PROBABILITÉS

Johann PAVIS (selon source : B. PERROT)

Table des matières

Chapitre 1

Généralités

1.1 Loi de probabilité

Les probabilités sont conçues pour être une mesure de la vraisemblance des évènements.

1.1.1 Définition générale

On considère Ω l'ensemble des évènements élémentaires qui constituent le phénomène que l'on veut étudier (par exemple, pour un jeu de dés ordinaire, à 6 faces, Ω sera $\{1,2,3,4,5,6\}$), les parties A de Ω représentent les évènements que l'on peut être amené à étudier ($A = \{2,4,6\}$ dans l'exemple précédent représente l'évènement "la valeur obtenue par le lancé de dé est un nombre pair") ; F sera la famille des parties A que l'on veut pouvoir considérer.

On est alors naturellement conduit à imposer les propriétés suivantes :

$$\begin{aligned}
&\Omega \in F \\
&A \in F \text{ et } B \in F \Rightarrow C_A(B) \in F \\
&\forall n, A_n \in F \Rightarrow (\bigcup A_n) \in F
\end{aligned}$$

Une famille F qui possède ces propriétés est appelée une <u>tribu</u>.

Une loi de probabilité est alors une application p telle que :

$$p : \left| \begin{array}{l} F \to [0,1] \\ A \mapsto p(A) \end{array} \right.$$

et qui vérifie les propriétés suivantes :

$$\begin{cases} p(\Omega) = 1 \\ (\mathrm{A}_n)_n \text{ étant une partition de } A \text{, alors : } p(A) = \sum p(A_n) \end{cases}$$

Remarque : il s'ensuit que $p(C(A)) = 1 - p(A)$ ($C(A)$ est l'évènement contraire de A).

1.1.2 Exemple ; Loi uniforme sur un ensemble fini

Ici Ω est une ensemble fini, cette loi uniforme doit traduire la notion intuitive de "hasard", c'est à dire : chaque évènement élémentaire à la même vraisemblance. On doit alors prendre pour F la tribu formée par toutes les parties de Ω et $\forall\{x\} \in \Omega(Card(\{x\}) = 1), p(\{x\}) = \frac{1}{Card(\Omega)}$.

Des propriétés du 1.1.1, on déduit :

$$p(A) = \frac{Card(a)}{Card(\Omega)}$$

qu'on traduit par la célèbre formule : $\frac{\text{Nombre de cas favorables}}{\text{Nombre de cas possibles}}$

1.1.3 Probabilités conditionnelles

On notera $p(A|B)$ la probabilité de A si l'on sait que B a lieu ($p(B) > 0$).

Ce qui donne, dans le cas de la loi uniforme sur un ensemble fini :

$$\left| \begin{array}{ll} p(A|B) = & \frac{Card(A\cap B)}{Card(B)} \\ = & \frac{Card(A\cap B)}{Card(\Omega)}.\frac{Card(\Omega)}{Card(B)} \\ = & \frac{p(A\cap B)}{p(B)} \end{array} \right.$$

C'est la justification la plus simple de la formule de Bayes ci dessous pour le cas général :

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

1.1.4 Évènements indépendants

Traduire que A et B sont indépendants se fait naturellement en écrivant :

$$\begin{cases} p(A|B) = p(A) \\ \text{ou bien} \\ p(A \bigcap B) = p(A).p(B) \end{cases}$$

Cette dernière formule se généralise de manière évidente au cas de n évènements indépendants.

1.1.5 Formule des probabilités totales

Des considérations ci-dessous, il découle la formule :

(B_n) étant une partition de Ω avec $p(B_n) > 0$, on a :
$p(A) = \sum p(A|B_n).p(B_n)$

1.2 Variables aléatoires

1.2.1 Définitions générales

Une variable aléatoire X est un ¡¡objet mathématique¿¿ (en fait, on peut l'interpréter comme une fonction) qui, dès lors d'une réalisation, peut prendre différentes valeurs possibles (l'ensemble de ces valeurs est appelé *espace des valeurs de* X.

Par exemple, pour le jeu de dés ordinaire à 6 faces, on peut considérer la variable aléatoire X = valeur que l'on pourrait obtenir si on lançait le dé.

Alors, l'évènement : "obtenir un nombre pair lors d'un lancé du dé" s'écrit : $X \in \{2, 4, 6\}$.

A une variable aléatoire X, on associe une probabilité : $A \rightarrow p(X \in A)$, appelée loi de X.

X peut être à valeur vectorielle (de dimension n) si l'on veut considérer des $n-$*upplets* de variables aléatoires comme une seule variable aléatoire.

Même, X pourrait être non numérique (par exemple une couleur, une réponse "oui"ou "non", etc) mais dans ce cas, on la numérise symboliquement.

Pour un couple (X,Y) de variables aléatoires, on appelle loi conjointe du couple, la loi de probabilité : $\Delta \rightarrow p((X, Y) \in \Delta)$, pour $\Delta \in$ la tribu considérée dans l'espace des valeurs de (X,Y) (souvent la considération des parties Δ de la forme $\Delta = A$xB suffit).

Ceci se généralise au cas de n variables aléatoires.

1.2.2 Variables aléatoires indépendantes

X et Y sont deux variables aléatoires indépendantes ssi "$X \in A$" et "$Y \in B$" sont des évènements indépendants pour tout A et B. Ceci se traduit par :

$$p(X \in A \text{ et } Y \in B) = p(X \in A).p(Y \in B)$$

Généralisation immédiate au cas de n variables aléatoires indépendantes.

1.2.3 Loi marginale

Si Ω_X désigne l'ensemble des valeurs possibles pour X, si on connaît la loi conjointe du couple (X,Y), on a : $p(Y \in B) = p(X \in \Omega_X \text{ et } Y \in B)$.

La loi de probabilité $B \rightarrow p(Y \in B)$ ainsi récupérée est appelée loi marginale de Y.

Idem pour X et généralisation immédiate au cas de n variables aléatoires.

Parmi toutes les lois de probabilités que l'on peut théoriquement définir, deux catégories sont utilisées dans la pratique (statistiques, traitement du signal, théorie de l'information, recherche opérationnelle, complexité d'algorithmes en moyenne, recouvrement aléatoire de graphes, etc). Se sont :

- les probabilités discrètes
- les probabilités diffuses

Chapitre 2

Loi de probabilité discrète

2.1 Probabilités discrètes

2.1.1 Construction générale

Ω est un ensemble dénombrable, à ce titre nous pouvons le représenter sous la forme d'une suite $(w_n)_{n\in\mathbb{N}}$. F est la tribu formée par toutes les parties de Ω. On considère une suite de réels $p(w_n)$ telle que :

$$\sum_{n}^{Card(\Omega)} p(w_n) = 1$$

Alors :

$$A \to p(A) = \sum_{n/w_n\in A} p(w_n)$$

donne la forme de toutes les lois de probabilités appelées discrètes.

Remarque : $p(\{w_n\}) = p(w_n)$ Remarquer aussi que la loi qui décrit le hasard (vue en 1.1.1) est une loi discrète.

2.1.2 Variable aléatoire discrète

On appelle variable aléatoire discrète une variable aléatoire dont la loi est discrète.

2.2 Espérance mathématique d'une variable aléatoire discrète

2.2.1 Définition

X étant une variable aléatoire discrète, l'ensemble des ses valeurs possibles Ω_X est dénombrable, donc représentables par une suite (x_n).

On appelle espérance mathématique (ou plus simplement espérance) d'une variable aléatoire discrète X et on note $E(X)$ la quantité :

$$E(X) = \sum_n x_n p(X = x_n)$$

lorsque la série de terme général $x_n p(X = x_n)$ est absolument convergente. Cette condition de convergence absolue est nécessaire pour savoir comment faire la sommation par "paquets". Pour signaler cette condition, on emploiera simplement la terminologie : "E(X) existe".

Dorénavant, lorsque nous utiliserons $E(X)$, cette condition sera toujours supposée vérifiée, sans que cela soit nécessairement rappelé.

Nous donnerons la signification heuristique de $E(X)$ au paragraphe 4. C'est elle qui justifie la terminologie "espérance". Et même, on peut dire qu'elle justifie toute la théorie des probabilités.

2.2.2 Formule de transfert

f étant une fonction numérique définie sur Ω_X, l'ensemble des valeurs possibles pour une variable aléatoire discrète X, l'ensemble des valeurs possibles $f(X)$ est dénombrable et donc représentable par une suite $(x_n^{'})$. Par définition, on a :

$$E(f(X)) = \sum_n x_n^{'} p(X = x_n)$$

La formule de transfert donne :

$$E(f(X)) = \sum_n f(x_n^{'}) p(X = x_n)$$

Preuve :

Naturellement, il est sous-entendu que $E(X)$ et $E(f(X))$ existent. Les ensembles : $paquet(n) = \{x \text{ tq } f(x) = x_n^{'}\}$ forment une partition de Ω_X

lorsque les $x_n^{'}$ parcourent l'ensemble des valeurs possibles pour $f(X)$. On effectue la sommation par $paquet(n)$. On obtient alors :

$$\sum_n \sum_{x \in paquet(n)} f(x)p(X = x_n) = \sum_n x_n^{'} \sum_{x \in paquet(n)} p(X = x_n)$$

L'union parcourant : $paquet(n)$ des évènements "X = x" est l'évènement : "$f(X) = x_n^{'}$". D'où :

$$\sum_{x \in paquet(n)} p(X = x) = p(f(X) = x_n^{'})$$

Soit :

$$\text{expression de départ} = \sum_n x_n^{'} p(f(X) = x_n^{'})$$

2.2.3 Linéarité de E

Pour α et β deux scalaires, et pour X et Y deux variables aléatoires, on a : $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$.

