MAT3071- Processus Stochastiques

Baccalauréats de Mathématiques, Statistique et d'actuariat -
Département de Mathématiques

Ces notes de cours ont été construite à partir du livre de Sheldon M. Ross Introduction to probability models, *tenth edition (Academic Press, 2010) et des notes de cours de Fabrice Larribe, Professeur à l'UQAM.*

Hélène Guérin
IRMAR, Université Rennes 1
Campus de Beaulieu, 35042 Rennes Cedex FRANCE.
helene.guerin@univ-rennes1.fr

Hiver 2012.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Table des matières

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Chapitre 1

Boîte à outils : Quelques notions de probabilités

1.1 Loi jointe, indépendance

Rappel : La formule des probabilités conditionnelles (ou formule de Bayes) :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B).$$

Exercice 1.1. *Cas discret*
On jette simultanément deux dés. On note X le nombre de chiffres pairs apparus et Y le maximum des deux chiffres obtenus.

1. A votre avis ces deux variables sont-elles indépendantes ?
2. Chercher la loi du couple (X, Y) ?
3. Vérifier votre opinion sur l'indépendance de X et Y en utilisant la question précédente.

Solution succinte :

1. *Non, car il existe un lien entre ces deux variables. Par exemple, si Y est pair on a forcément $X \geqslant 1$.*
2.

$Y \backslash X$	0	1	2	*Loi de Y*
1	1/36	0	0	1/36
2	0	1/18	1/36	1/12
3	1/12	1/18	0	5/36
4	0	1/9	1/12	7/36
5	5/36	1/9	0	1/4
6	0	1/6	5/36	11/36
Loi de X	1/4	1/2	1/4	

3. *On a effectivement $\mathbb{P}(X = 0, Y = 1) \neq \mathbb{P}(X = 0) \times \mathbb{P}(Y = 1)$.*

△

Exercice 1.2.
On effectue une suite infinie de lancers indépendants d'un dé équilibré. On numérote les lancers à partir de 1. On définit les deux variables aléatoires :

X égale au numéro du lancer qui donne le premier 6,
Y égale au nombre de 5 obtenus avant le premier 6.

Déterminer la loi du couple (X, Y).

Solution succinte : On sait que X suit une loi géométrique (loi du premier succès). X est à valeurs dans $\mathbb{N}^*$ et Y dans $\mathbb{N}$. Deplus on a forcément $Y \geqslant X$. Soit $(k,n) \in \mathbb{N}^* \times \mathbb{N}$, alors si $n \geqslant k$, $\mathbb{P}(X = k, Y = n) = 0$. Si $n < k$,

$$\mathbb{P}(X = k, Y = n) = \begin{pmatrix} k-1 \\ n \end{pmatrix} (1/6)^{n+1}(4/6)^{k-1-n}.$$

△

Exercice 1.3. *Cas continu*
Soit (X, Y) un couple de variables aléatoires de densité

$$f_{X,Y}(x,y) = \begin{cases} c\,\frac{\exp(-x)}{y^2} & \text{si } x > 0, \text{ et } y > 1 \\ 0 & \text{sinon.} \end{cases}$$

Calculer la constante c. Les variables aléatoires X et Y sont-elles indépendantes ?

Solution succinte : Les variables sont indépendantes car $f_{(X,Y)}(x,y) = ce^{-x}\mathbb{1}_{x>0} \times 1/y^2\mathbb{1}_{y>1}$ peut s'écrire comme le produit d'une fonction qui ne dépend que de x et d'une fonction qui ne dépend que de y. En calculant l'intégrale de la densité jointe sur $\mathbb{R} \times \mathbb{R}$, on obtient $c = 1$. △

Exercice 1.4. *Les amoureux des bancs publiques*
Deux personnes se donnent rendez-vous. L'heure d'arrivée de chacune de ces deux personnes sur les lieux est une variable uniforme entre midi et une heure. Ces deux variables sont indépendantes. Quelle est la probabilité qu'ils arrivent au même instant ? Quelle est la probabilité que le premier arrivé doive attendre plus de 10 minutes ?
Si les deux personnes se donnent un rendez-vous plus précis, à midi exactement par exemple. La loi uniforme est-elle bien adaptée au problème ?

