Eléments de probabilités et de statistiques

A- ELEMENTS DE STATISTIQUE DESCRIPTIVE

La plupart du temps les données se présentent sous la forme suivante : on a relevé *p* variables numériques sur *n* individus.

On ne s'intéresse ici qu'à une variable X, appelée caractère, dont on possède *n* valeurs. La synthèse de ces données se fait sous forme de tableaux, de graphiques et de résumés numériques. C'est ce que l'on appelle couramment la « statistique descriptive ».

I- Tableaux statistiques et représentations graphiques

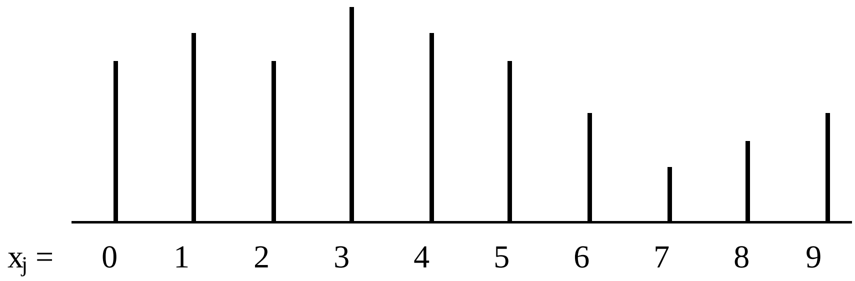
I.1- Variables discrètes

Pour chaque valeur x_j de la variable, on note n_j le nombre d'occurrence (ou effectif) de x_j dans l'échantillon, $\sum n_j = n$, et f_j la fréquence correspondante, $f_j = n_j/n$.

Le tableau statistique se présente sous la forme

x_j	n_j	f_j

Dans un graphique en bâtons, on porte en ordonnée f_j en fonction de x_j



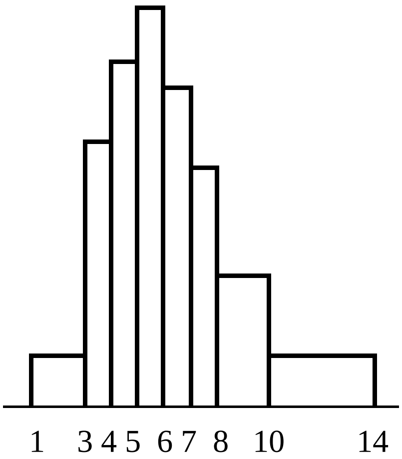
I.2- Valeurs continues

On regroupe les valeurs en k classes d'extrémités e_0, e_1, ..., e_k et l'on note pour chaque classe $[e_{j-1}, e_j[$ (d'amplitude h_j) l'effectif n_j et la fréquence f_j ainsi que les fréquences cumulées $F_j = \sum_{l=1}^{j} f_l$.

Le tableau statistique se présente sous la forme

e_{j-1}			F_{j-1}
	n_j	f_j	
e_j			F_j

Dans un histogramme, le rectangle construit sur chaque classe a une surface égales à la fréquence de la classe:



II- Résumés numériques

Des indicateurs synthétiques permettent de résumer une série de n observations x_1, x_2, x_3, …, x_n.

On note F la fréquence cumulée.

I.1- Caractéristiques de tendance centrale

La médiane est la valeur M telle que F(M) = 0,5. Elle est déterminée par interpolation linéaire dans le cas d'un nombre d'observation paires ou dans le cas d'une répartition par classes.

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

Le mode est la valeur la plus fréquente pour une distribution discrète ou la classe correspondant au pic de l'histogramme pour une variable continue (sa détermination est donc malaisée puisque reposant sur le découpage en classes).

I.2- Caractéristiques de dispersion

L'étendue ou intervalle de variation est égal à $|x_{max} - x_{min}|$

L'intervalle interquartile est égal à $|Q_3 - Q_1|$ où $F(Q_1) = 0,25$ et $F(Q_3) = 0,75$

La variance $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ et l'écart type $s = \sqrt{s^2}$

Les calculs sont simplifiés par l'utilisation de la formule $s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \overline{x}^2$

B – ELEMENTS DE PROBABILITES

I - Quelques définitions

I.1- Définitions préliminaires

On parle d'expérience ou d'épreuve aléatoire lorsque l'on ne peut pas prévoir à l'avance de résultat. Il peut s'agir d'une expérience que l'on peut répéter plusieurs fois dans les mêmes conditions et qui peut avoir des résultats différents (exemple : lancer de dé) ou d'une expérience par nature unique (exemple : observation de la durée de vie d'un individu).

L'ensemble des résultats d'une expérience aléatoire est appelé l'univers des possibles ; il est généralement noté Ω.

Un événement est une affirmation relative au résultat de l'expérience aléatoire. Par exemple, dans le cas du lancer de dé, les propositions « le résultat du lancer de dé est 4 » ou « le résultat du lancer de dé est supérieur à 5 » constituent chacune un événement possible.

Un événement peut être considéré comme une partie de l'ensemble Ω, c'est-à-dire comme un ensemble lui-même. C'est pourquoi l'on utilise fréquemment les notations ensemblistes.

Deux événements A et B peuvent être incompatibles : $A \cap B = \varnothing$

A un événement A correspond un événement contraire noté $\overline{A}$

I.2- Définition de la notion de probabilité

A chaque événement $A \subset \Omega$, on associe un nombre entre 0 et 1, noté P(A) et appelé probabilité de l'événement A. Il faut, de plus, que :

$$P(\Omega) = 1$$

et $$P(A \cup B) = P(A) + P(B) \text{ pour tous événements A et B incompatibles.}$$

Au delà de cette définition mathématique peu « éclairante », on peut proposer trois approches de la notion de probabilité.

La notion classique (équiprobabilité des cas) :

Il s'agit de compter tous les cas de figures où l'événement se produit et de le rapporter au nombre total de cas envisageables.

$$\text{Probabilité} = \frac{\text{nombre de cas favorables}}{\text{nombre total de cas}}$$

Cette notion est très utile pour comprendre sur des exemples simples ce qu'est une probabilité.

La notion expérimentale :

Elle repose sur un théorème que nous verrons plus tard (loi des grands nombres) : lorsqu'une expérience peut être répétée une « infinité » de fois, la fréquence d'un événement tend vers sa probabilité. Autrement dit :

$$\text{Probabilité} = \lim_{\text{grand nombre d'observations}} \frac{\text{nombre de fois où l'événement a lieu}}{\text{nombre total d'observations}}$$

Dans la plupart des problèmes de la vie courante, la probabilité d'un événement ne peut être déterminé par un dénombrement des cas ; d'où l'importance de la définition expérimentale (et, plus généralement, le développement de la statistique).