Preuve : D'après la formule de transfert :

$$E(\alpha X + \beta Y) = \sum_{n,m} (\alpha x_n + \beta y_m) p(X = x_n \text{ et } Y = y_n)$$

$$E(\alpha X + \beta Y) = \sum_{n,m} \alpha x_n p(X = x_n \text{ et } Y = y_m) + ...$$

$$E(\alpha X + \beta Y) = \alpha \sum_n \sum_m p(X = x_n \text{ et } Y = y_m) + ...$$

$$E(\alpha X + \beta Y) = \alpha \sum x_n p(X = x_n) + ...$$

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

2.2.4 Variable aléatoire constante

On assimile une constante c^{ste} et une variable aléatoire X dont la loi est : $p(X = \mathrm{c}^{ste}) = 1$.

Ainsi : $E(c^{ste}) = E(X) = \mathrm{c}^{ste}$.

Donc : $\forall Y$ variable aléatoire, $E(Y - E(Y)) = 0$. On dit que $Y - E(Y)$ est centrée.

2.2.5 Formule à propos de E(XY)

Si X et Y sont deux variables aléatoires **indépendantes**, on a alors : $E(X.Y) = E(X).E(Y)$

Preuve :

$$E(XY) = \sum_{n,m} x_n y_m p(X = x_n \text{ et } Y = y_n)$$

$$E(XY) = \sum_{n} x_n p(X = x_n) \sum_{m} y_m p(Y = y_m) = E(X)E(Y)$$

2.3 Variance. Covariance

Dans ce qui suit (par soucis de simplicité), les notations sont adaptées au cas de variables aléatoires à valeurs réelles. L'adaptation au cas de variables aléatoires à valeur complexe est aisée, puisqu'il suffit de remplacer la forme bilinéaire symétrique $C_{X,Y} = E(XY) - E(X)E(Y)$ par la forme sesquilinéaire hermitienne : $E(X\overline{Y}) - E(X)E(\overline{Y})$.

2.3.1 Variance

On montre aisément (en majorant séparément les $|x_n|$ dans la série $\sum |x_n| p(X = x_n)$, selon que $|x_n| > 1$ ou que $|x_n| \leq 1$, que : $E(X^2)$ existe $\Rightarrow E(X)$ existe (en règle générale, la réciproque est fausse).

Dans le cas où $E(X^2)$ existe, on appelle variance de X la quantité : $Var(X) = E(X^2) - (E(X))^2$

On a la formule : $Var(X) = E(X^2) - (E(X))^2 = E((X - E(X))^2)$

Preuve : $E((X - E(X))^2) = E(X^2 + (E(X))^2 - 2E(X)X) = E(X^2) - (E(X))^2$ Cette formule permet de donner la signification heuristique de la variance : $Var(X)$ caractérise l'importance des variations de X autour de son espérance.

α étant un scalaire, on a : $Var(\alpha x) = \alpha^2 Var(X)$
Pour cette raison, on introduit la notion d'écart type de X : $\sigma(X) = \sigma_X = \sqrt{Var(X)}$ On a ainsi : $\forall \alpha \in \mathbb{R}^+$, $\sigma(\alpha X) = \alpha\ \sigma(X)$

2.3.2 Covariance

Considérons un couple (X,Y) tel que $E(X^2)$ et $E(Y^2)$ existent. On a alors : $E(XY)$ existe. On appelle alors covariance du couple (X,Y) et on note $C_{X,Y}$ la forme bilinéaire symétrique positive :
$C_{X,Y} = E(XY) - E(X)E(Y) = E((X - E(X))(Y - E(Y))$
(la preuve est immédiate avec la linéarité).

Remarque : $Var(X, X) = C_{X,X}$

D'après 2.2.5, si X et Y sont indépendantes, $E(XY) = E(X)E(Y)$. On a alors : $C_{X,Y} = 0$ (la réciproque de cette propriété est généralement fausse).

2.3.3 Coefficient de corrélation

On appelle coefficient de corrélation de X et de Y la quantité :

$$\rho_{X,Y} = \frac{C_{Y,Y}}{\sigma_X.\sigma_Y}$$

L'intérêt de cette notion est que le coefficient de corrélation est invariant par changement d'unités.
Attention à ne pas confondre les notions de corrélation et de causalité.

L'inégalité de Cauchy-Schwartz appliquée à la forme bilinéaire symétrique positive $C_{X,Y}$ donne : $\begin{cases} C_{X,Y} \leq C_{X,X}C_{Y,Y} \\ |C_{X,Y}| \leq \sigma_X\sigma_Y \end{cases}$
D'où : $|\rho_{X.Y}| \leq 1$

2.3.4 Variance d'une somme

$$Var(X + Y) = Var(X) + Var(Y) + 2C_{X,Y}$$

Preuve : immédiate par développement de la forme bilinéaire symétrique $C_{X+Y,X+Y}$.

Corollaire : Si X et Y sont **indépendantes**, alors $C_{X,Y} = 0$. on a alors :

$$Var(X + Y) = Var(X) + Var(Y)$$

Généralisation immédiate au cas de n variables aléatoires indépendantes, ce qui donne le résultat important ci-dessous :

$X_1, X_2, \ldots, X_n$ étant n variables aléatoires indépendantes suivant la même loi que X (on dit que les X_i sont un *échantillon* de modèle X et de taille n), on considère alors la variable aléatoire $\overline{X}$, définie par :

$$\overline{X} = \sum_{i=1}^{n} X_n$$

Cette variable aléatoire est appelée estimateur de moyenne (appellation justifiée dans la partie "Statistiques inductives"). On a alors :

$$Var(\overline{X}) = \frac{1}{n} Var(X)$$

2.4 Variable Aléatoire à valeurs entières

2.4.1 Fonction génératrice

X étant une variable aléatoire à valeurs entières, on appelle fonction génératrice la série entière : $G(z) = \sum p(X = n) z^n$, de rayon de convergence supérieur à 1 (la somme des probabilités vaut 1).

La fonction génératrice caractérise la loi de X, car on a la formule :

$$p(X = n) = \frac{1}{n!} G^{(n)}(0)$$

Supposons que le rayon de convergence soit ¿ 1. Alors, en dérivant la série entière, terme à terme au point $z = 1$, on obtient :

$$E(X) = G^{'}(1)$$

Et en dérivant une seconde fois :

$$Var(X) = G^{''}(1) + G^{'}(1) - (G^{'}(1))^2$$

Si X et Y sont indépendantes, la fonction génératrice de la somme $X + Y$ est le produit des fonctions génératrices de X et de Y.

Preuve :
Posons $S = X + Y$. On a alors :

$$p(S = k) = \sum_n p(X = n \text{ et } Y = k - n)$$

$$p(S = k) = \sum_n p(X = n)p(Y = k - n)$$

Donc on a :

$$G_S(z) = \sum_{k=0}^{\infty} p(S = k)z^k = \sum_{k=0}^{\infty} \sum_{n=0}^{k} p(X = n)p(Y = k - n)z^k$$

$$G_S(z) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} p(X = n)z^n p(Y = k - n)z^{k-n} = G_X(z)G_Y(z)$$

Généralisation immédiate dans le cas de n variables aléatoires indépendantes.

2.4.2 Loi binomiale

Au cours d'une certaine expérience, un évènement A a la probabilité d de se réaliser. On considère la variable aléatoire X = nombre d'occurence de cet évènement A, que l'on obtiendrait en effectuant N telles expériences de manière indépendante (les variables aléatoires caractéristiques de cet évènement A au cours de ces expériences sont aussi appelées *variables de Bernouilli*).

La loi de X est alors : $p(X = k) = \begin{cases} C_N^k d^k (1-d)^{N-k} \text{ , si } 0 \leq k \leq N \\ 0 \text{ , sinon} \end{cases}$
où C_N^k est le coefficient du binôme de Newton.

On appelle cette loi "*loi binomiale d'ordre N et de paramètre d*".

Sa fonction génératrice est : $G(z) = (d.z + (1 - d))^N$

Preuve :

$$G(z) = \sum_{k=0}^{\infty} C_N^k d^k (1 - d)^{N-k} = \sum_{k=0}^{\infty} C_N^k d^k.z^k (1 - d)^{N-k}$$

On a donc : $E(X) = N.d$ et $Var(X) = N.d(1 - d)$

2.4.3 Loi de Poisson

On appelle *loi de Poisson* de paramètre a ($a \in \mathbb{R}^{+*}$) la loi entière :

$$\forall k \in \mathbb{N} \text{ , } p(X = k) = \mathbf{e}^{-a}\frac{a^k}{k!}$$

<u>Théorème</u> 1 *On pose $a = N.d$. Pour a fixé :*

$$\lim_{N\to\infty} C_N^k d^k (1-d)^{N-k} = \boldsymbol{e}^{-a}\frac{a^k}{k!}$$

<u>Preuve</u> :

$$C_N^k d^k (1-d)^{N-k} = \frac{N}{N} \dots \left(\frac{N-k+1}{N}\right) . \left(1 - \frac{a}{N}\right)^{-k} . \left(1 - \frac{a}{N}\right)^N \frac{a^k}{k!}$$

Dans la pratique dès que N est assez grand, on est amené à approcher la loi Binomiale de paramètre d et d'ordre N, par la loi de Poisson de paramètre $a = N.d$, car celle-ci est d'un emploi plus commode.

