,6)^{x_i}}{x_i!}$	Prévision du nombre de journées où l'on aurait dû constater un nombre x_i d'accidents
x_i	f_i	$x_i \cdot f_i$	$p(x_i)$	$f_i^* = J \cdot p(x_i)$
0	110	0	0,549	109,8
1	65	65	0,329	65,8
2	21	42	0,099	19,8
3	3	9	0,020	4,0
4	1	4	0,003	0,6
5	0	0	0,0003	0,06
6	0	0	0,0000	0,008
Totaux :		$\sum x_i f_i = 120$		
Nombre moyen d'accidents par jour de travail :				
$\bar{x} = \frac{\sum x_i f_i}{J} = \frac{120}{200} = 0,6$				

L'addition des chiffres de la troisième colonne fournit le nombre total $\sum x_i f_i$ d'accidents relevés au cours des J journées observées, ce qui conduit à une moyenne.

$$\bar{x} = \frac{\sum x_i f_i}{J} = \frac{120}{200} = 0,6 \text{ accident/jour}$$

On adoptera cette valeur comme paramètre de la loi de Poisson la mieux indiquée pour représenter la distribution observée. La colonne 4 reproduit les valeurs des probabilités

$$p(x_i) = e^{-0,6} \cdot \frac{0,6^{x_i}}{x_i!}$$

lues sur la table \bar{x} du recueil de tables numériques.

La colonne 5 indique, en fonction du nombre x_i d'accidents journaliers, les prévisions du nombre f'_i de journées où l'on devrait, au cours de 200 jours de travail, observer un nombre x_i d'accidents.

$$f'_i = J \times p(x_i) = 200 \times p(x_i)$$

La comparaison des colonnes 5 et 2 montre que les écarts $f'_i - f_i$ sont très faibles. L'ajustement d'une loi de Poisson de paramètre $\nu = 0,6$ est excellent et fournit une représentation très satisfaisante des résultats observés.

On peut conclure de cette étude que, pour un ouvrier de cet atelier, le risque journalier p d'avoir un accident du travail à jour vaut :

$$p = \frac{np}{n} = \frac{\nu}{n} = \frac{0,6}{500} = 1,2 \times 10^{-3}$$

Le fait que la série des observations puisse être représentée par une loi de Poisson laisse à penser que les accidents du travail sont, dans cet atelier, des événements aléatoires indépendants. Le risque doit donc être identique pour chaque ouvrier et à n'importe quelle époque de l'année. Ce risque

apparaît comme assimilable au tirage au sort d'une boule noire dans une urne contenant 12 boules noires pour un nombre total de 10.000 boules.

V - 23 Addition de variables de Poisson indépendantes.

Lorsque plusieurs variables indépendantes x_1, x_2, \dots, x_k sont distribuées suivant des lois de Poisson de paramètres respectifs $\nu_1, \nu_2, \dots, \nu_k$, on peut montrer⁽¹⁾ que leur somme $S = \sum x_i$ suit une loi de Poisson de paramètre $\nu = \sum \nu_i$.

On sait, en effet, (voir paragraphe III - 5) que la seconde fonction caractéristique d'une somme de variables indépendantes est égale à la somme des secondes fonctions caractéristiques de chacune d'elles. Par suite :

$$\psi(u)_S = \sum_{i=1}^{i=k} \psi(u)_{x_i} = \sum_{i=1}^{i=k} \nu_i (e^{iu} - 1) = (e^{iu} - 1) \sum_{i=1}^{i=k} \nu_i$$

soit en posant

$$(V - 52) \left\{ \begin{array}{l} \nu = \sum_{i=1}^{i=k} \nu_i \\ \psi(u)_S = \nu (e^{iu} - 1) \end{array} \right.$$

ce qui démontre la propriété énoncée.

V - 23 - i Exemple d'application.