Solution succinte : On note X et Y les instants d'arrivé de chacune des personnes. On peut modéliser les lois de X et Y par la loi uniforme entre midi et une heure, qui a pour densité $\mathbb{1}_{[0,1]}$. Vu que l'on travaille avec des variables continus $\mathbb{P}(X = Y) = 0$. Par ailleurs, on veut évaluer $\mathbb{P}(|X - Y| > 1/6) = \mathbb{P}(\max(X,Y) > \min(X,Y) + 1/6)$. On trouve $5/6$.
La loi uniforme n'est plus adaptée dans le cas où le rendez-vous est plus précis, on pourrait alors utiliser une loi exponentielle ou une loi de Laplace (de densité $1/2e^{-|x|}$). △

1.2 Fonction génératrice des moments

Définition 1.5. *Soit X une variable aléatoire. On appelle* **fonction génératrice des moments** *(ou tranformée de Laplace) la fonction*

$$\phi X(t) = \mathbb{E}\big[e^{tX}\big] \in [0, +\infty].$$

Cette fonction caractérise la loi de la variable aléatoire X, par conséquent connaitre la fonction génératrice suffit à connaître la loi d'une variable aléatoire.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Propriété 1.6. *1.* $\phi(0) = 1$

2. On retrouve facilement les moments d'une variable X à partir de sa fonction génératrice des moments :

$$\mathbb{E}[X] = \phi'(0),\ \mathbb{E}\big[X^2\big] = \phi''(0),\ \mathbb{E}[X^n] = \phi^{(n)}(0) \text{ pour } n \geqslant 0.$$

3. Si X et Y sont indépendantes alors pour tout $t \in \mathbb{R}$

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t).$$

Exercice 1.7. *Loi de Poisson*

1. Calculer la fonction génératrice d'une loi de Poisson de paramètre λ.
2. Soient X et Y deux variables indépendantes, de loi de Poisson de paramètre respectif λ et μ. Montrer que $X + Y$ suit la loi de Poisson de paramètre $\lambda + \mu$ en utilisant la fonction génératrice.
3. Déterminer la loi conditionnelle de X sachant $\{X + Y = n\}$.

Solution succinte :

1. On obtient $\phi_X(t) = e^{\lambda(e^t-1)}$.

2. Par indépendance on a $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = e^{(\lambda+\mu)(e^t-1)}$. D'où le résultat.

3. Si $k > n$, on a $\mathbb{P}(X = k|X + Y = n) = 0$. Si $k \leqslant n$, par la formule de Bayes, on a

$$\mathbb{P}(X = k|X + Y = n) = \frac{\mathbb{P}(X = k, Y = n-k)}{\mathbb{P}(X + Y = n)} = \begin{pmatrix} n \\ k \end{pmatrix} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k}.$$

On reconnaît la loi Binomiale $\mathcal{B}(n, \frac{\lambda}{\lambda+\mu})$.

△

1.3 Théorèmes limites et inégalités bien utiles

1.3.1 Quelques notions sur les variables gaussiennes

Définition 1.8. *Une variable aléatoire* **gaussienne** *(ou* **normale***) X d'espérance m et de variance σ^2 a pour densité*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad \text{avec } x \in \mathbb{R}.$$

On dit que X suit la loi $\mathcal{N}(m, \sigma^2)$.

Proposition 1.9. *1. La transformée de Laplace d'une variable aléatoire X de loi $\mathcal{N}(m, \sigma^2)$ est :*

$$\varphi_X(t) = \mathbb{E}[e^{tX}] = e^{mt}e^{\frac{\sigma^2t^2}{2}} \quad \text{pour } t \in \mathbb{R}.$$

2. Si $a, b \in \mathbb{R}$, X une variable gaussienne $\mathcal{N}(m, \sigma^2)$, alors $Y = aX + b$ est une variable gaussienne $\mathcal{N}(am + b, a^2\sigma^2)$.