Voici quelques exemples numériques :

N=30 et d=0.333	k = 5		k = 10		k = 15	
	Binom	Poisson	Binom	Poisson	Binom	Poisson
p(X=k)	2.3 E-2	3.8 E-2	1.53	1.25	2.5 E-2	3.5 E-2
N=50 et d = 0.1	k = 1		k = 5		k = 10	
	Binom	Poisson	Binom	Poisson	Binom	Poisson
p(X=k)	2.9 E-2	3.3 E-2	0.849	0.775	1.5 E-2	1.8 E-2
N=100 et d = 0.1	k = 1		k = 10		k = 20	
	Binom	Poisson	Binom	Poisson	Binom	Poisson
p(X=k)	3 E-4	4 E-4	0.13	0.125	1.7 E-3	1.8 E-3
N=200 et d = 0.1	k = 10		k = 20		k = 30	
	Binom	Poisson	Binom	Poisson	Binom	Poisson
p(X=k)	4.5 E-3	5.8 E-3	9.4 E-2	8.9 E-2	6.8 E-3	8.3 E-3

La fonction génératrice est :

$$G(z) = \mathbf{e}^{-a}\sum_{n=0}^{\infty}\frac{a^n}{n!}.z^n = \mathbf{e}^{-a}\sum_{n=0}^{\infty}\frac{(a.z)^n}{n!}$$

D'où : $E(X) = a$ et $Var(X) = a$

<u>Théorème</u> 2 *Si X et Y sont <u>indépendantes</u> et suivent des lois de Poisson de paramètres respectifs a et b, alors $S = X + Y$ suit une loi de Poisson de paramètre $a + b$*

<u>Preuve</u> :

$$G_S(z) = G_X(z).G_Y(z) = \mathbf{e}^{a.(z-1)}.\mathbf{e}^{b.(z-1)} = \mathbf{e}^{(a+b)(z-1)}$$

Généralisation immédiate au cas de N variables aléatoires.

<u>exemple</u> : Soit $X_1, X_2, \ldots, X_N$ un échantillon de modèle X qui suit la loi de Poisson de paramètre a. Alors, $N.\overline{X}$ suit la loi de Poisson de paramètre $N.a$

Chapitre 3

Variable aléatoire diffuse

3.1 Généralités

3.1.1 Construction générale

Ω est une partie de $\mathbb{R}$ ou de $\mathbb{R}^n$. La densité f est une fonction numérique positive et continue sur Ω telle que : $\int_\Omega f = 1$

On pose alors : $\forall A$ mesurable $\in \Omega$, $p(A) = \int_A f$

On constate donc qu'on est obligé de restreindre la tribu F (qui contient toutes les parties pour lesquelles on veut calculer la probabilité) à ne contenir que des parties mesurables. Cette restriction théorique ne pose aucun problème pratique, car dans tous les cas rencontrés, les parties A utiles ont une forme très simple, ce qui assure qu'elles sont mesurables.
Dans tout ce qui suit, cette restriction théorique est supposée vérifiée.

Ceci est le procédé de construction des probabilités diffuses.

3.1.2 Exemple : Loi uniforme sur Ω borné

La loi uniforme sur un ensemble borné Ω traduit la notion intuitive de "hasard". Donc, on prend pour une densité constante et égale à $\frac{1}{Mesure(\Omega)}$, pour assurer la condition $\int_\Omega f = 1$.
On a alors :

$$p(A) = \frac{\int_A f}{\int_\Omega f} \text{ , formule qu'on peut aussi énoncer :} \frac{\text{Mesure des cas favorables}}{\text{Mesure des cas possibles}}$$

3.1.3 Remarque

On pourrait aussi envisager des probabilités qui contiennent une partie discrète et une partie diffuse. (w_n) étant une suite de points de Ω, $(p(w_n))$ une suite de réels positifs et f une fonction positive continue sur Ω telles que :

$$\sum_{n=0}^{\infty} p(w_n) + \int_{\Omega} f = 1$$

on définit : $p(A) = p_{\text{discrète}}(A) + p_{\text{diffuse}}(A)$, où $p_{\text{discrète}}(A)$ est définie comme au 2.1.1 et $p_{diffuse}(A)$ comme en 3.1.1.

3.1.4 Fonction de répartition d'une loi de probabilité sur une partie de $\mathbb{R}$

Ω étant une partie de $\mathbb{R}$, on appelle fonction de répartition la fonction :

$$x \to F(x) = p(\{w \in \Omega / w \leq x\})$$

On constate aisément que F est une fonction croissante de 0 à 1 (les valeurs 0 et 1 pouvant être des valeurs asymptotiques).

Si la loi de probabilité est diffuse avec densité continue f, alors F est une primitive de f.

Si la loi de probabilité contient une partie discrète et une partie diffuse (comme dans la remarque 3.1.3), alors, en chaque point w_n, F présente une discontinuité avec un saut de hauteur $p(w_n)$. En dehors de ces points, F est une primitive de f.

Remarque : L'utilisation de la fonction de répartition F est le moyen pratique pour chercher la densité d'une loi de probabilité.

3.1.5 Variables aléatoires diffuses

Une variable aléatoire sera appelée diffuse (avec densité continue) si sa loi de probabilité est diffuse (avec densité continue). Par la suite : “avec densité continue” sera sous-entendu et on dira simplement “diffuse”.

La fonction de répartition F d'une variable aléatoire X à valeurs réelles est la fonction de répartition de sa loi de probabilité.

Ainsi :

$$F(x) = p(X \leq x) = \int_{t<x} f(t)\text{dt}\text{pour une loi diffuse de densité } f$$

3.1.6 Loi conjointe d'un couple (X,Y) de variables aléatoires diffuses

Pour un couple (X,Y), la densité est une fonction de deux variables $h(x, y)$ et :

$$p\left(X \in A \text{ et } Y \in B\right) = \int_{A \times B} h(x,y)\text{dx dy}$$

Ainsi la loi marginale de Y que l'on récupère est :

$$p(Y \in B) = \int_B \left(\int_\Omega h(x,y)\text{dx} \right) \text{dy}$$

La densité de cette loi marginale de Y est donc :

$$y \to \int_\Omega h(x,y)\text{dx}$$

Idem pour la loi marginale de X et généralisation immédiate au cas d'un n-upplet.

3.1.7 Cas de variables aléatoires diffuses, indépendantes

Propriété 3 *X et Y sont indépendantes ssi la densité $h(x, y)$ de la loi conjointe du couple (X,Y) est le produit des densités respectives des lois marginales de X et de Y.*

Preuve :

$$p(X \in A \text{ et } Y \in B) = \int_{A \times B} \text{dx dy}$$

Or X et Y sont indépendantes ssi $p(X \in A \text{ et } Y \in B) = p(X \in A).p(Y \in B)$

$$\text{ssi } p(X \in A \text{ et } Y \in B) = \int_A f(x)\text{dx} \int_B g(y)\text{dy}$$

où on a noté f la densité de X et g celle de Y.

3.1.8 Densité d'une somme

Soient X et Y étant deux variables aléatoires diffuses **indépendantes**. On note f la densité de X, g celle de Y et w celle de (X+Y). On a alors : $w = f * g$, produit de convolution de f et de g.

Preuve :

$$p(X+Y<s) = \int_{x+y<s} f(x)g(y)\mathrm{dx}\ \mathrm{dy}$$

(d'après 3.2.3)

On effectue le changement de variable : $\left\{ \begin{array}{l} u = x+y \\ v = y \end{array} \right.$ l'intégrale devient :

$$\int_{u<s} \left(\int f(u-v)g(v)\mathrm{dv} \right) du$$

donc la densité de la somme $X+Y$ est :

$$w = \int f(u-v)g(v)\mathrm{dv}$$

3.1.9 Densité d'une variable aléatoire image

X est une variable aléatoire diffuse (de densité f)à valeurs dans Ω. φ est une fonction injective par morceau sur Ω. On considère la variable aléatoire $\varphi(X)$ de densité g. On appelle ψ les bijections réciproques de φ sur chaque morceau (par exemple si $\varphi(X) = X^2$, il y a deux fonctions réciproques ψ : l'une correspond à $y \to \sqrt{y}$ (morceau des réels positifs) et l'autre à $y \to -\sqrt{y}$ (morceau des réels négatifs)). On a alors :

$$g(y) = \sum_{\psi} f(\psi(y)).\left| det\ Jacobienne\,(\psi(y)) \right|$$

Preuve :

$$p(\varphi(X) \in B) = p\left[\bigcup_{\psi} (X \in \psi(B)) \right] = \sum_{\psi} p\left[X \in \psi(B) \right] = \sum_{\psi} \int_{\psi(B)} f(x)\mathrm{dx}$$

Dans les intégrales, effectuons un changement de variable : $x = \psi(y)$. On obtient alors :

$$\sum_{\psi} \int_{B} f(\psi(y)).\left| det\ Jacobienne\,(\psi(y)) \right| \mathrm{dy} =$$

$$\int_{B} \sum_{psi} f(\psi(y)).\left| det\ Jacobienne\,(\psi(y)) \right| \mathrm{dy}$$

3.2 Espérance mathématique d'une variable aléatoire diffuse

3.2.1 Définition

Dans le cas d'une variable aléatoire diffuse X de densité f, l'espérance mathématique de X (ou plus simplement espérance de X) est définie par :

$$E(X) = \int x.f(x)\mathrm{dx}$$

où, naturellement, on suppose que l'intégrale est absolument convergente. Cette hypothèse qui s'énonce simplement ("*E(X) existe*") sera toujours sous-entendue par la suite.