Une usine comporte 4 ateliers distincts dans lesquels les effectifs employés ainsi que les risques moyens d'accident sont fournis par l'ouvrage sont reproduits dans les deux premières colonnes du tableau suivant :

- (*) Lorsqu'une somme de variables indépendantes obéit à une distribution de même forme que chacun de ces termes, on dit que le type de distribution correspondant est "stable". A ce titre la loi de Laplace-Gauss est également une distribution stable.

Tableau V - 23 - 1

Ateliers	Effectifs n_i	Risque journalier par ouvrier p_i	Produits
			$V_i = n_i \times p_i$
I	450	$2,0 \times 10^{-4}$	900×10^{-4}
II	1230	$0,5 \times 10^{-4}$	615×10^{-4}
III	80	$8,0 \times 10^{-4}$	640×10^{-4}
IV	340	$0,1 \times 10^{-4}$	34×10^{-4}
TOTAUX:	2100		2189×10^{-4}

En admettant que dans chaque atelier les accidents sont distribués suivant des lois de Poisson, on demande de déterminer une limite supérieure raisonnable du nombre d'accidents qui peuvent se produire dans toute l'usine au cours d'une même journée de travail, sachant que la limite ainsi calculée devra avoir moins de 1 chance sur 1000 d'être dépassée.

Il résulte de ces données que les accidents sont distribués dans chaque atelier, selon des lois de Poisson dont les paramètres $V_i = n_i p_i$ figurent⁽¹⁾ dans la dernière colonne du tableau V-23-1.

Le nombre total S d'accidents journaliers qui peuvent se produire dans toute l'usine suit donc une loi de Poisson de paramètre

$$V = \sum n_i p_i = 0,2189 \text{ accidents/jour}$$

L'abaque IX du recueil de tables numériques montre que, pour une loi de Poisson de paramètre $V = 0,22$ il y a :

(1) On notera que les produits $V_i = n_i p_i$ permettent de comparer les nombres moyens d'accidents journaliers dans chaque ateliers.

} une probabilité de $0,9985$ pour que $S \leq 2$
 } une probabilité de $0,99992$ pour que $S \leq 3$

Soit encore :

} un risque de $1,5 \times 10^{-3}$ pour que $S > 2$
 } un risque de $0,8 \times 10^{-4}$ pour que $S > 3$

La limite cherchée est donc de 3 accidents par jour pour l'ensemble de l'usine.

On notera que l'on aura, en moyenne, à déplorer $0,2189$ accidents journaliers, soit environ 10 accidents pour 46 jours ouvrables. Ces chiffres correspondent à un risque moyen d'accident journalier par ouvrier de :

$$p = \frac{v}{\sum n_i} = \frac{2.189 \times 10^{-4}}{2.100} = 1,04 \times 10^{-4}$$

V-24 Application de La Loi de Poisson à l'étude des suspensions dispersées.

Comme dernier exemple d'application de la Loi de Poisson nous nous proposons de montrer comment cette distribution peut s'appliquer à l'étude de milieux très dispersés.

Lorsque des particules insolubles, non susceptibles de s'agglutiner, sont dispersées dans un liquide sous forme de suspension, on peut admettre que ces particules n'exercent pratiquement plus aucune influence les unes sur les autres lorsque la dilution est suffisante pour que la distance moyenne entre particules soit grande vis à vis de leurs dimensions propres. Leur distribution tend alors à devenir statistiquement uniforme dans un volume de grandes dimensions. Nous caractériserons la suspension étudiée par sa concentration macroscopique λ , ou nombre moyen de particules par unité de volume. Nous nous proposons, dans ces conditions, de calculer la probabilité :

$$P\{\alpha | \lambda, V\}$$

de constater la présence de α particules dans un

volume V donné de la suspension.