L'approche expérimentale permet de mieux comprendre la notion de probabilité : attachée à un événement, elle n'a qu'une valeur de « prévision », fondée sur l'étude de cas similaires.

La notion subjectiviste :

Certains phénomènes aléatoires ne se répètent jamais tout à fait dans les mêmes conditions (exemple : évolution de la bourse). On peut être amené à estimer de façon tout à fait subjective des « probabilités » d'événements et essayer d'établir des prévisions en fonction de ces probabilités établies a priori.

I.3- Propriétés élémentaires

1°) $P(\emptyset) = 0$

2°) $P(\overline{A}) = 1 - P(A)$

3°) $P(A) \leq P(B)$ si $A \subset B$

4°) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

5°) Théorème des probabilités totales :

si les événements B_i sont incompatibles deux à deux et que $\cup B_i = \Omega$, on dit qu'ils forment un « système complet d'événements » et, quel que soit l'événement A, on a :

$$P(A) = \sum_i P(A \cap B_i)$$

II - Probabilités conditionnelles

On peut s'intéresser à la probabilité de réalisation d'un événement A sachant qu'un événement B est réalisé. On définit alors la probabilité de A sachant B, notée P (A/B) comme :

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

II.1- Evénements indépendants

Deux événements sont dits indépendants si la connaissance de l'un des événements ne change rien à la probabilité de survenance de l'autre événement.

Par exemple, soit une urne contenant 4 boules de couleurs différentes (bleu, jaune, rouge, noir).

Si l'on effectue deux tirages successifs sans remise, il semble intuitif que l'événement « la deuxième boule tirée est rouge » n'est pas indépendant de l'événement « la première boule tirée est noire ». La connaissance du premier événement nous apporte une information supplémentaire (la boule rouge est encore dans l'urne donc on peut encore l'obtenir au deuxième tirage) : le nombre total de cas possibles n'est plus le même que si nous n'avions aucune information sur le résultat du premier tirage.

En revanche, si l'on remet dans l'urne la première boule tirée avant le second tirage (et qu'on remélange les boules), alors les deux événements sont indépendants l'un de l'autre.

Mathématiquement, cette notion s'exprime ainsi : un événement A est indépendant d'un événement B si la probabilité de A sachant B est égale à la probabilité de A, soit :

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

On démontre alors facilement le résultat très important suivant :

Si deux événements A et B sont indépendants, alors :

$$\boxed{P(A \text{ et } B) = P(A \cap B) = P(A).\,P(B)}$$

II.2- Formules de Bayes

Les formules de Bayes seront utiles pour certains problèmes d'actuariat. Elles permettent de passer de P(A/B) à P(B/A).

1ère formule de Bayes :

$$P(B/A) = \frac{P(A/B).\,P(B)}{P(A)}$$

2ème formule de Bayes :

si les événements B_i sont incompatibles deux à deux et que $\cup B_i = \Omega$, alors :

$$P(B_i/A) = \frac{P(A/B_i).\,P(B_i)}{\sum_k P(A/B_k).\,P(B_k)}$$

III - Variables aléatoires réelles

On parle de variables aléatoires réelles quand les événements possibles (les « résultats » de l'expérience aléatoire) peuvent être décrits par des nombres réels.

On s'intéressera alors aux événements écrits sous la forme : X = a où a est un réel et X est la « variable aléatoire réelle » étudiée.

Exemple : tirage du loto, d'un dé, d'une partie de pile ou face (en posant par exemple la convention suivante : pile = 0, face = 1), âge de décès d'un assuré, montant d'un sinistre dommage, etc.

On distingue :

- les variables aléatoires discrètes, qui prennent leurs valeurs dans un nombre fini ou dénombrable de réels ;

- les variables aléatoires continues, qui prennent leurs valeurs dans un intervalle continu de R.

III.1- Quelques définitions

Fonction de répartition d'une variable aléatoire réelle X

Cette fonction est définie ainsi :

$$F(x) = P(X<x)$$

La fonction de répartition permet de calculer simplement la probabilité de X dans tout intervalle. En effet :

$$P(a \leq X<b) = F(b) - F(a)$$

<u>Densité de répartition</u> d'une variable aléatoire continue

La description d'une variable aléatoire continue ne peut se faire sous la forme P(X=a) comme pour les variables aléatoires discrètes.

On introduit (lorsque c'est possible) la notion de « densité de probabilité » :

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

ce qui permet, par intégration, d'avoir $P(a<X<b) = \int_a^b f(x)\,dx = F(b) - F(a)$

La densité de probabilité de X est la dérivée de sa fonction de répartition.

III.2- Espérance

Pour une variable aléatoire discrète, l'espérance est définie par :

$$E(X) = \sum_{\text{ensemble des x possibles}} x\, P(X=x)$$

E(X) a le sens d'une moyenne de toutes les valeurs possibles de X, pondérées par leur probabilité. Intuitivement, c'est la moyenne des valeurs que l'on observe lorsque le nombre d'observations successives de X tend vers l'infini.

Pour une variable continue de densité de probabilité f, on définit de même :

$$E(X) = \int_R x.\, f(x).\, dx$$

Propriétés de l'espérance

Si a est une constante,

$E(a) = a$ (événement certain)

$E(X+a) = E(X) + a$

$E(aX) = a.\, E(X)$

Soient deux variables aléatoires X et Y,

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

Si, de plus, X et Y sont des variables <u>indépendantes</u>,

$E(XY) = E(X).\, E(Y)$

Théorème de l'espérance totale

Si on note E(Y/X) la variable aléatoire qui prend pour valeurs $E(Y/X=x)=\sum_y y.P(Y=y\ /\ X=x)$ avec les probabilités P(X=x), on a :

$$E[E(Y/X)]=E(Y)$$

III.3- Variance

La variance est définie par :

$$V(X) = E\,[(X-E(X))^2]$$

On voit dans cette définition que l'on cherche à mesurer « en moyenne » (en espérance), l'écart de la variable aléatoire à E(X). V(X) traduit la dispersion des résultats de X autour de la « moyenne ».

L'écart type (noté σ) de X est la racine carrée de la variance V(X)

$$\sigma = \sqrt{V(X)}$$

Propriétés de la variance

Autre formule de la variance : $V(X) = E(X^2) - (E(X))^2$

Inégalité de Bienaymé-Tchébychev :

$$P(\,|\,X - E(X)\,| > k\,\sigma) < \frac{1}{k^2}$$

Cette inégalité est la première relation entre espérance et écart type qui permet d'apprécier la dispersion d'une variable aléatoire.