3.2.2 Formule de transfert

X est une variable aléatoire diffuse de densité f) à valeurs dans Ω. φ est une fonction injective par morceau sur Ω. On considère la variable aléatoire $\varphi(X)$ de densité g. D'après 3.1.9, on a :

$$g(y) = \sum_{\psi} f(\psi(y)).\left|det\ Jacobienne\,(\psi(y))\right|$$

où on appelle ψ les bijections réciproques de φ sur chaque morceau. Par définition :

$$E(\varphi(X)) = \int y.g(y)\mathrm{dy}$$

Par la formule de transfert, on a :

$$E(\varphi(X)) = \int \varphi(x).f(x)\mathrm{dx}$$

<u>Preuve</u> : Dans l'intégrale, on effectue le changement de variable :

$$x = \psi(y)) \text{ et on obtient : } \int y.g(y)\mathrm{dy}$$

3.2.3 Propriétés

Propriété 4 *De manière analogue au cas discret, on montre :*

- *La **linéarité** de E*
- *Si X et Y sont **indépendantes**, alors $E(XY) = E(X).E(Y)$.*

3.3 Variance. Covariance. Coefficient de corrélation

Les notions de *variance*, *covariance*et de *coefficient de corrélation* sont définies directement à partir de l'espérance (qui a les mêmes propriétés que dans le cas discret). Il n'y a donc, formellement, rien à changer à ce qui est écrit au 2.3.

3.4 Exemples de variables aléatoires diffuses

3.4.1 Fonction caractéristique

X étant une variable aléatoire diffuse avec une densité f, on définit sa fonction caractéristique ϕ par : $\phi(u) = E\left(\mathbf{e}^{\mathbf{i}.u.X}\right)$

D'après la formule de transfert, on a :

$$\phi(u) = \int \mathbf{e}^{\mathbf{i}.u.x} f(x)\mathrm{dx}$$

ϕ est appelée *transformée de Fourier* de f au point $-\frac{u}{2\pi}$.

Il vient alors : $\begin{cases} E(X) = -\mathbf{i}\phi'(0) \\ Var(X) = -\phi''(0) + \left(\phi'(0)\right)^2 \end{cases}$

Si X et Y sont indépendantes la fonction caractéristique de $(X+Y)$ est le produit des fonctions caractéristiques de X et de Y.

Preuve : On a vu en 3.1.8 que la densité de la somme $(X+Y)$ était le produit de convolution des densités de X et de Y. Or la transformée de Fourier d'un produit de convolution est le produit des transformées de Fourier.

Cette propriété se généralise évidemment au cas de n variables aléatoires indépendantes.

3.4.2 Loi exponentielle

On rencontre cette loi essentiellement dans les problèmes de durée de vie. Si un appareil était mis en marche à l'instant t_0, X désignerait l'instant où il tomberait en panne. Cherchons la loi de X.

On note F la fonction de répartition : $F(t) = p(X \leq t)$.
On a :

$$\begin{aligned} p(X \leq t + dt | X > t) &= a(t).dt + \varepsilon(dt).dt \text{ (développement de Taylor)} \\ &= \tfrac{p(X \leq t+dt \cap X > t)}{p(X>t)} \\ &= \tfrac{F(t+dt)-F(t)}{1-F(t)} \end{aligned}$$

D'où :

$$\frac{F(t+dt) - F(t)}{dt} = (a(t) + \varepsilon(t)).(1 - F(t))$$

On suppose une hypothèse de non-vieillissement : $\forall t$, $a(t) = a$. En faisant tendre dt vers 0, on obtient : $F^{'}(t) = a.(1 - F(t))$

En résolvant cette équation différentielle, on obtient : $F(t) = 1 + c.\mathbf{e}^{-\mathbf{at}}$. On suppose $F(t_0) = 0$. On a alors : $c = -\mathbf{e}^{\mathbf{at_0}}$.

Ainsi le résultat est :

$$F(t) = 1 - \mathbf{e}^{a(t_0 - t)} \text{ , pour } t > t_0$$

D'où :

$$f(t) = a\mathbf{e}^{a(t_0 - t)} \text{ , pour } t > t_0$$

On dit qu'il s'agit de <u>la loi exponentionnelle de paramètre a et de début t_0</u>.
On trouve :

$$E(X) = \frac{1}{a} + t_0 \text{ et } Var(X) = \frac{1}{a^2}$$

.

3.4.3 Loi de Gauss ou loi Normale

C'est la loi la plus importante pour les applications aux statistiques. On appelle loi de Gauss ou loi Normale $N(0{,}1)$ la loi de densité :

$$\forall x \in \mathbb{R} \ , \ f_{0,1}(x) = \frac{1}{\sqrt{2}}\mathbf{e}^{-\frac{1}{2}\curvearrowleft^2}$$

Si X suit une loi Normale $N(0{,}1)$, un calcul immédiat montre que : $E(X) = 0$ et $\sigma_X = 1$.

Par contre, $F(x) = p(X \leq x)$ ne peut se calculer analytiquement et nécessite un calcul numérique. Les résultats sont consignés dans des tables (voir appendice).

Si X suit la loi Normale $N(0{,}1)$, la variable aléatoire $Y = \sigma.X + m$ (où $\sigma > 0$) suit la loi de densité $f(y) = \frac{1}{\sigma}.f_{0,1}(\frac{y-m}{\sigma})$ d'après 3.1.9. Cette loi est appelée loi Normale $N(m,\sigma)$.

On a alors : $E(Y) = m$ et $\sigma_Y = \sigma$.

Conséquence pratique : pour faire des calculs de probabilité concernant Y qui suit la loi $N(m,\sigma)$, on pose $X = \frac{Y-m}{\sigma}$. X suit la loi $N(0{,}1)$ et on fait alors les calculs avec les tables.

Théorème 5 *Si Y_1 suit la loi $N(m_1,\sigma_1)$ et Y_2 la loi $N(m_2,\sigma_2)$, si Y_1 et Y_2 sont indépendantes, alors $Y_1 + Y_2$ suit une loi normale N(m,σ)*

Preuve : La fonction caractéristique de Y_1 est $\mathbf{e}^{\mathbf{i}um_1 - \frac{1}{2}(u\sigma_1)^2}$ (c'est un calcul de transformée de Fourier). Celle de Y_2 est $\mathbf{e}^{\mathbf{i}um_2 - \frac{1}{2}(u\sigma_2)^2}$

Donc celle de la somme $Y_1 + Y_2$ en est le produit (d'après 3.4.1 grâce à l'indépendance de Y_1 et Y_2) : $\mathbf{e}^{\mathbf{i}um - \frac{1}{2}(u\sigma)^2}$, avec $m = m_1 + m_2$ et $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Il s'ensuit que toute combinais on linéaire de variables aléatoires indépendantes qui suivent des lois normales suit une loi normale.

Cas particulier important : Soit $X_1, X_2, \ldots, X_N$ un échantillon de modèle X qui suit la loi N(m,σ). Alors $\overline{X}$ suit la loi N(m,$\frac{\sigma}{\sqrt{N}}$).

N-uplet Gaussien

On se donne des coefficients m_i, $\sigma_i > 0$ et des coefficients $r_{i,j}$ tels que $\forall(i,j) \in [1,N]^2$, $|r_{i,j}| < 1$ et $r_{i,i} = 1$. On appelle C la matrice de $\mathcal{M}_N(\mathbb{R})$ dont les coefficients sont $c_{i,j} = r_{i,j}.\sigma_i.\sigma_j$

A tout vecteur réel $\overrightarrow{x}$ de dimension N on associe le vecteur $\overrightarrow{x}_m$ (matrice uni-colonne) de composantes $x_i - m_i$. ${}^t\overrightarrow{x}_m$ désigne alors la matrice transposée (uni-ligne).

On considère la fonction :

$$f(x) = \frac{1}{(2\pi)^N\sqrt{det(C)}}.\mathbf{e}^{-\frac{1}{2}({}^t x_m C^{-1} x_m)}$$

Cette fonction est appelée loi gaussienne en dimension N, loi de densité de f.

On peut calculer que la loi marginale de chaque X_i est la loi $N(m_i,\sigma_i)$ et les coefficients $r_{i,j}$ sont les coefficients de corrélation entre X_i et X_j.

Explicitons la formule de f en dimension 2 :

$$f(x_1,x_2) = \frac{\mathbf{exp}\left[-\frac{1}{2.(1-r^2)}.\left(\left(\frac{x_1-m_1}{\sigma_1}\right)^2 - 2r.\frac{(x_1-m_1).(x_2-m_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-m_2}{\sigma_2}\right)^2\right)\right]}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}$$

On remarque que dans ce cas si le coefficient de corrélation entre X_1 et X_2 est nul ($r = 0$), alors la densité de f peut s'écrire comme le produit d'une fonction de x_1 et d'une fonction de x_2. Ceci montre que les variables aléatoires X_1 et X_2 sont indépendantes.

Chapitre 4

Loi des grands nombres et théorème central limite

4.1 Loi des grands nombres

4.1.1 Inégalité de Markoff

<u>**Théorème**</u> **6** *Soient $a > 0$ et X une variable aléatoire à valeurs positives. On a alors : $p(X \geq a.E(X)) \leq \frac{1}{a^2}$.*

<u>Preuve</u> : X peut être discrète ou diffuse. On fera la démonstration dans le cas diffus avec densité f (le cas discret est analogue). Si $a \leq 1$, c'est trivial. Sinon, :

$$E(X) = \int_0^\infty x.f(x)\mathrm{dx} \geq \int_{a.E(X)}^\infty x.f(x)\mathrm{dx}$$

$$\geq a.E(X) \int_{a.E(X)}^\infty f(x)\mathrm{dx} \geq a.E(X)p(X \geq a.E(X))$$

4.1.2 Inégalité de Tchebitcheff

<u>**Théorème**</u> **7** *Soient $b > 0$ et Y une variable aléatoire telle que $E(Y^2)$ existe. On note σ son écart-type. On a alors :*

$$p\left(|Y - E(Y)| \geq b.\sigma\right) \leq \frac{1}{b^2}$$

<u>Preuve</u> : Il suffit d'appliquer l'inégalité de Markoff à la variable aléatoire $X = (Y - E(X))^2$, avec $a = b^2$ (car $E(X) = \sigma^2$)

4.1.3 Convergence en probabilité et loi des grands nombres

On dit qu'une suite de variables aléatoires (Y_n) converge en probabilité vers un nombre a ssi $\forall \varepsilon > 0$, $p(|Y_n - a| > \varepsilon)$ converge vers 0 pour n tendant vers ∞.