Pour simplifier le problème nous supposons que les particules sont approximativement sphériques et de mêmes dimensions. Nous désignerons par Δv le volume du parallélépipède circonscrit à une particule. Le nombre n de cellules élémentaires Δv contenues dans le volume V est donc égal à :

$$n = \frac{V}{\Delta v}$$

La probabilité p de rencontrer une particule dans une cellule élémentaire choisie au hasard a pour valeur :

$$p = \lambda \cdot \Delta v$$

De même, la probabilité q pour qu'une cellule élémentaire soit vide est égale à :

$$q = 1 - \lambda \cdot \Delta v$$

Le volume V contenant n cellules Δv qui peuvent être considérées, quant à leur contenu, comme indépendantes les unes des autres, on en conclut que la probabilité pour que x de ces cellules soient occupées par des particules de la substance en suspension a pour valeur :

$$P\{x, |\lambda, V\} = C_n^x p^x q^{n-x}$$

Lorsque la concentration λ est faible et le volume Δv petit, la probabilité

$$p = \lambda \cdot \Delta v$$

sera alors très faible. On sait (voir Par. V-18) que, dans ces conditions, la loi de Bernoulli peut être remplacée par la loi de Poisson de paramètre :

$$v = np = \frac{V}{\Delta v} \cdot \lambda \cdot \Delta v = \lambda \cdot V$$

La probabilité cherchée a donc pour expression :

$$(V-53) \quad P\{x | \lambda, V\} = e^{-\lambda \cdot V} \cdot \frac{(\lambda \cdot V)^x}{x!}$$

On en déduit, en particulier, que l'espérance mathématique du nombre x de particules contenues dans un volume V de la suspension a pour valeur :

$$(V-54) \quad E[x] = \lambda \cdot V$$

et pour écart-type

$$(V-55) \quad \sigma_x = \sqrt{\lambda \cdot V}$$

La formule V.54 confirme que le paramètre λ est bien égal au nombre moyen de particules par unité de volume de la suspension.

V-24-1 Application.

Dans les processus de numération globulaire on dilue un volume connu de sang dans un volume défini d'une solution isotonique. On dénombre ensuite, à l'aide d'un microscope, le nombre de globules rouges contenus dans un volume donné de la suspension ainsi obtenue. Connaisant la dilution de la suspension, il est facile d'en déduire une estimation du nombre de globules rouges par millimètres cubes du sang examiné.

En supposant que les opérations de dilution, les mesures de volume ainsi que le dénombrement des globules sous le microscope sont effectués sans erreurs, on demande de calculer le nombre minimum

théorique de globules qu'il est nécessaire de compter pour obtenir, dans 95 numérasions sur 100, une estimation de la concentration globulaire ne s'écartant pas de plus de 10%, 5% ou 1% de sa valeur vraie.

En désignant par λ la valeur vraie de la concentration globulaire dans la suspension isotonique et par V le volume examiné sous le microscope, on sait que le nombre x de globules dans le volume V est distribué suivant la loi de Poisson.

$$P\{x|\lambda, V\} = e^{-\lambda V} \cdot \frac{(\lambda V)^x}{x!}$$

On rappelle que la loi de Poisson de paramètre V peut être remplacée par la loi de Gauss d'espérance mathématique $E[x] = V$ et d'écart-type $\sigma_x = \sqrt{V}$ lorsque V est assez grand, en pratique supérieur à 25. Sous réserve de vérification en fonction des conclusions obtenues, nous adopterons l'hypothèse que $\lambda V > 25$. Dans ces conditions on sait que le nombre x de globules contenus dans un échantillon de volume V sera, dans 95% des numérasions effectuées, compris dans l'intervalle:

$$E[x] \pm 1,96 \sigma_x = \lambda V \pm 1,96 \sqrt{\lambda V}$$

L'erreur relative maximum à craindre restera donc, dans 95 numérasions sur 100, inférieure à:

$$\frac{\Delta x}{E[x]} = \frac{1,96 \sqrt{\lambda V}}{\lambda V} = \frac{1,96}{\sqrt{\lambda V}}$$

Elle sera donc:

- { inférieure à 10% pour $\lambda V \geq 385$
- { inférieure à 5% pour $\lambda V \geq 1540$
- { inférieure à 1% pour $\lambda V \geq 38.500$