3. Soient $X_1, \ldots, X_n$ des variables gaussiennes réelles indépendantes d'espérance m_i et de variance σ_i^2 respectives. Soient $a_1, \ldots, a_n$ des réels, alors $Y = a_1X_1 + \ldots + a_nX_n$ est une variable gaussienne $\mathcal{N}(\sum_{i=1}^n a_im_i, \sum_{i=1}^n a_i^2\sigma_i^2)$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Démonstration. 1. Calculons la transformée de Laplace d'une variable gaussienne :

$$\begin{aligned}
\varphi_X(t) &= \mathbb{E}[e^{tX}] = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} e^{tx} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-m)^2-2\sigma^2 tx}{2\sigma^2}} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-m-\sigma^2 t)^2-\sigma^4t^2-2\sigma^2t^2m}{2\sigma^2}} dx \\
&= e^{\frac{\sigma^4t^2+2\sigma^2mt}{2\sigma^2}} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-m-\sigma^2t)^2}{2\sigma^2}} dx = e^{mt} e^{\frac{\sigma^2t^2}{2}}.
\end{aligned}$$

On reconnaît la densité de $\mathcal{N}(m+\sigma^2 t, \sigma^2)$ dans l'intégrale, on a donc $\frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{(x-m-\sigma^2t)^2}{2\sigma^2}} dx = 1$.

2. On calcule la transformée de Laplace de Y :

$$\begin{aligned}
\varphi_Y(t) &= \mathbb{E}[e^{atX+bt}] = e^{bt}\mathbb{E}[e^{(at)X}] = e^{bt}e^{mat}e^{\sigma^2a^2t^2/2} \\
&= e^{(am+b)t}e^{(a\sigma)^2t^2/2}
\end{aligned}$$

On reconnaît la transformée de Laplace de la loi gaussienne $\mathcal{N}(am+b, a^2\sigma^2)$, et comme la transformée de Laplace caractérise la loi, on a le résultat.

3. On utilise encore la transformée de Laplace. Comme les variables X_i sont indépendantes, la transformée de Laplace de la somme est le produit des transformées de Laplace :

$$\varphi_Y(t) = \mathbb{E}[e^{ta_1X_1+\ldots+ta_nX_n}] = \prod_{i=1}^n \mathbb{E}[e^{ta_iX_i}] = \prod_{i=1}^n e^{\sum_{i=1}^n a_im_it} e^{\sum_{i=1}^n a_i^2\sigma_i^2t^2/2}$$

La variable Y suit par conséquent la loi gaussienne $\mathcal{N}(\sum_{i=1}^n a_im_i, \sum_{i=1}^n a_i^2\sigma_i^2)$.

$\square$

Remarque 1.10. *Soit X une variable de loi $\mathcal{N}(0,1)$, alors les valeurs de X sont "concentrées" entre -4 et 4. En effet, la probabilité $\mathbb{P}(|X|>4) = 5.10^{-5}$ est très faible.*

1.3.2 Inégalités importantes

L'inégalité de Markov

Soit X une variable positive et un réel $a > 0$, alors

$$\mathbb{P}(X > a) \leqslant \frac{\mathbb{E}[X]}{a}.$$

Démonstration. La variable X étant positive, on a

$$\mathbb{E}[X] = \mathbb{E}[X\mathbb{I}_{X>a}] + \mathbb{E}[X\mathbb{I}_{X\leqslant a}] \geqslant \mathbb{E}[X\mathbb{I}_{X>a}] \geqslant a\mathbb{P}(X>a).$$

$\square$

L'inégalité de Bienaymé-Chebichev

Soit X une variable avec $\mathbb{E}[X^2] < \infty$ et un réel $a > 0$, alors

$$P\left(|X - \mathbb{E}[X]| > a\right) \leqslant \frac{Var(X)}{a^2}.$$

Démonstration. On applique l'inégalité de Markov à la variable $|X-\mathbb{E}[X]|^2$ et on obtient le résultat car $\mathbb{P}(|X-\mathbb{E}[X]| > a) = \mathbb{P}(|X-\mathbb{E}[X]|^2 > a^2)$. $\square$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

1.3.3 Différentes notions de convergence

On considère une suite de variables $(X_n)_{n\geqslant 1}$ et la variable X.