<u>**Théorème**</u> **8** ***Loi des grands nombres.***

Soit X une variable aléatoire telle que $E(X)$ existe. $X_1, X_, \ldots, X_N$ désigne un échantillon de modèle X de taille N. On a alors :

$\overline{X}$ converge en probabilité vers $E(X)$ quand N tend vers ∞.

<u>Preuve</u> : On applique l'inégalité de Tchebitcheff à $\overline{X}$. Cela donne :

$$p\left(|\overline{X} - E(X)| \geq b.\sigma_X N^{\frac{1}{2}}\right) \leq \frac{1}{b^2} \left(\text{choisir } b = \varepsilon \frac{N^{\frac{1}{2}}}{\sigma_X}\right)$$

4.2 Théorème central limite

4.2.1 Convergence en loi

On dit qu'une suite de variables aléatoire (X_N) à valeurs réelles converge en loi vers une variable aléatoire X à valeurs réelles ssi la suite des fonctions de répartition desX_N converge simplement vers la fonction de répartition de X.

On a montré, par exemple en 2.4.3, que si les X_N suivent la loi binomiale d'ordre N et de paramètre $\frac{a}{N}$ (avec a fixe), alors les X_N convergent en loi vers X qui suit la loi de Poisson de paramètre a.

4.2.2 Cas de variables aléatoires diffuses

Théorème **(P. Levy) 9** *Les X_N étant diffuses, elles convergent en loi vers X ssi leurs fonctions caractéristiques convergent simplement vers la fonction caractéristique de X.*

Théorème Central Limite 10 *Soit X une variable aléatoire diffuse de densité f telle que $E(X^3)$ existe et $E(X) = 0$. Soit $X_1, X_2, \ldots, X_N$ un échantillon de modèle X et de taille N. On considère la suite de variables aléatoires définies par :*

$$Y_N = \frac{1}{sqrtN} \sum_{k=1} NX_k$$

On a alors : (Y_N) converge en loi vers la loi Normale $N(0,\sigma_X)$.

Preuve : Considérons ϕ la fonction caractéristique de $\frac{1}{\sqrt{N}}X$.
On a donc :

$$\phi(u) = E\left(\mathbf{e}^{\mathbf{i}\frac{u}{sqrtN}X}\right) = \int f(t).\mathbf{e}^{\mathbf{i}\frac{u}{\sqrt{N}}}\mathrm{dt}$$

On développe l'exponentielle par la formule de Taylor :

$$\mathbf{e}^{\mathbf{i}\frac{u.t}{\sqrt{N}}} = 1 + \mathbf{i}\frac{ut}{\sqrt{N}} - \frac{u^2t^2}{N} - \mathbf{i}\frac{u^3t^3}{6N^{\frac{3}{2}}}$$

Donc :

$$\phi(u) = \int f(t).(1 + \mathbf{i}\frac{ut}{\sqrt{N}} - \frac{u^2t^2}{N} - \mathbf{i}\frac{u^3t^3}{6N^{\frac{3}{2}}})\mathrm{dt}$$

$$\phi(u) = 1 + 0 - \frac{u^2.Var(X)}{2N} - reste_{(N)}$$

Grâce à l'hypothèse $E(X^3)$ existe, on majore $|reste_{(N)}|$ par $\mathrm{c}^{ste}.N^{\frac{2}{3}}$. Comme Y_N est une somme de N variables aléatoires **indépendantes** et suivant la loi de X, la fonction caractéristique de Y_N est :

$$\phi(u)^N = \left(1 - \frac{u^2.Var(X)}{2N} - reste_{(N)}\right)^N$$

La majoration du reste montre que la limite (lorsque N tend vers ∞) est :

$$\mathbf{exp}\left(-\frac{1}{2}(u.\sigma_X)^2\right)$$

qui est la fonction caractéristique de la loi Normale $N(0,\sigma_X)$.

L'utilisation heuristique (en statistiques) de ce théorème central limite fait que l'on cherchera une approximation de la loi d'une variable aléatoire Y (qui est la somme d'un grand nombre de variables aléatoires indépendantes) sous la forme d'une loi Normale. C'est ce qui explique le rôle important joué par cette loi normale en statistique.

Extension de l'utilisation de ce type d'approximation :

$$\text{Si une variable aléatoire } Y \text{ se présente sous la forme : } \sum_{k=1}^{N} X_k$$

avec N suffisamment grand (X est non nécessairement centrée), alors on a (d'après ce qui précède) :

$$\frac{1}{\sqrt{N}}\left(Y - \sum_{k=1}^{N} E(X_k)\right) = \frac{1}{\sqrt{N}}\sum_{k=1}^{N}(X - k - E(X_k)) \text{ suit une loi } N(0,\sigma)$$

$$\text{Donc : Y suit une loi } N\left(\sum_{k=1}^{N} E(X_k)\ ,\ N^{\frac{1}{2}}\sigma\right)$$

4.2.3 Exemple : loi du Chi-deux

Soit $X_1, X_2, ..., X_N$ un échantillon de modèle X qui suit la loi N(0,1). On considère la variable aléatoire U_n définie par :

$$U_n = \sum_{k=1}^{N} X_k^2$$

On dit que U_n suit une loi du Chi-deux à n degrés de liberté.

Remarque : Les calculs de probabilités concernant U_n nécessitent un calcul d'intégrale qui ne peut se faire analytiquement (cependant comme il s'agit d'un calcul d'intégrale en dimension n on ne peut non plus procéder directement avec un calcul numérique ; on commence par faire un changement de variables qui permet de se ramener au calcul d'une intégrale en dimension 1 (voir détails dans le cours de statistiques)). Les résultats de ces calculs numériques d'intégrales sont consignés dans des tables. Cependant, il faut une table pour chaque valeur du paramètre n. On doit donc bien se limiter pour n.

Il a été vérifié que pour $n > 30$ on avait une bonne approximation de la loi de U_n par la loi normale $N(n, \sqrt{(2n)})$.

Chapitre 5

Simulation numérique des lois de probabilités

5.1 Principe général

5.1.1 Définition d'une simulation numérique

Simuler numériquement une loi de probabilité p consiste à donner les valeurs $x_1, x_2, \ldots, x_N$ obtenues lors d'une réalisation de $X_1, X_2, \ldots, X_N$, un échantillon de X suivant la loi p.

On notera R une variable aléatoire qui suit la loi uniforme sur l'intervalle [0,1]. Puis, à partir de R, on forme une variable aléatoire X qui suit la loi p demandée.

Ainsi, à partir d'une simulation $r_1, r_2, \ldots, r_N$ (en principe connue) de R, on obtient les $x_1, x_2, \ldots, x_N$ cherchées et qui simulent la loi p.

5.1.2 Obtention de la simulation de la loi uniforme sur [0,1]

Quelques fois, il est utilisé des propriétés de phénomènes physiques pour obtenir les $r_1, r_2, \ldots, r_N$ qui simulent la loi uniforme sur [0,1].

Nous allons ici simplement décrire la méthode numérique la plus utilisée pour obtenir ces valeurs r_i.

Considérons la formule de récurrence : $x_{n+1} = a.x_n[M]$ où M est un nombre entier (le plus grand possible pour la machine considérée afin de retarder le moment où on obtient un entier déjà obtenu). On prend $M = 2^q$ pour une machine à q bits et a premier avec M.

On prend $a = 3^k$, avec k tel que $a \simeq \frac{M}{2}$.

Le premier nombre entier x_0 est pris tel que $0 < x_0 < M$ (il est initialisé avec l'horloge ou un autre moyen physique). On obtient ainsi une suite (x_n) de nombres entiers tels que $0 < x_n < M$.

Alors, les r_n recherchés sont les réels $r_n = \frac{x_n}{M}$. On obtient ainsi une suite de réels tels que : $0 < r_n < 1$.

En réalisant des tests statistiques (par exemple Chi-deux que l'on verra dans la partie Statistiques de ce module) on est amené à considérer que cette suite simule correctement la loi uniforme sur [0,1].

5.1.3 Méthode de la transformée inverse

La méthode (dite de transformation inverse) pour obtenir une simulation numérique d'une loi de densité continue f se déduit immédiatement du théorème ci-dessous.

Théorème 11 *Soit f la densité d'une loi. Soit F la fonction de répartition de cette loi. F est une bijection croissante de l'intérieur du support de f sur]0,1[. Donc il existe une bijection F^{-1}.*

Alors la variable aléatoire $X = F^{-1}(R)$ suit la loi de densité f (où R, conformément aux notations du 5.1, suit la loi uniforme sur]0,1[).

Preuve : La fonction de répartition de X est :

$$p(X \leq x) = p(F^{-1}(R) \leq x) = p(R \leq F(x)) = F(x)$$

5.2 Exemples de simulation

5.2.1 Simulation de la loi uniforme sur un intervalle [a,b]

C'est une application de la méthode de la trnasformation inverse, vue en 5.1.3. La loi à simuler est la loi de densité $f = constante = \frac{1}{b-a}$. Sur cet intervalle $[a,b]$, sa fonction de répartition est $F(x) = \frac{x-a}{b-a}$.