Le produit $\lambda \cdot V$ représentant l'espérance mathématique du nombre total de globules présents dans l'échantillon contrôlé, on en conclut que, pour obtenir la précision relative demandée il faudra choisir la dilution et le volume total de l'échantillon examiné⁽¹⁾ de manière à ce que le nombre théorique total de globules à décompter soit supérieur à :

385 globules pour que l'erreur relative soit inférieure à 10%

1540 " " " " " " " " 5%

38.500 " " " " " " " " 1%

On remarquera que l'erreur relative ne dépend pas comme l'inverse du nombre de globules dénombrés mais seulement comme la racine carrée de ce nombre.

On notera enfin que l'erreur relative à craindre sur la composition globulaire du sang examiné ne dépend que du nombre total de particules dénombrées, quel que soit le nombre moyen de globules par millimètre cubes du sang analysé ou la dilution de la suspension isotonique examinée.

(1) On subdivisera éventuellement l'échantillon contrôlé en fractions de volume plus faible si le volume V excède la capacité de l'hématomètre utilisé.

DOMAINES D'EMPLOI DES LOIS HYPERGÉOMétriques, DE BERNoulli, DES GRANIS NOMBRES ET DE POISSON.

TIRAGES EXHAUSTIFS

LOI HYPERGÉOMétrique

$$P\{x | N, x, n\} = \frac{\binom{x}{N} \cdot \binom{n-x}{N-x}}{\binom{n}{N}}$$

Pour $n < N/10$, la loi hypergéométrique peut être remplacée avec une bonne approximation par la loi de Bernoulli:
 en posant $p = \frac{x}{N}$

PRÉUVES REPÉTÉES

LOI DE BERNoulli

$$P\{x | n, p\} = C_n^x \cdot p^x \cdot q^{n-x}$$

Pour $p < 0.1$, la loi de Bernoulli peut être remplacée avec une bonne approximation par la loi de Poisson en posant $\lambda = n \cdot p$

Pour $\sqrt{npq} > 5$, la loi de Bernoulli peut être remplacée avec une bonne approximation par une loi normale

LOI DE POISSON

$$P\{x | \nu\} = e^{-\nu} \cdot \frac{\nu^x}{x!}$$

Pour $\lambda = np > 25$ (ou $\sigma_x = \sqrt{\lambda} = \sqrt{np} > 5$), la loi de Poisson peut être remplacée avec une bonne approximation par la loi de Laplace-Gauss de paramètres

LOI DES GRANDS NOMBRES
 Pour $\sqrt{npq} > 5$ la variable

$$t = \frac{x - np}{\sqrt{npq}}$$

est normale réduite

Table générale des matières traitées dans les différents tomes

Introduction au cours d'analyse statistique et de calcul des probabilités.

première partie

statistique descriptive

distribution à un caractère

distribution à deux caractères

fonctions multivariantes

ch. I tome I

ch. II "

ch. III "

deuxième partie

calcul des probabilités

événements aléatoires

les épreuves multiples

processus aléatoires

notion de fiabilité

ch. IV tome I

ch. V "

ch. VI tome II

ch. VII "

troisième partie

théorie de l'échantillonnage et tests statistiques

quatrième partie

méthode d'analyse des observations

analyse de la variance à simple entrée

analyse de la variance à double entrée

analyse de la variance à triple entrée

régression linéaire à une variable

régression linéaire multiple

régressions non linéaires

ch. X^{II} tome III

ch. X^{IV} "

ch. X^V "

ch. X^{VI} tome IV

ch. X^{VII} "

ch. X^{VIII} "