Variance d'une somme de variables aléatoire :

- cas général :

$$V(X+Y) = V(X) + V(Y) + 2\,\mathrm{cov}\,(X, Y)$$

où la quantité cov (X,Y), appelée <u>covariance</u> de X et de Y vaut : E(XY) - E(X). E(Y)

- si X et Y sont indépendantes,

$$V(X+Y) = V(X) + V(Y)$$

Théorème de la variance totale

Si on note V(Y/X) la variable aléatoire qui prend pour valeurs $V(Y/X=x)= E[(Y-E(Y/X=x))^2 / X=x]$ avec les probabilités P(X=x), on a :

$$V(Y) = E[V(Y/X)] + V[E(Y/X)]$$

C- EXEMPLES DE LOIS DE PROBABILITES DISCRETES

I - La Loi Binomiale

I.1- Loi de Bernoulli

On étudie la probabilité p de survenance d'un événement A et celle de son contraire $\overline{A}$, de probabilité 1-p. Tout naturellement on s'intéresse alors à la fonction indicatrice de cet événement X, variable aléatoire telle que :

X = 1 si l'événement A survient (probabilité p)

X = 0 sinon (probabilité 1-p).

On montre facilement que :

$E(X) = p$

$V(X) = p\,(1-p)$

I.2- Epreuves n fois répétées, loi binomiale

On répète n fois l'expérience aléatoire précédente. Chaque fois, l'événement A a une probabilité p de survenir. La question est la suivante : quelle est la probabilité que l'événement A surviennent k fois lors de n répétitions de l'expérience aléatoire ?

On note X_i la variable indicatrice de la i-ème expérience aléatoire identique effectuée. On peut considérer que les X_i (i variant de 1 à n) sont indépendantes (par construction).

On étudie alors la variable aléatoire $X = \sum_{i=1}^{n} X_i$ correspondant au nombre de fois où l'événement A peut survenir au cours de la répétition de n épreuves aléatoires identiques.

On montre que la loi de X est :

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

De plus :

$E\,(X) = np$

Et $V(X) = np(1-p)$ soit aussi $\sigma = \sqrt{np(1-p)}$

<u>Description de la loi binomiale</u> :

P (X=k) croît puis décroît quand k augmente.

La valeur la plus probable de k est la valeur la plus proche de np.

La loi binomiale peut être approximée :

- par la loi de Poisson si p est petit (en pratique $p<0,1$) et n est assez grand ($n>50$), ou encore $n.p < 10$;

- par la loi de Gauss quand n devient grand (en pratique si $n.p > 30$).

II - La Loi de Poisson

La variable aléatoire X, à valeurs dans N, suit une loi de Poisson de paramètre λ (notée $P(\lambda)$) si elle vérifie la formule suivante :

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

On montre que :

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

$$V\,(X) = \lambda$$

Propriété : addition de deux lois de Poisson indépendantes.

Soit X_1 et X_2 deux variables aléatoires indépendantes suivant respectivement les lois de Poisson $P(\lambda_1)$ et $P(\lambda_2)$, la variable aléatoire $X_1 + X_2$ suit une loi de poisson de paramètre $\lambda_1+\lambda_2$.

La loi de Poisson est très utile en actuariat dommage ainsi que dans tous les domaines où il s'agit de modéliser le nombre de fois où un événement (sinistre en assurance) survient dans un intervalle de temps donné.

II.1- La loi de Poisson comme limite de lois binomiales

On montre qu'une loi binomiale de paramètres n et p peut être approchée si n est grand et que $np \to \lambda$ vers la loi de Poisson de paramètre λ.

Si l'on découpe un intervalle de temps T en n parties égales (n grand) et que l'on suppose que dans chaque petit intervalle de temps ainsi obtenu un seul sinistre a la possibilité de survenir, avec une probabilité p, le nombre (aléatoire) de sinistres survenus dans la période de temps T est décrit par une loi binomiale de paramètres n et p.

Si l'on affine le découpage, par exemple en n' = 2n parties, il semble intuitif que la probabilité de survenance du sinistre sur un intervalle élémentaire est divisée par 2 : p' = p/2. On remarque qu'on peut ainsi diviser « à l'infini » l'intervalle de temps T, tout en maintenant constante la quantité np = n'p' ... = λ La loi de Poisson décrit donc le « cas limite » suivant : le risque de survenance d'un sinistre est le même à chaque « instant » (limite d'un petit intervalle de temps).

II.2- Le processus de Poisson

Un processus de Poisson est défini par les hypothèses « naturelles » suivantes : on suppose que les événements sont indépendants et que le nombre d'événements survenant pendant une période T ne dépend que de la durée de cette période (et non par exemple des dates de début et de fin de cette période). On suppose enfin que le nombre moyen d'événements par unité de temps est constant, on le note c (« cadence » ou « intensité de fréquence » du processus).

Soit N le nombre aléatoire de sinistres se produisant dans la période de temps T, on montre que N suit une loi de Poisson de paramètre λ = c T. En particulier, E (N) = c T.

D - LES LOIS DE PROBABILITES CONTINUES

I - Généralités sur les variables aléatoires continues

I.1 - Densité de probabilité d'une variable aléatoire continue

Une variable aléatoire réelle est dire **continue** lorsqu'elle prend ses valeurs dans un intervalle continu de R. La description d'une variable aléatoire continue X ne peut se faire, comme dans le cas d'une variable aléatoire discrète, par la définition des probabilités P(X=x), car le nombre de valeurs possibles n'est pas dénombrable : on ne peut définir que la probabilité pour que X se trouve «au voisinage » de x.

On exprime cela en utilisant les notations mathématiques infinitésimales : dx désigne la longueur d'un intervalle infiniment petit, au voisinage de x. Au lieu de probabilités individuelles, on définit la **densité de probabilité de X**: si f(x) est la densité de probabilité au point x, la probabilité pour que X se situe entre x et x+dx est égale à : f(x) * dx.

Par intégration, on obtient ensuite la probabilité pour que X se situe entre deux points a et b :

$P(a < X < b) = \int f(x) * dx = F(b) - F(a)$,

où F est la **fonction de répartition** de la variable aléatoire X :

$\mathbf{F(x) = P(X < x)}$

La fonction de densité f est donc **la dérivée** de la fonction de répartition F : dans la pratique, on cherchera souvent à déterminer la fonction de répartition de la variable aléatoire, puis on obtiendra la densité de probabilité en dérivant cette fonction de répartition.