Donc la fonction réciproque est $F^{-1}(r) = a + r.(b-a)$.

La suite simulat la loi uniforme sur [a,b] est donc : $x - N = a + r_n.(b-a)$.

5.2.2 Simulation de la loi exponentionnelle

C'est une application de la méthode de la transformation inverse (vue en V.1.3). La loi exponenetielle, de début 0 et de paramètre a, admet pour densité la fonction $f(x) = a.\mathbf{e}^{-a.x}$ pour $x > 0$ et $f(x) = 0$ pour $x < 0$.

On a donc : $F(x) = 1 - \mathbf{e}^{-a.x}$ qui est une bjection de l'ensemble des réels positifs sur l'intervalle]0,1[.

$$\text{Sa fonction réciproque vaut : } F^{-1}(r) = -\frac{1}{a.ln(1-r)}$$

La suite qui simule la loi exponentielle voulue est :

$$(x_n)_{n\in\mathbb{N}} = \left(-\frac{1}{a.ln(1-r_n)}\right)_{n\in\mathbb{N}}$$

Remarquer que l'on peut remplacer la suite $(1 - r_n)$ par la suite (r_n).

5.2.3 Simulation de la loi de Poisson de paramètre a

Elle se déduit immédiatement du théorème ci-dessous (car on vient de voir comment simuler la loi exponenetielle de début 0 et de paramètre a).

Théorème 12 *Soit $Y_1, Y_2, \dots Y_n$ un échnatillon de modèle Y qui suit la loi exponentielle de début 0 et de paramètre a. Soit X le plus grand entier telle que :*

$$\sum_{i=1}^{X} Y_i \leq 1$$

Alors X suit la loi de Poisson de paramètre a.

Preuve :

$$p(X = k) = p(\sum_{i=1}^{k} Y_i \leq 1 et \sum_{i=1}^{k+1} > 1$$

$$p(X = k) = \int_A a^{k+1}\mathbf{e}^{-a.\sum_{i=1}^{k+1} y_i}.\mathrm{d}y_1\ \mathrm{d}y_2 \dots \mathrm{d}y_n$$

$$\text{où } A \text{ est définie par : } \sum_{i=1}^{k} y_i \leq 1 \text{ et } \sum_{i=1}^{k+1} y_i > 1 \text{ et } y_i > 0$$

(car on rappelle que la densité de la loi exponentielle de début 0 et de paramète a est $f(y_i) = a.\mathbf{e}^{-a.y_i}$ pour $y_i > 0$).

On calcule cette intégrale :

$$a^{k+1}\int_{y_i>0 \text{ et } \sum_{i=1}^{k} y_i<1} \left(\int_{1-\sum_{i=1}^{k}}^{\infty} \mathbf{e}^{-a.y_{k+1}}\mathrm{d}y_{k+1}\right).\mathbf{exp}\left(-a.\sum_{i=1}^{k} y_i\right)\mathrm{d}y_1\ldots\mathrm{d}y_k$$

Or on a :

$$\int_{1-\sum i=1^k}^{\infty} \mathbf{e}^{-a.y_{k+1}}\mathrm{d}y_{k+1} = \frac{1}{a}.\mathbf{exp}\left[-a.\left(1-\sum_{i=1}^{k} y_i\right)\right]$$

Donc l'intégrale considérée vaut :

$$a^k.\mathbf{e}^{-a}\int_{y_i>0 \text{ et } \sum i=1^k y_i<1} \mathrm{d}y_1\ldots\mathrm{d}y_k = a^k\frac{\mathbf{e}^{-a}}{k!}$$

(le calcul de la dernière intégrale ci dessus se fait par récurrence).

5.2.4 Simulation de la loi normale réduite

Elle se déduit immédiatement du théorème ci-dessous :

Théorème 13 *Soient ρ suivant la loi exponentielle de début 0 et de paramètre 1 et θ suivant la loi uniforme sur [0,2π], ϕ et θ étant **indépendantes**.*

On a alors, en posant $r=\sqrt{2\rho}$, $X=r.\cos\theta$ et $Y=r.\sin\theta$: X et Y sont indépendantes et suivent la loi normale réduite N(0,1).

Preuve :

Le couple (θ,ρ) admet $f(\theta,\rho)=\frac{1}{2\pi}.\mathbf{e}^{-\rho}$comme densité sur$[0,2.\pi]x\mathbb{R}^+$

On considère la bijection $\varphi : (\theta,\rho)\rightarrow(x=r.\cos\theta, y=r.\sin\theta)$ et on note ψ la bijection réciproque. (X,Y) est le couple image de (θ,ρ) par la bijection φ ci-dessus.

Donc sa densité (d'après 3.1.9) est :

$$g(x,y)=f(\psi(x,y)).|Jacobienne(\psi(x,y))|$$

$$\text{Soit : } g(x,y)=\frac{1}{2\pi}\mathbf{e}^{-\frac{1}{2}.(x^2+y^2)}$$

(car $|Jacobienne(\psi(x,y))|=|Jacobienne(\varphi(x,y))|=1$)

Conclusion D'après tout ce que l'on vient de considérer, pour simuler une loi normale réduite on simule le couple (θ,ρ) en prenant par exemple la suite $(\theta_n\ ,\ \rho_n) = (2\pi r_n\ ,\ -\ln r'_n)$ (remarquer qu'on prend les mêmes r_n pour former θ_n et ρ_n sinon on n'aurait pas l índépendance de θ et de ρ) Enfin on forme les suites $x_n = \sqrt{2\rho_n}\cos\theta_n$ et $y_n = \sqrt{2\rho_n}\sin\theta_n$.

On en déduit une simulation de la loi normale générale $N(m,\sigma)$ à partir d'une simulation de la loi normale réduite $N(0,1)$ du fait que si X suit une loi normale réduite, alors $Y = m + \sigma.X$ suit la loi normale $N(m,\sigma)$.

Chapitre 6

Processus de Markov discrets

6.1 Généralités

6.1.1 Définitions

On considère une variable aléatoire qui évolue par étapes successives. On note X_n cette variable aléatoire à l'étape numéro n. On dit qu'on a un processus aléatoire discret en temps. Si chaque X_n est une variable aléatoire discrète, on dit que le processux aléatoire est discret en valeurs. Un processus aléatoire discret en *en temps* et en *en valeurs* sera simplement appelé **processus aléatoire discret**.

Par la suite, nous ne considérerons que des processus aléatoires discrets.

Notons $e_k (k \in \mathbb{N})$ les valeurs possibles des variables aléatoires discrets X_n. Un processus aléatoire discret sera appelé processus aléatoire discret ed Markov ssi il vérifie :

$$\begin{aligned}\forall n > m\ ,\ & p(X_{m+1} = e_{m+1}, \ldots, X_n = e_n | X_m = e_m, X_{m-1}, \ldots, X_0 = e_0) \\ & = p(X_{m+1} = e_{m+1}, \ldots, X_n = e_n | X_m = e_m)\end{aligned}$$

Si on considère que m représente le présent, la propirété ci-dessus se résume par la phrase suiavnte (qui permet de reconnaître facilement qu'un processus aléatiore dee Markov) : "Les états futurs du processus aléatoires ne dépendent que du présent, mais pas du passé"

Exemple Les jeux se jouant en plusieurs parties (sans mémoire du passé) sont des processus discrets de Markov.

On appelle *transitions* les : $p(X_k = e_k | X_{k-1} = e_{k-1})$

6.1.2 Propriété

Propriété 14 *La loi d'un processus aléatoire discret de Markov ne dépend que des transitions et de la loi de* X_0.

Preuve : La loi du processus aléatoire discret de Markov est la famille des probabilités $p(X_n = e_n, \ldots, X_0 = e_0)$ pour tout choix de n et des e_k. Or :

$$p(X_n = e_n, X_{n-1}, \ldots, X_0 = e_0) =$$

$$p(X_n = e_n | X_{n-1} = e_{n-1}, | \ldots, X_= e_0).p(X_{n-1} = e_{n-1}, \ldots, X_0 = e_0)$$

(c'est la formule de Bayes des probabilités conditionnelles).

Or :

$$p(X_n = e_n | X_{n-1} = e_{n-1}, \ldots, X_0 = e_0) = p(X_n = e_n | X_{n-1} = e_{n-1})$$

(c'est la propriété de Markov)

Donc : $p(X_n = e_n, \ldots, X_0 = e_0) =$
$p(X_n = e_n | X_{n-1} = e_{n-1}).p(X_{n-1} = e_{n-1}, \ldots, X_0 = e_0)$

En itérant ce procédé de calcul, on obtient :

$$p(X_n = e_n, \ldots, X_0 = e_0) = \prod_{k=1}^{n} p(X_k = e_k | X_{k-1} = e_{k-1}).p(X_0 = e_0)$$

(*loi conjointe* = produit des *transitions* fois *loi de* X_0)

6.2 Chaîne de Markov

6.2.1 Définitions

Un processus aléatiore discret de Markov est appelé *chaîne de Markov* ssi il est *invariant par translation dans le temps.* C'est à dire :

$$\forall (k, n) \ , \forall (e, e^{'}) \ , p(X_n = e^{'} | X_0 = e) = p(X_{k+n} = e^{'} | X_k = e)$$

Exemple : Les jeux se jouant en plusieurs parties (sans mémoirre du passé) et ne changeant pas de règle au cours du temps sont des chaînes de Markov.