Table des matières du tome I

Introduction au cours d'analyse statistique et de calcul des probabilités

I statistique descriptive	P. 1
II événements aléatoires et probabilités	P. 2
III Théorie de l'échantillonnage et tests statistiques	P. 3
IV Méthode d'analyse des observations	P. 4

première partie

statistique descriptive

chapitre I : distribution si un seul caractère

individus et population	I 1
I 1 distribution d'une variable discrète	I 2
I 2 distribution d'une variable continue	I 3
I 3 localisation approximative d'une distribution	I 4
I 4 écarts (par rapport à la moyenne arithmétique)	I 5
I 5 paramètres d'étalement	I 5
I 6 valeur numérique des écarts types	I 6
I 7 modèle mathématique d'une distribution	I 8
I 8 densité d'une distribution (variable continue)	I 9
I 9 modèle mathématique d'une distribution discrète	I 11
I 10 espérance mathématique d'une variable x	I 11
I 11 variable entrée	I 12
I 12 écart type théorique σ ; variance de x	I 13
I 13 variable réduite	I 13
I 14 moments d'une distribution	I 14
I 15 moments rapportés à une origine quelconque l	I 14
I 16 moments entrés M_q	I 14
I 17 inégalité de Bienaymé-Chebychev	I 16
I 18 fonction caractéristique d'une distribution	I 18
I 19 développement en série de Mac Laurin de la fonction caractéristique.	I 19
I 20 deuxième fonction caractéristique	I 20
I 21 changements de variable et fonctions caractéristiques	I 21
I 22 passage de la fonction caractéristique à la distribution	I 22

I 23	suite de distribution ; recherche de la distribution limite	I 23
I 24	population normale ; loi de soprême Gauss.	I 25
I 25	loi normale réduite	I 25
I 26	fonction caractéristique de la loi normale	I 26
I 27	moments centrés de la loi normale	I 27
I 28	tables de la loi normale réduite	I 28
I 29	écrive de Henry.	I 30

chapitre II distribution à deux caractères

II	1	le concept de relation fonctionnelle entre deux variables.	II	1
II	2	insuffisance du concept de relation fonctionnelle pour représenter une loi plus générale de dépendance entre deux variables	II	2
II	3	étude qualitative d'une corrélation	III	3
II	4	table de contingence	III	3
II	5	fréquence globale - histogramme	III	5
II	6	fréquences marginales	III	5
II	7	fréquences liées	III	6
II	8	expression de la fréquence globale en fonction des fréquences marginales et liées	III	7
II	9	variables indépendantes	III	7
II	10	modèle mathématique	III	8
II	11	fréquences marginales	III	10
II	12	fréquences liées	III	12
II	13	équations des courbes de régression	III	13
II	14	variables indépendantes	III	14
II	15	espérances mathématiques des variables x et y	III	14
II	16	variables entrées	III	15
II	17	moments d'une distribution	III	15
II	18	moments entrés	III	15
II	19	moments entrés d'ordre 2	III	15
II	20	covariance de deux variables indépendantes	III	16
II	21	coefficiant de corrélation	III	17
II	22	matrice des variances et covariances	III	18
II	23	ellipse indicatrice	III	19
II	24	variables réduites	III	21
II	25	indicatrice réduite	III	21
II	26	ellipses de concentration; généralisation de l'inégalité de Bienaymé-Chebychev.	III	23
II	27	devise des moindres carrés	II	25
II	28	fonction caractéristique	II	28
II	29	fonction caractéristique des distributions marginales	II	29

II 30	expression de la fonction caractéristique lorsque les variables x_1 et x_2 sont indépendantes	II 30
II 31	changement du système de référence	II 31
II 32	changement de vecteur	II 33
II 33	développement en série de Mac Laurin de la fonction caractéristique	II 34
II 34	loi normale de Laplace-Gauss à 2 variables	II 36
II 35	tables numériques de la loi normale à 2 variables	II 36
II 36	distributions marginales et distributions liées de la loi normale	II 37
II 37	calcul de la fonction caractéristique de la loi normale	II 38
II 38	forme canonique de la loi normale; couples de variables réduites indépendantes	II 40