I.2 - Espérance et variance d'une variable aléatoire continue

Définitions

Les grandeurs définies pour les variables aléatoires discrètes ont leur analogue dans le cas de variables aléatoires continues : les probabilités individuelles de réalisation de x : P(X=x) sont simplement remplacées par des densités de probabilité au voisinage de x : f(x)*dx, et les sommations sont remplacées par **des intégrales**.

Le fait que f définisse une densité de probabilité se vérifie en calculant l'intégrale sur R $\int f(x) * dx$, qui doit être égale à 1, puisque la probabilité de l'univers des possibles est égale à 1.

L'espérance d'une variable aléatoire continue de fonction de densité f se définit ainsi :

E(X) = $\int x * f(x) * dx$

La **variance** se définit comme : $V(X) = E((X - E(X))^2)$

Comme pour une variable discrète, **l'écart type** se définit simplement comme la racine carrée de la variance.

Propriétés

Les propriétés de l'espérance et de la variance sont les mêmes dans le cas continu que dans le cas discret.

- Si X et Y sont deux variables aléatoires continues **quelconques** :

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

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

V(X+Y) = V(X) + V(Y) + 2 * COV(X,Y) , où **COV(X,Y) = E(X*Y) - E(X)*E(Y)**

• Si X et Y sont deux variables aléatoires continues **indépendantes** :

E(X*Y) = E(X) * E(Y)

V(X+Y) = V(X) + V(Y)

II.2 - Loi uniforme

Définition et principales propriétés

Soit a > 0. On définit la **loi de probabilité uniforme** sur l'intervalle [0 ; a] par la densité de probabilité f suivante :

f(x) = 1/a si $x \in [0 ; a]$

f(x) = 0 sinon

Sa fonction de répartition s'écrit F(x) = P(X < x) = $\int_{-\infty}^{x} f(y) * dy$

F(x) = 0 **si x < 0**

F(x) = 1 **si x > 1**

F(x) = $\int_{0}^{x} 1/a * dx$ = **x / a** **si 0 < x < a**

E(X) = $\int_{0}^{a} x/a * dx = [\, x^2 / 2a \,]_0^a$ = **a / 2**

V(X) = $E(X^2) - E(X)^2 = \int_{0}^{a} x^2 / a * dx - a^2/4 = [\, x^3 / 3a \,]_0^a - a^2/4 = a^2/3 - a^2/4 =$ $\mathbf{a^2 / 12}$

II.3 - La loi exponentielle

Définition et principales propriétés

La loi exponentielle de paramètre a est définie par la densité de probabilité suivante.

f(x) = a * exp (-a * x) pour $x \in]0 ; +\infty[$ f(x) = 0 sinon

Pour que f(x) soit une densité de probabilité, il faut que

$$1 = \int f(x) * dx = \int_{0}^{+\infty} a * e^{-a*x} * dx = [-\exp(-a*x)]_0^{+\infty}$$

ce qui est vrai **si et seulement si a > 0**

La fonction de répartition s'écrit : **F(x) = 1 - exp(-a*x)**

$$E(X) = \int_{0}^{+\infty} a * x * e^{-a*x} * dx = [-x*\exp(-a*x)]_0^{+\infty} - \int_{0}^{+\infty} -e^{-a*x} * dx \quad \text{(intégration par parties)}$$

E(X) = $0 - [1/a * \exp(-a*x)]_0^{+\infty}$ = **1/a**

$V(X) = E(X^2) – E(X)^2$

$$E(X^2) = \int_0^{+\infty} a * x^2 * e^{-a*x} * dx = [-x^2*\exp(-a*x)]_0^{+\infty} - \int_0^{+\infty} -2*x*e^{-a*x}*dx = 2/a * E(X)$$

D'où : **V(X)** = $2/a^2 - 1/a^2$ = $\mathbf{1/a^2}$

L'écart type d'une loi exponentielle est donc égal à son espérance, soit 1/a.

Loi de survie exponentielle

On s'intéresse à la **fonction de survie** d'un assuré : R(t) = P (X > t), où X est la date aléatoire de décès de l'assuré et t > 0.

On se propose de modéliser la loi de survie de l'assuré par la fonction R(t) = exp(-c*t).

La probabilité (conditionnelle) pour que l'assuré décède avant t_2, sachant qu'il a vécu jusqu'à t_1, s'écrit : $P(X < t_2 / X > t_1)$

$= P(X < t_2 \text{ et } X > t_1) / P(X > t_1) = (P(X > t_1) - P(X > t_2)) / P(X > t_1) = \mathbf{(R(t_1) - R(t_2)) / R(t_1)}$

Si R(t) = exp(-c*t), alors : $(R(t_1) - R(t_2)) / R(t_1) = \mathbf{1 - \exp(-c*(t_2-t_1))}$

Dans cette modélisation, la probabilité de survie de t_1 à t_2 **ne dépend que de la longueur de l'intervalle de temps entre t_1 et t_2, indépendamment de l'âge t_1 atteint** : par exemple, la probabilité de survie jusqu'à 80 ans d'un assuré âgé de 60 ans est la même que la probabilité de survie jusqu'à 100 ans d'un assuré de 80 ans.

Si on se place sur un intervalle de temps infinitésimal dt, on définit ainsi **le taux de mortalité instantané** à l'âge t, noté $\mu(t)$:

(probabilité de décès entre t et t+dt) / (probabilité de survie jusqu'à l'âge t) = $\mu(t)$ * dt

Selon la modélisation proposée, on a : $\mu(t)$ * dt = (exp(-c*t) - exp(-c*(t+dt))) / exp(-c*t)

En faisant tendre dt vers 0 et en dérivant, on obtient : $\mathbf{\mu(t)}$ = - (-c*(exp(-c*t))) / exp(-ct) = **c**

Le modèle proposé correspond donc à un **taux de mortalité constant quel que soit l'âge** : cela n'est bien sûr pas adapté à la description de la mortalité humaine, le taux de mortalité étant en réalité croissant avec l'âge : pour une durée fixée, la probabilité de survie sur cette durée décroît avec l'âge atteint.