6.2.2 Matrice de translation

Dans le cas où les valeurs possibles sont en nombre **fini** N, on notera simplement celles-ci : $1, 2, \dots, N$ (au lieu de $e_1, \dots, e_N$) et on notera $G_{i,j}$ les transitions $p(X_1 = j | X_0 = i)$. G est la matrice dont les coefficients sont les $G_{i,j}$ et est appelée *matrice de transition* .

Remarque : Une matrice de transition G vérifie :

$$\forall i \in [1, N] \ , \ \sum_{j=1}^{N} G_{i,j} = 1$$

Propriété 15 *Une matrice de transition G admet 1 comme valeur propre. Ses autres valeurs propres λ sont telles que $|\lambda| \leq 1$.*

Preuve : D'après la remarque ci-dessus, le vecteur dont toutes les composantes sont égales à 1 est un vecteur propre pour la valeur propre 1.

Soit λ une autre valeur propre et X le vecteur propre associé. On a :

$$\forall i \in [1, N] \ , \ \sum_{j=1}^{N} G_{i,j}.X_j = \lambda.X_i$$

On note i_0 l'indice tel que : $|X_{i_0}| = sup_{i \in [1,n]} |X_i|$

On a donc, pour $i = i_0$:

$$\sum_{j=1 \text{ et } j \neq i_0}^{N} G_{i_0,j}.X_j = (\lambda - G_{i_0,i_0}).X_{i_0}$$

D'où :

$$|(\lambda - G_{i_0,i_0})|.|X_{i_0}| \leq \sum_{j \neq i_0} G_{i_0,j}.|X_j| \leq |X_{i_0}|.\sum_{j \neq i_0} G_{i_0,j}$$

Donc, on a :

$$|(\lambda - G_{i_0,i_0})| \leq \sum_{j \neq i_0} G_{i_0,j} = 1 - G_{i_0,i_0}$$

Ainsi :

$$|\lambda| \leq 1 - G_{i_0,i_0} + G_{i_0,i_0} = 1$$

6.2.3 Théorème fondamental

L'évolution du processus est décrite par la matrice $G^{(k)}$ dont les coefficents sont les $G_{i,j}^{(k)} = p(X_k = j | X_0 = i)$

Théorème 16 $\forall k \in [1, N]$, $G^{(k)} = G^k$

Preuve : Par récurrence sur k.

La propriété est vraie pour $k = 1$. Supposons la propriété vraie jusqu'à l'indice k. On a alors :

$$G_{i,j}^{(k+1)} = p(X_{k+1} = j | X_0 = i) = \sum_{n=1}^{N} p(X_{k+1} = j | X_k = n).p(X_k = n | X_0 = i)$$

(formule des probabilités totales)

Or : $p(X_{k+1} = j | X_k = n) = G_{i,j}^{(k+1)}$ car c'est une chaîne de Markov.

De plus : $p(X_k = n | X_0 = i) = G_{i,n}^{(k)} = G_{i,n}^k$, par hypothèse de récurrence.

On a donc :

$$G_{i,j}^{(k+1)} = \sum_{n=1}^{N} G_{i,n}^k.G_{n,j} = G_{i,j}^{k+1}$$

6.2.4 Exemple de marche aléatoire

A chaque étape un "*promeneur* " peut se déplacer soit vers la gauche (avec une probabilité p), soit ver la droite (avec une probabilité q), soit rester à la même place (avec une probabilité $r = 1 - p - q$). Il y a un nombre fini de positions possibles. C'est à dire qu'il y a des bords (absorbants ou réfléchissants, selon la nature du problème).

C'est un processus aléatoire discret de Markov, qui est, de plus, une chaîne de Markov.

Considérons le problème avec 4 positions possibles et bords absorbants, pour $p = q = \frac{1}{2}$.

La matrice des transitions G est dons la matrice suivante : $\begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Le calcul de G^n donne les probabiltés pour le promeneur d'être, à l'étape n$^\circ$ n, à la place n$^\circ$ j sachant qu'il est parti de la place n$^\circ$ i.

On a :

$$G^n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} + \frac{1}{2^{n+1}}. \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$+\frac{(-1)^n}{2^{n+1}}. \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{3} & 1 & -1 & \frac{1}{3} \\ \frac{1}{3} & -1 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

6.2.5 Etude asymptotique

On veut étudier l'évolution du processus lorsque le nombre d'étapes n tend vers l'infini. Dans l'exemple précédent (en 6.2.4) : $\lim\limits_{n\to\infty} G^n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Réfléchir sur l'interprétation de ce résultat.

6.3 Ergodicité

6.3.1 Vecteur d'état d'une chaîne de Markov X_k, où $k \in [1, N]$

Notons q^k le vecteur dont la composante n?i est : $q_i^k = p(X_k = i)$.

<u>**Remarque**</u> $\forall k \in [1, N]$, $||q^k||_1 = 1$

Notons P la matrice transposée de la matrice de transition G : $P = {}^tG$

<u>**Propriété**</u> **17** $\forall k \in [1, N-1]$, $P.q^k = q^{k+1}$

<u>Preuve</u> : Soit $k \in [1, n-1]$. On a : $(P.q^k)_i = \sum_{j=1}^{N} P_{i,j}.q_j^k = \sum_{j=1}^{N} G_{j,i}.q_j^k$

Soit : $(P.q^k)_i = \sum j = 1^N p(X_{k+1} = i|X_k = j).p(X_k = j) = p(X_{k+1} = i) = q^{k+1}$
(car pour une chaîne de Markov, on a : $G_{i,j} = p(X_{k+1} = i|X_k = j)$)

<u>**Corollaire**</u> **18** $\forall k \in [1, N-1]$, $q^k = P^k.q0$

6.3.2 Vecteur d'état asymptotique d'une chaîne de Markov

On appelle vecteur d'état asymptotique la limite, lorsqu'elle existe des vecteurs d'états q^k. On notera cette limite q^∞.

$$\text{On a donc : } q^\infty = \lim_{k\to\infty} q^k$$

<u>Remarque</u> : Comme on a : $\forall k \in \mathbb{N} ||q^k|| = 1$, on a aussi : $||q^\infty|| = 1$.

<u>Définition</u> 19 *On dit que la chaîne de Markov est ergodique ssi q^∞ existe et ne dépend pas de q^0.*

<u>Théorème</u> 20 *Si 1 est <u>valeur propre **simple**</u> de G (matrice de transition) et si les autres valeurs propres λ sont telles que $|\lambda| < 1$, alors la chaîne de Markov est ergodique.*

<u>Preuve</u> : G et P ont les mêmes valeurs propres. Supposons, pour des raisons de simplification des calculs, que P soit diagonalisable (sinon, il faut remplacer la diagonalisation par *une réduction de Jordan*, ce qui rend les calculs plus "techniques").

Considérons une base de vecteurs propres de P, soit '$(e_1, \ldots, e_N)$, telle que :

$\{beginarrayce_1$ est associé à la valeur propre $1 \forall i \in [1, N]$, $||e_i|| = 1$

On décompose q^0 sur cette base.

$$\text{On obtient : } q^0 = a_1.e_1 + \sum_{i=2}^{N} a_i.e_i$$

Or : $\forall k \in \mathbb{N}$, $q^k = P^k.q^0$.

$$\text{Donc on a : } \forall k \in \mathbb{N} \text{ , } q^k = a_1.e_1 + \sum_{i=2}^{N} a_i.(\lambda_i)^k.e_i$$

On sait que : $\forall i \in [1, N]$, $|\lambda_i| < 1$. Donc, par passage à la limite sur k, on obtient : $q^\infty = a_1.e_1$.

Or, les normes 1 de q^∞ et de e_1 étant égales à 1, on a : $a_1 = \pm 1$. Comme les q_i^k sont positifs pour tout k dans $\mathbb{N}$ et $i \in [1, N]$, on a aussi : $\forall i \in [1, N]$, $q_i^\infty \geq 0$. q^∞ est donc **le** vecteur propre de P, unitaire pour la norme 1, de composantes toutes positives (il ne dépend pas de q^0 donc).

6.3.3 Exemple

Considérons le cas de la marche aléatoire (vue en V.2.4) avec 3 positions possibles, des bords réfléchissants et $p = q = \frac{1}{3}$. Les matrices G et P sont donc les suivantes :

$$G = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 1 & 0 \end{pmatrix} \text{ et } P = \begin{pmatrix} 0 & \frac{1}{3} & 0 \\ 1 & \frac{1}{3} & 1 \\ 0 & \frac{1}{3} & 0 \end{pmatrix}$$

Les valeurs propres de P (et de G) sont 1, 0 et $-\frac{2}{3}$. Les hypothèses du théorème du VI.3.2 sont donc vérifiées. Cette chaîne de Markov est donc ergodique et le vecteur q^∞ existe bien et est le vecteur propre de P, unitaire pour la norme 1 et associé à la valeur propre 1 et de composantes toutes positives. C'est le vecteur :

$$q^\infty = \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \\ \frac{1}{5} \end{pmatrix}$$

Réfléchir sur l'interprétation de ce résultat.