chapitre III fonctions multivariées

III 1	espérance mathématique d'une somme	III 1
III 2	espérance mathématique d'un produit	III 1
III 3	écart type d'une somme	III 2
III 4	écart type d'une moyenne arithmétique	III 3
III 5	fonction caractéristique d'une somme de variables indépendantes	III 3
III 6	composition de variables normales	III 4
III 7	écart-type des petites fluctuations d'une fonction quelconque autour d'une valeur fixe	III 5
III 8	tolérances d'ajustement	III 6

autre exemple étude statistique des diviseurs de tension

I	objet de l'étude	III 7
II	expression de la force de faibles fluctuations ξ_1 et ξ_2	III 8
III	fluctuations du rapport à lorsque les résistances utilisées sont distribuées suivant des lois normales d'écart-types σ_1 et σ_2 données	III 9
IV	fluctuations du rapport à lorsque les résistances utilisées sont distribuées uniformément dans leurs intervalles de tolérance d'amplitudes relatives p_1 et p_2	III 12
V	distribution du rapport à lorsque les résistances sont distribuées suivant deux lois en voie de	III 18

deuxième partie

calcul des probabilités

chapitre IV événements aléatoires et probabilités

IV 1	événement aléatoire	1
IV 2	probabilité d'un événement aléatoire	2
IV 3	restriction de Von Mises	2
IV 4	signification de l'échelle de probabilité	3
IV 5	estimation d'une probabilité	4
IV 6	variable aléatoire. population-mère. loi de probabilité.	5
IV 7	événements composés	6
IV 8	principe des probabilités totales.	8
IV 9	événements exclusifs	10
IV 10	généralisation du principe des probabilités totales	11
IV 11	principe des probabilités composés	12
IV 12	événements aléatoires indépendants	13
IV 13	théorème de Bayes. probabilité des causes	14

chapitre V les épreuves multiples

A	échantillonnages exhaustifs	1
V 1	échantillonnage par tirage exhaustif dans une population contenant des éléments de 2 catégories différentes	1
V 2	tirage exhaustif dans une urne contenant des boules de 9 catégories différentes	5
V 3	tirage exhaustif de plusieurs échantillons dans une urne contenant des boules de 9 catégories différentes	6

B	loi des épreuves répétées	8
V 4	expression de la loi de Bernoulli ou loi des épreuves répétées	8
V 5	utilisation de la loi de Bernoulli comme expression approchée d'un tirage exhaustif	9
V 6	fonction caractéristique de la loi de Bernoulli	10
V 7	premiers moments de la loi de Bernoulli	12
V 8	fonction de répartition de la loi de Bernoulli	13
V 9	généralisation de la loi de Bernoulli	17

V 10	fonction caractéristique de la loi généralisée de Bernoulli	V 19
V 11	matrice des variances de la loi de Bernoulli	E 20

C loi des grands nombres

V 12	diagramme de fréquence de la loi de Bernoulli	V 23
V 13	fonction caractéristique de la variable réduite de Bernoulli	V 24
V 14	forme réduite de la loi de Bernoulli lorsque le nombre des épreuves augmente indéfiniment	E 25
V 15	exemple numérique	V 25
V 16	généralisation de la loi des grands nombres	V 30
V 17	expression de la loi des grands nombres généralisée	V 32

D loi des petites probabilités ou loi de Poisson

V 18	la loi de Poisson comme approximation de la loi de Bernoulli	V 39
V 19	modèle de la distribution de Poisson	E 41
V 20	propriétés de la distribution de Poisson	E 42
V 21	fonction de répartition de la loi de Poisson	E 43
V 22	ajustement d'une loi de Poisson sur une série statistique d'observations	E 45
V .22.1	exemple d'ajustement	V 46
V 23	addition de variables de Poisson indépendantes	E 48
V 24	application de la loi de Poisson à l'étude des solutions dispersées	E 50
V 25	domaine d'emploi des lois hypergéométriques, de Bernoulli - des grands nombres - de Poisson.	E 55