II.3 – La loi gamma

Définition et principales propriétés

Une variable aléatoire positive suit une loi gamma de paramètres (r,a), notée $\boldsymbol{\gamma}$**(r,a)**, si sa densité est de la forme :

$\mathbf{f(x) = a^r / \Gamma(r) * \exp(-a*x) * x^{r-1}}$ pour $x \in]0 ; +\infty[$ $a > 0$ et $r \neq 0$

$\Gamma(r)$ est un facteur de normalisation introduit pour que f(x) soit une densité de probabilité, i.e. $\int f(x)*dx = 1$.

$$\mathbf{\Gamma(r)} = \int_0^{+\infty} e^{-y} * y^{r-1} * dy$$

On montre facilement, en effectuant une intégration par parties, que :

$$G(r+1) = \int_0^{+\infty} e^{-y} * r * y^{r-1} * dy = r * G(r)$$

Et comme $\Gamma(1) = \int_0^{+\infty} e^{-y} * dx = [-\exp(-y)]_0^{+\infty} = 1$, on a de façon générale **: G(r) = (r-1)!** pour tout r entier strictement positif, ce qui explique que cette fonction soit également appelée fonction factorielle.

$$E(X) = \int_0^{+\infty} a^r / \Gamma(r) * e^{-a*x} * x^r * dx = \int_0^{+\infty} e^{-y} / \Gamma(r) * y^r * dy / a = \Gamma(r+1) / (a*\Gamma(r))$$

(changement de variable y = a*x)

E(X) = r / a

$$V(X) = \int_0^{+\infty} a^r / \Gamma(r) * e^{-a*x} * x^{r+1} * dx - r^2/a^2 = \int_0^{+\infty} e^{-y} / \Gamma(r) * y^{r+1} * dy / a^2 - r^2/a^2$$

$V(X) = \Gamma(r+2) / (a^2*\Gamma(r)) - r^2/a^2 = r*(r+1)/a^2 - r^2/a^2$

V(X) = r / a^2

Remarque : lorsque r = 1, la densité de la loi gamma s'écrit :

f(x) = a * exp (-a*x) pour $x \in]0 ; +\infty[$ avec a > 0

On retrouve **la loi exponentielle, qui est un cas particulier de loi gamma, avec un paramètre r égal à 1**.

II.4 - La loi de Pareto

Définition et principales propriétés

La loi de Pareto est notamment utilisée en **actuariat de la réassurance** : elle permet de décrire une partie de la distribution d'une variable aléatoire continue, par exemple la partie excédant une certaine valeur (priorité d'un traité de réassurance XL, la variable aléatoire représentant alors le coût du sinistre).

Une loi de Pareto est définie par la fonction de répartition suivante :

$F(x) = 1 - (x_0 / x)^a$ pour $x \in [x_0 ; +\infty[$ (avec a > 0 et $x_0 > 0$)

La fonction de densité de probabilité f est la dérivée de la fonction de répartition F :

f(x) = $(-a) * (-x_0 / x^2) * (x_0 / x)^{a-1}$ = **$a * x_0^a * x^{-a-1}$**

$E(X) = \int_{x_0}^{+\infty} a * x_0^a * x^{-a} * dx$: on sait que cette intégrale converge **si et seulement si a > 1**

L'espérance d'une variable aléatoire suivant une loi de Pareto n'est définie que si a > 1

Dans ce cas : **E(X)** = $a * x_0^a / (1-a) * [x^{1-a}]_{x_0}^{+\infty}$ = **$a / (a-1) * x_0$**

$E(X^2) = \int_{x_0}^{+\infty} a * x_0^{\ a} * x^{1-a} * dx$: on sait que cette intégrale converge **si et seulement si a > 2**

La variance d'une variable aléatoire suivant une loi de Pareto n'est définie que si a > 2

Dans ce cas : $E(X^2) = a * x_0^a / (2-a) * [x^{2-a}]_{x_0}^{+\infty} = a / (a-2) * x_0^2$

D'où : $V(X) = E(X^2) - E(X)^2 = (a / (a-2) - a^2 / (a-1)^2) * x_0^2$

V(X) = $\dfrac{a}{(a-2)*(a-1)^2} * x_0^2$

Loi de Pareto et loi conditionnelle

Soit X variable aléatoire suivant une loi de Pareto dont la fonction de répartition est donnée par :

$F(x) = 1 - (x_0 / x)^a$ pour $x \in [x_0 ; +\infty[$ (avec $a > 0$)

On s'intéresse à la loi conditionnelle de X sachant X > M.

$P(X > x / X > M) = P(X > x \text{ et } X > M) / P(X > M)$

- Si $x \leq M$: $P(X > x \text{ et } X > M) = P(X > M)$ et la probabilité conditionnelle vaut **1**
- Si $x > M$: $P(X > x \text{ et } X > M) = P(X > x) = 1 - F(x) = (x_0 / x)^a$ car $x > x_0$

$P(X > M) = 1 - F(M) = (x_0 / M)^a$ car $M > x_0$

D'où : $\mathbf{P(X > x / X > M) = (x_0 / x)^a / (x_0 / M)^a} = \mathbf{(M / x)^a}$

On en déduit que : $F(x / X > M) = P(X < x / X > M) = 1 - (M / x)^a$ pour $x \in [M ; +\infty[$ et 0 sinon.

Si X suit une loi de Pareto de paramètres a et x_0, la loi conditionnelle de X sachant que X > M (où $M > x_0$) est **une loi de Pareto, de paramètres a et M** qui n'est définie que si a > 1,.

Dans ce cas : $\mathbf{E(X / X > M) = a / (a-1) * M}$

Lien entre loi de Pareto et loi exponentielle

Soit X variable aléatoire suivant une loi de Pareto dont la fonction de répartition est donnée par :

$F(x) = 1 - (x_0 / x)^a$ pour $x \in [x_0 ; +\infty[$ (avec $a > 0$)

On considère la variable aléatoire $\mathbf{Y = \ln (X / x_0)}$.

La densité de probabilité de la loi de X est : $f(x) = a * x_0^a * x^{-a-1}$

On pose $Y = \ln (X / x_0)$, et on appelle g la fonction de densité de la loi suivie par Y.

X variant sur $[x_0 ; +\infty[$, Y varie sur $[0 ; +\infty[$

Le changement de variable s'écrit : $x = x_0 * \exp(y)$

On doit avoir : $g(y) * dy = f(x) * dx = f(x_0*\exp(y)) * x_0*\exp(y)*dy = a*x_0^a*(x_0*\exp(y))^{-a-1} * x_0*\exp(y)*dy$

D'où : $\mathbf{g(y)} = a * \exp(y)^{-a} = \mathbf{a * \exp(-a*y)}$

La variable aléatoire Y suit donc **une loi exponentielle, de paramètre a**.

Y, contrairement à X, possède une espérance et une variance quelle que soit la valeur de a > 0 (égales respectivement à 1/a et $1/a^2$).