Annexe 1 : Valeur de la fonction de répartition ϕ de la variable réduite de Gauss Z

$$\underline{\text{Rappel}} : \phi(t) = P_Z(]-\infty, t[= \frac{1}{\sqrt{2\pi}} . \int_{-\infty}^{t} \mathbf{e}^{-\frac{z^2}{2}} dz$$

t	$\phi(t)$	t	$\phi(t)$	t	$\phi(t)$	t	$\phi(t)$	t	$\phi(t)$	t	$\phi(t)$
0,00	0,5000	1,00	0,8413	2,00	0,9772	0,50	0,6915	1,50	0,9332	2,50	0,9938
0,01	0,5040	1,01	0,8437	2,01	0,9779	0,51	0,6950	1,51	0,9345	2,51	0,9940
0,02	0,5080	1,02	0,8461	2,02	0,9783	0,52	0,6985	1,52	0,9357	2,52	0,9941
0,03	0,5120	1,03	0,8485	2,03	0,9788	0,53	0,7019	1,53	0,9370	2,53	0,9943
0,04	0,5160	1,04	0,8508	2,04	0,9793	0,54	0,7054	1,54	0,9382	2,54	0,9945
0,05	0,5199	1,05	0,8531	2,05	0,9798	0,55	0,7088	1,55	0,9394	2,55	0,9946
0,06	0,5239	1,06	0,8554	2,06	0,9803	0,56	0,7123	1,56	0,9406	2,56	0,9948
0,07	0,5279	1,07	0,8577	2,07	0,9808	0,57	0,7157	1,57	0,9418	2,57	0,9949
0,08	0,5319	1,08	0,8599	2,08	0,9812	0,58	0,7190	1,58	0,9429	2,58	0,9951
0,09	0,5359	1,09	0,8621	2,09	0,9817	0,59	0,7224	1,59	0,9441	2,59	0,9952
0,10	0,5398	1,10	0,8643	2,10	0,9821	0,60	0,7257	1,60	0,9452	2,60	0,9953
0,11	0,5438	1,11	0,8665	2,11	0,9826	0,61	0,7291	1,61	0,9463	2,61	0,9955
0,12	0,5478	1,12	0,8686	2,12	0,9830	0,62	0,7324	1,62	0,9474	2,62	0,9956
0,13	0,5517	1,13	0,8708	2,13	0,9834	0,63	0,7357	1,63	0,9484	2,63	0,9957
0,14	0,5557	1,14	0,8729	2,14	0,9838	0,64	0,7389	1,64	0,9495	2,64	0,9959
0,15	0,5696	1,15	0,8749	2,15	0,9842	0,65	0,7422	1,65	0,9505	2,65	0,9960
0,16	0,5636	1,16	0,8770	2,16	0,9846	0,66	0,7454	1,66	0,9515	2,66	0,9961
0,17	0,5675	1,17	0,8790	2,17	0,9850	0,67	0,7486	1,67	0,9525	2,67	0,9962
0,18	0,5714	1,18	0,8810	2,18	0,9854	0,68	0,7517	1,68	0,9535	2,68	0,9963
0,19	0,5753	1,19	0,8830	2,19	0,9857	0,69	0,7549	1,69	0,9545	2,69	0,9964
0,20	0,5793	1,20	0,8849	2,20	0,9861	0,70	0,7580	1,70	0,9554	2,70	0,9965
0,21	0,5832	1,21	0,8869	2,21	0,9864	0,71	0,7611	1,71	0,9564	2,71	0,9966
0,22	0,5871	1,22	0,8888	2,22	0,9868	0,72	0,7642	1,72	0,9573	2,72	0,9967
0,23	0,5910	1,23	0,8907	2,23	0,9871	0,73	0,7673	1,73	0,9582	2,73	0,9968
0,24	0,5948	1,24	0,8925	2,24	0,9875	0,74	0,7703	1,74	0,9591	2,74	0,9969
0,25	0,5987	1,25	0,8944	2,25	0,9878	0,75	0,7734	1,75	0,9599	2,75	0,9970
0,26	0,6026	1,26	0,8962	2,26	0,9881	0,76	0,7764	1,76	0,9608	2,76	0,9971
0,27	0,6064	1,27	0,8980	2,27	0,9883	0,77	0,7794	1,77	0,9616	2,77	0,9972
0,28	0,6103	1,28	0,8997	2,28	0,9887	0,78	0,7823	1,78	0,9625	2,78	0,9973
0,29	0,6141	1,29	0,9015	2,29	0,9890	0,79	0,7852	1,79	0,9633	2,79	0,9974
0,30	0,6179	1,30	0,9032	2,30	0,9893	0,80	0,7881	1,80	0,9641	2,80	0,9974
0,31	0,6217	1,31	0,9049	2,31	0,9896	0,81	0,7910	1,81	0,9649	2,81	0,9975
0,32	0,6255	1,32	0,9066	2,32	0,9898	0,82	0,7939	1,82	0,9656	2,82	0,9976
0,33	0,6293	1,33	0,9082	2,33	0,9901	0,83	0,7967	1,83	0,9664	2,83	0,9977
0,34	0,6331	1,34	0,9099	2,34	0,9904	0,84	0,7995	1,84	0,9671	2,84	0,9977
0,35	0,6368	1,35	0,9115	2,35	0,9906	0,85	0,8023	1,85	0,9678	2,85	0,9978
0,36	0,6406	1,36	0,9131	2,36	0,9909	0,86	0,8051	1,86	0,9686	2,86	0,9979
0,37	0,6443	1,37	0,9147	2,37	0,9911	0,87	0,8078	1,87	0,9693	2,87	0,9979
0,38	0,6480	1,38	0,9162	2,38	0,9913	0,88	0,8106	1,88	0,9699	2,88	0,9980
0,39	0,6517	1,39	0,9177	2,39	0,9916	0,89	0,8133	1,89	0,9706	2,89	0,9981
0,40	0,6554	1,40	0,9192	2,40	0,9918	0,90	0,8159	1,90	0,9713	2,90	0,9981
0,41	0,6591	1,41	0,9207	2,41	0,9920	0,91	0,8186	1,91	0,9179	2,91	0,9982
0,42	0,6628	1,42	0,9222	2,42	0,9922	0,92	0,8212	1,92	0,9726	2,92	0,9982
0,43	0,6664	1,43	0,9236	2,43	0,9925	0,93	0,8238	1,93	0,9732	2,93	0,9983
0,44	0,6700	1,44	0,9251	2,44	0,9927	0,94	0,8264	1,94	0,9738	2,94	0,9984
0,45	0,6736	1,45	0,9265	2,45	0,9929	0,95	0,8298	1,95	0,9744	2,95	0,9984
0,46	0,6772	1,46	0,9279	2,46	0,9931	0,96	0,8315	1,96	0,9750	2,96	0,9985
0,47	0,6808	1,47	0,9292	2,47	0,9932	0,97	0,8340	1,97	0,9756	2,97	0,9985
0,48	0,6844	1,48	0,9306	2,48	0,9934	0,98	0,8365	1,98	0,9761	2,98	0,9986
0,49	0,6879	1,49	0,9319	2,49	0,9936	0,99	0,8389	1,99	0,9767	2,99	0,9986

Grandes valeurs absolues de t

t	$\phi(t)$
$4,5$	$0,999997$
$4,0$	$0,999968$
$3,8$	$0,999928$
$3,6$	$0,999841$
$3,5$	$0,99976$
$3,4$	$0,99966$
$3,3$	$0,99952$
$3,2$	$0,99931$
$3,1$	$0,99904$
$3,0$	$0,99865$

Annexe 2 : Table de K. PEARSON ou table de χ^2

Valeurs de χ^2 ayant une probabilité p d'être dépassée

Remarque : ν est le nombre de degré de liberté.

	0,990	0,975	0,950	0,900	0,100	0,050	0,025	0,010
1	0,0002	0,001	0,004	0,016	2,71	3,84	5,02	6,63
2	0,02	0,05	0,10	0,21	4,61	5,99	7,38	9,21
3	0,12	0,22	0,35	0,58	6,25	7,81	9,35	11,34
4	0,30	0,48	0,71	1,06	7,78	9,49	11,14	13,28
5	0,55	0,83	1,15	1,61	9,24	11,07	12,83	15,09
6	0,87	1,24	1,64	2,20	10,64	12,59	14,45	16,81
7	1,24	1,69	2,17	2,83	12,02	14,07	16,01	18,47
8	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09
9	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67
10	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21
11	3,05	3,82	4,57	5,58	17,27	19,67	21,92	24,72
12	3,57	4,40	5,23	6,30	18,55	21,03	23,24	26,22
13	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69
14	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14
15	5,23	6,26	7,26	5,55	22,31	25,00	27,49	30,58
16	5,81	6,91	7,96	9,31	23,54	26,30	28,84	32,00
17	6,41	7,56	8,67	10,08	24,77	27,59	30,19	33,41
18	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,80
19	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19
20	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57
21	8,90	10,28	11,59	13,24	29,61	32,67	35,48	38,93
22	9,54	10,98	12,34	14,04	30,81	33,92	36,78	40,29
23	10,20	11,69	13,09	14,85	32,01	35,17	38,08	41,64
24	10,86	12,40	13,85	15,66	33,20	36,41	39,37	42,98
25	11,52	13,12	14,61	16,47	34,38	37,65	40,65	44,31
26	12,20	13,84	15,38	17,29	35,36	38,88	41,92	45,64
27	12,88	14,57	16,15	18,11	36,74	40,11	43,19	46,96
28	13,57	15,31	16,93	18,94	37,92	41,34	44,46	48,28
29	14,26	16,05	17,71	19,77	39,09	42,56	45,72	49,59
30	14,95	16,79	18,49	20,60	40,26	43,77	46,98	50,89

Annexe 3 : Bibliographie

J. BASS : *Elément de calcul des probabilités*, MASSON.

M. GAULTIER : *Analyse, probabilités et méthodes statistiques*, Tome 1, VUIBERT.

J. PELLAUMAIL : *Probabilité, statistique, files d'attente*, DUNOD.

G. SAPORTA : *Probabilités, analyse des données et statistique*, TECHNIP, ISBN : 2-7108-0565-0.