Pour estimer le paramètre a, on utilisera le résultat précédent, en travaillant sur les valeurs des logarithmes des observations de X, variable qui en réassurance non-vie, correspondra généralement au coût des sinistres

II.5 – La loi normale ou de Laplace-Gauss

La loi normale joue un rôle fondamentale en théorie des statistiques.

Définition et principales propriétés

Une variable aléatoire suit une **loi normale ou de Laplace-Gauss de paramètres m et s** si sa densité de probabilité est donnée par la formule :

f(x) = $\frac{1}{s\sqrt{2p}} * \exp(\frac{-(x-m)^2}{2s^2})$ pour tout $x \in$ R

On notera cette loi : **N (m ; s)**

E(X) = 1 / (σ $\sqrt{2p}$) * $\int_{-\infty}^{+\infty} x * e^{-(x-m)^2/2s^2} * dx$

En effectuant le changement de variable y = (x-m)/σ, on obtient :

E(X) = 1 / (σ $\sqrt{2p}$) * $\int_{-\infty}^{+\infty} (s * y + m) * e^{-y^2/2} * s * dy$

E(X) = (σ / $\sqrt{2p}$) * $\int_{-\infty}^{+\infty} y * e^{-y^2/2} * dy$ **+ (**m / $\sqrt{2p}$) * $\int_{-\infty}^{+\infty} e^{-y^2/2} * dy$

La fonction y*exp(-y^2/2) est anti-symétrique par rapport à 0 (f(-y) = -f(y) pour tout réel y), donc son intégrale calculée sur]-∞ ; +∞[est égale à 0 : le premier terme disparaît.

Par ailleurs, on doit avoir : $\int f(x) * dx = 1$, ce qui s'écrit : $\int_{-\infty}^{+\infty} e^{-(x-m)^2/2s^2} * dx$ = σ* $\sqrt{2p}$

ou encore, après le changement de variable : y = (x - m) / σ (dx = σ*dy)

$\int_{-\infty}^{+\infty} e^{-y^2/2} * s * dy$ = σ * $\sqrt{2p}$

soit $\int_{-\infty}^{+\infty} e^{-y^2/2} * dy = \sqrt{2p}$

La fonction exp(-y^2/2) est symétrique par rapport à 0, donc son intégrale calculée sur]-∞ ; +∞[est égale au double de son intégrale calculée sur]0 ; +∞[.

D'où : $\int_{0}^{+\infty} e^{-y^2/2} * dy = \sqrt{p/2}$ **,** que l'on remplace dans le second terme de l'expression de E(X).

ce qui donne **E(X) = m**

V(X) = E((X-m)^2) = 1 / (σ $\sqrt{2p}$) * $\int_{-\infty}^{+\infty} (x-m)^2 * e^{-(x-m)^2/2s^2} * dx$

En effectuant de nouveau le changement de variable y = (x-m)/σ, on obtient :

V(X) = 1 / (σ $\sqrt{2p}$) * $\int_{-\infty}^{+\infty} s^2 * y^2 * e^{-y^2/2} * s * dy$

$$V(X) = \sigma^2 / \sqrt{2\pi} * \int_{-\infty}^{+\infty} y^2 * e^{-y^2/2} * dy = \sigma^2 * \sqrt{2/\pi} * \int_{0}^{+\infty} y^2 * e^{-y^2/2} * dy$$

$$V(X) = \sigma^2 * \sqrt{2/\pi} * \int_{0}^{+\infty} (-y) * (-y * e^{-y^2/2}) * dy$$

En intégrant par parties : $V(X) = \sigma^2 * \sqrt{2/\pi} * \left([-y * \exp(-y^2/2)]_0^{+\infty} - \int_0^{+\infty} -e^{-y^2/2} * dy \right)$

$V(X) = \sigma^2 * \sqrt{2/\pi} * \int_0^{+\infty} e^{-y^2/2} * dy = \sigma^2$ puisque $\int_0^{+\infty} e^{-y^2/2} * dy = \sqrt{\pi/2}$

Une loi normale $N(m, \sigma)$ a pour espérance m et pour variance σ^2 (donc pour écart type σ).

La loi normale centrée réduite

On a souvent utilisé le changement de variable $y = (x-m) / \sigma$

Si X suit la loi normale $N(m,\sigma)$, alors la variable **$Y = (X-m) / \sigma$ suit la loi $N(0,1)$**, dite **loi normale centrée réduite**.

En effet, si on note g la fonction de densité de Y, on a :

$g(y)*dy = f(x)*dx = 1 / (\sigma \sqrt{2\pi}) * \exp(-y^2/2) * \sigma * dy$

$g(y) = 1/\sqrt{2\pi} * \exp(-y^2/2)$, ce qui correspond bien à la densité de la loi $N(0,1)$.

Somme de deux lois normales

On a la propriété importante suivante (conservation du caractère « normal » par addition).

Si X et Y sont deux variables aléatoires indépendantes suivant respectivement les lois normales $N(m_1, \sigma_1)$ et $N(m_2, \sigma_2)$, alors **la variable $Z = X + Y$ suit la loi normale $N(m_1+m_2 , \sqrt{\sigma_1^2 + \sigma_2^2})$.**

Lien entre loi normale et loi gamma

Soit X une variable aléatoire réelle suivant la loi normale $N(0 ; \sigma)$.

On s'intéresse à la variable aléatoire $Y = X^2$.

On a, en notant f et g les densités des variables X et Y :

$x = y^{1/2}$ et $dx = 1/2*y^{-1/2}*dy$

La probabilité pour Y d'être au voisinage de y est égale au double de la probabilité de X d'être au voisinage de $x = y^{1/2}$ car la fonction carré est symétrique et car la densité de la loi $N(0 ; \sigma)$ est également symétrique par rapport à 0.

D'où : $g(y)*dy = 2*f(x)*dx = f(y^{1/2})*y^{-1/2}*dy$

$$g(y) = \frac{1}{\sigma\sqrt{2\pi}} * e^{\frac{-y}{2\sigma^2}} * y^{-1/2}$$

En notant : **$a = 1 / 2\sigma^2$ et $r = 1/2$**

g(y) = $a^{1/2}$ * $\frac{1}{\sqrt{p}}$ * exp (-a*y) * y^{r-1}

Ce qui est la fonction de densité de **la loi gamma :** $g(\frac{1}{2};\frac{1}{2s^2})$

On a (facteur de normalisation) :

G(1/2) = $\sqrt{p}$

II.6 - La loi log-normale

Définition et principales propriétés

Soit Y une variable aléatoire réelle strictement positive.

On dit que **Y suit une loi log-normale si la variable X = ln(Y) suit une loi normale**.

Si X = ln(Y) suit la loi normale N(m,σ), alors :

x = ln(y) donc on a : dx = dy / y

On a : f(y)*dy = g(x)*dx , où g(x) est la densité de la loi normale N(m,σ).

Donc : f(y)*dy = 1/(σ $\sqrt{2p}$) * exp (-(ln(y)-m)^2 / (2*σ^2)) * dy/y

D'où : **f(y) = 1 / (y*s $\sqrt{2p}$) * exp (- (ln(y)-m)^2 / (2s ^{2}))**

E(Y) = $\int y*f(y)*dy$ = 1 /(σ$\sqrt{2p}$) * $\int_0^{+\infty} e^{-(\ln y-m)^2/2s^2}*dy$

On procède au changement de variable : x = (ln(y) – m) / σ équivalent à : y = exp(σ*x+m)

E(Y) = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-x^2/2}*e^{s*x+m}*dx$ = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-x^2/2+sx+m}*dx$

E(Y) = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-(x-s)^2/2}*dx$ * exp(m + σ^2/2)

E(Y) = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-y^2/2}*dy$ * exp(m + σ^2/2)

E(Y) = exp(m + s ^{2}/2)

E(Y^2) = 1/$\sqrt{2p}$ * $\int_0^{+\infty} y*e^{-(a*\ln(y)+b)^2/2}*dy$

On procède de nouveau au changement de variable : x = a*ln(y) + b

E(Y^2) = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-x^2/2}*e^{2*s*x+2*m}*dx$ = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-x^2/2+2sx+2m}*dx$

E(Y^2) = 1/$\sqrt{2p}$ * $\int_{-\infty}^{+\infty} e^{-(x-2s)^2/2}*dx$ * exp(2m + 2σ^2) = exp(2m + 2σ^2)

Donc : V(Y) = E(Y^2) - E(Y)^2 = exp(2m + 2σ^2) - exp(2m + σ^2)

$V(Y) = exp(2m + s^2) * (exp(s^2) - 1)$

Dans le cas particulier ou $m = 0$ et $\sigma = 1$, c'est-à-dire si $ln(Y)$ suit la loi normale centrée réduite, on a :

$E(Y) = exp(1/2) = \sqrt{e}$

$V(Y) = e * (e-1)$

E - LOI DES GRANDS NOMBRES ET THEOREME DE LA LIMITE CENTRALE

L'idée directrice de ce chapitre est la suivante : on considère une variable aléatoire X, et on s'intéresse à des observations X_i de X ; quelle que soit la loi de X, la moyenne des X_i (pour i compris entre 1 et n), notée M_n, ou la somme des X_i, notée S_n, ont certaines propriétés lorsque n est suffisamment « grand » : on parle de **propriétés asymptotiques**.

Ces propriétés se traduisent par **les deux théorèmes fondamentaux** connus sous les dénominations suivantes : la loi des grands nombres et le théorème de la limite centrale.

I - La loi des grands nombres

I.1 - La loi faible des grands nombres

On s'intéresse au comportement asymptotique (c'est-à-dire lorsque n tend vers $+\infty$) des sommes :

$S_n = X_1 + X_2 + \ldots + X_n$

Où les X_i sont des variables aléatoires réelles indépendantes équidistribuées.

On note E(X) et V(X) l'espérance et la variance de la loi commune à ces n variables aléatoires, et on suppose que V(X) est finie.

On sait que $E(S_n) = n * E(X)$ (linéarité de l'espérance)

Et que $Var(S_n) = Var(X_1) + Var(X_2) + \ldots + Var(X_n) = n * V(X)$ (indépendance des X_i)

En appelant $M_n = S_n / n$ la moyenne des X_i, on a :

$\mathbf{E(M_n) = E(X)}$

$\mathbf{Var(M_n)} = Var(S_n / n) = n*V(X) / n^2$ $\mathbf{= V(X) / n}$

Par conséquent, la variable aléatoire moyenne M_n a la même espérance que les X_i, mais possède une variance inférieure, qui tend vers 0 lorsque n tend vers $+\infty$.

Cela s'écrit :

$E[(M_n - E(X))^2] = Var(M_n) = V(X) / n \to 0$ lorsque $n \to +\infty$

On dit que **M_n converge vers E(X) en moyenne quadratique**.

En appliquant l'inégalité de Bienaymé-Chebyshev, dont la formulation générale est :

$P [| Y - E(Y) | > a*\sigma (Y)] < 1 / a^2$ pour $a > 0$

On obtient avec $Y = M_n$ et $a = \varepsilon / \sigma (M_n)$, où ε est un réel positif arbitrairement petit :

P [| M_n - E(X) | > e] < $Var(M_n) / \varepsilon^2 = V(X) / (n*\varepsilon^2)$ **® 0 lorsque n ® +¥**

Autrement dit, pour tout α et tout ε arbitrairement petits, il existe un N tel que pour tout n supérieur à N, la probabilité pour que M_n s'écarte de E(X) de plus de ε soit inférieure à α.

Ce résultat constitue **la loi faible des grands nombres**.

I.2 - La loi forte des grands nombres

La loi forte des grands nombres s'exprime ainsi :

Pour toute suite (X_i) de variables aléatoires réelles indépendantes et équidistribuées, d'espérance E(X) finie :

P (1/n * (X_1 + X_2 + … + X_n) ® E(X) lorsque n ® +¥) = 1

On dit que **M_n converge vers E(X) avec une probabilité égale à 1**.

Mathématiquement, ce résultat est beaucoup plus fort que le premier (et plus difficile à démontrer, faisant intervenir la théorie de la mesure de Lebesgue et le théorème de Borel-Cantelli).

Dans la pratique, il suffit de retenir que la moyenne des X_i tend vers E(X) lorsque n tend vers $+\infty$, la loi forte ajoutant que cette convergence a lieu « presque partout » dans l'univers des possibles.

La seconde grande loi statistique vise à décrire à quelle vitesse cette convergence a lieu : c'est la théorème de la limite centrale, parfois appelé théorème central-limite.

II - Le théorème de la limite centrale ou « central-limite »

II.1 - Formulation du théorème

On reprend les hypothèses précédentes : on considère une suite (X_i) de variables aléatoires réelles indépendantes équidistribuées, d'espérance E(X), d'écart type fini $\sigma(X)$. On désigne par S_n la variable aléatoire somme de n variables X_i et par M_n la variable aléatoire moyenne de n variables X_i.

La question se pose de savoir « à quelle vitesse » M_n converge vers E(X). Le théorème central-limite apporte une réponse, en décrivant comment se comporte asymptotiquement l'écart entre M_n et E(X) : pour n « grand », cet écart suit approximativement une loi normale centrée.

Le théorème de la limite centrale s'énonce ainsi :

Les lois de probabilités des variables aléatoires (moyennes partielles normalisées)

$\sqrt{n}$ * (M_n - E(X)) / s (X)

convergent lorsque n tend vers +¥ vers la loi normale centrée réduite N (0;1).

Ce qui signifie que pour n « assez grand », la loi de $\sqrt{n}$ * (M_n - E(X)) / σ(X) « ressemble » à une loi gaussienne centrée réduite.

On peut également écrire le théorème central-limite ainsi :

(S_n - n*E(X)) / ($\sqrt{n}$ * s (X)) ® N (0;1)

Remarque : ce théorème est encore valable lorsque X_1, X_2, ... X_n ne suivent pas toutes la même loi, si certaines conditions sur les $\sigma_i = \sigma(X_i)$ sont respectées (par exemple, si les s_i sont bornées). Le théorème s'écrit alors :

$$\frac{X_1 + X_2 + ... + X_n - [E(X_1)+E(X_2)+ ... +E(X_n)]}{\sqrt{n} * \sigma} \rightarrow N(0,1)$$

où : $\sigma^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2$

II.2 - Application à la loi binomiale

Si B_n suit une loi binomiale de paramètres (n, p), B_n est la somme de n variables aléatoires indépendantes X_i suivant toutes la loi de Bernoulli de paramètre p.

Comme E(X) = p et V(X) = p*(1-p), le théorème central-limite (seconde formulation) indique que :

(B_n – n*p) / $\sqrt{n*p*(1-p)}$ ® N (0;1)

Cela permet d'approcher la loi binomiale B(n ; p) par la loi normale N(n*p ; $\sqrt{n*p*(1-p)}$ **)**

D'où la formule de calcul approché d'une probabilité de réalisation de la loi binomiale :

$$P(B_n = x) \approx P\ [\ x - 1/2 < n*p + \sqrt{n*p*(1-p)} * U < x + 1/2\ ]$$

où U suit une loi normale centrée réduite.

$$\mathbf{P\ (B_n\ = x)}\ \text{»}\ \mathbf{P}\ [\ \frac{x-\frac{1}{2}-n*p}{\sqrt{n*p*(1-p)}}\ <\ \mathbf{U}\ <\ \frac{x+\frac{1}{2}-n*p}{\sqrt{n*p*(1-p)}}\ ]$$

La probabilité pour que U se trouve dans l'intervalle ci-dessus s'obtient ensuite à partir d'une **table de la loi normale** : celle-ci donne, pour x compris entre 0,00 et, souvent, 2,99, la fonction de répartition F(x) (c'est-à-dire la probabilité pour qu'une variable aléatoire suivant la normale centrée réduite prenne une valeur inférieure à x).

On a : P [x_1 < U < x_2] = F(x_2) – F(x_1)

Si x_1 ou x_2 est négatif, on utilisera la relation : F(x) = 1 - F(-x) pour se ramener à des valeurs figurant dans la table.

II.3 - Généralisation : application pratique du théorème à une loi quelconque

Plus généralement, le principe d'application pratique du théorème central-limite sera le même que pour le cas particulier de la loi binomiale.

Étant donnée une suite (X_i) de variables réelles aléatoires réelles indépendantes, suivant une même loi d'espérance E(X) et d'écart type σ(X), la loi de la variable aléatoire somme partielle S_n = X_1 + X_2 + … + X_n pourra être approchée, pour les valeurs élevées de n, par la loi normale :

Loi de S_n » N (n*E(X) ; $\sqrt{n}$ * s (X))

De même, la loi de la variable aléatoire moyenne partielle M_n = (X_1 + X_2 + … + X_n) / n pourra être approchée, pour les valeurs élevées de n, par la loi normale :

Loi de M_n » N (E(X) ; s (X) / $\sqrt{n}$)

<u>Remarque</u> : on ne peut pas dire de façon rigoureuse que les lois des S_n et des M_n tendent vers une loi normale, car la loi normale approchée dépend de n : ce n'est qu'après **normalisation** (soustraction de n*E(X) puis division par $\sqrt{n}$ *σ(X) pour la somme, soustraction de E(X) puis multiplication par $\sqrt{n}$ /σ(X) pour la moyenne) que l'on obtient une suite de lois tendant vers une loi normale unique, la loi centrée réduite de Laplace-Gauss.

Pour n élevé, on aura alors l'approximation suivante :

$$\mathbf{P\ [\ x_1} < \frac{S_n - n*E(X)}{\sqrt{n}*s(X)} \mathbf{< x_2\ ]\ =\ P\ [\ x_1} < \sqrt{n}*(\frac{M_n - E(X)}{s(X)}) \mathbf{\ < x_2\ ]}\ \text{»}\ \mathbf{F(x_2) - F(x_1)}$$

Où F est la fonction de répartition de la loi normale centrée réduite, donnée par la table de Laplace-Gauss.

Dans la pratique, on souhaitera souvent obtenir un **intervalle de confiance** pour la moyenne M_n, intervalle centré autour de l'espérance E(X).

On cherchera par exemple A tel que : P [-A < M_n - E(X) < A] = 95%

On devra donc avoir :

$$\text{P [-A} * \frac{\sqrt{n}}{s(X)} < \sqrt{n}*(\frac{M_n - E(X)}{s(X)})) < \text{A} * \frac{\sqrt{n}}{s(X)}\ ] = 95\%$$

$$F\left(A * \frac{\sqrt{n}}{s(X)}\right) - F\left(-A * \frac{\sqrt{n}}{s(X)}\right) = 95\%$$

(où F est la fonction de répartition de la loi normale centrée réduite)

Cette équation donne :

$$2 * F\left(A * \frac{\sqrt{n}}{s(X)}\right) - 1 = 95\%$$

$$F\left(A * \frac{\sqrt{n}}{s(X)}\right) = 97,5\%$$

On lit alors dans la table de Laplace-Gauss la valeur de x donnant une répartition de 97,5% (c'est : x = 1,96), et on en déduit la valeur de A.

Le problème pourra être posé différemment : A sera donné et on cherchera à partir de quelle valeur de n l'intervalle [-A ; A] est un intervalle de confiance à 95% (ou 99%, ou 99,9%, etc…) pour l'écart entre la moyenne M_n et l'espérance E(X).