• Convergence en probabilité
La suite $(X_n)_{n\geqslant 1}$ converge en probabilité vers X ssi

$$\forall \varepsilon > 0 \quad \lim_{n\to+\infty} P\left(|X_n - X| > \varepsilon\right) = 0.$$

On note alors $X_n \underset{n\to+\infty}{\overset{proba}{\longrightarrow}} X$.
La convergence en probabilité signifie que la probabilité que X_n soit loin de X (à distance plus grande que ε) converge vers 0 quand n tend vers $+\infty$.

• Convergence presque-sûre
La suite $(X_n)_{n\geqslant 1}$ converge presque sûrement vers X ssi

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

"Presque tout le temps X_n converge vers X", les valeurs pour lesquelles il n'y a pas convergence sont négligeables (i.e., appartiennent à un ensemble de probabilité nulle).
On note alors $X_n \underset{n\to+\infty}{\overset{p.s.}{\longrightarrow}} X$.

• Convergence en loi
La suite $(X_n)_{n\geqslant 1}$ converge en loi vers X ssi

$$\forall I \text{ intervalle} \quad \mathbb{P}(X_n \in I) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(X \in I).$$

On note alors $X_n \underset{n\to+\infty}{\overset{loi}{\longrightarrow}} X$.
La convergence en loi signifie que la loi de X_n converge vers la loi de X quand n tend vers $+\infty$. Par conséquent l'ensemble des valeurs de X_n converge vers l'ensemble des valeurs de X.

Proposition 1.11. *On a $X_n \underset{n\to+\infty}{\overset{loi}{\longrightarrow}} X$*
s'il y a convergence des fonctions de répartition : $F_{X_n}(x) \underset{n\to+\infty}{\longrightarrow} F_X(x) \quad \forall x$ point de continuité de F_X
ou si, dans le cas discret, $\forall k$, $\mathbb{P}(X_n = k) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(X = k)$
ou s'il y a convergence des fonctions génératrices des moments : $\phi_{X_n}(t) \underset{n\to+\infty}{\longrightarrow} \phi_X(t)$ pour tout t

• Lien entre les différentes notions de convergence

$$\text{"convergence p.s."} \Rightarrow \text{"convergence en proba"} \Rightarrow \text{"convergence en loi"}$$

Attention les réciproques sont fausses ! Cependant, on a le résultat suivant :

Remarque 1.12. *Si $X_n \underset{n\to+\infty}{\overset{loi}{\longrightarrow}} a$ et a est constante, alors $X_n \underset{n\to+\infty}{\overset{proba}{\longrightarrow}} a$*

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

1.3.4 La loi des grands nombres

Considérons une suite $(X_n)_{n\geqslant 0}$ de variables aléatoires. On s'intéresse à la variable aléatoire $\overline{X}_n = \frac{X_1+X_2+...+X_n}{n}$ appelée **moyenne empirique** et on souhaite connaître son comportement quand n grandit. On a le théorème suivant :

Théorème 1.13. (Loi faible des grands nombres)
Soit X une varialbe aléatoire avec $\mathbb{E}[|X|] < \infty$.
Soient $(X_n)_{n\geqslant 0}$ une suite de variables indépendantes et de même loi que X. Alors

$$\overline{X}_n = \frac{X_1+X_2+...+X_n}{n} \underset{n\to+\infty}{\longrightarrow} \mathbb{E}[X] \quad \textit{en proba.}$$

Démonstration. Preuve dans le cas où $\mathbb{E}[X^2] < \infty$. On suppose $\mathbb{E}[X] = 0$ (sinon il suffit de travailler avec la suite $(X_n - \mathbb{E}[X])_{n\geqslant 0}$). Soit $\varepsilon > 0$. On a

$$P\left(\frac{|X_1+\ldots+X_n|}{n} \geqslant \varepsilon\right) \overset{\text{Bienaymé -Chebichev}}{\leqslant} \frac{Var(X_1+\ldots+X_n)}{n^2\varepsilon^2} \overset{\text{Indépendance}}{\leqslant} \frac{Var(X)}{n\varepsilon^2}.$$

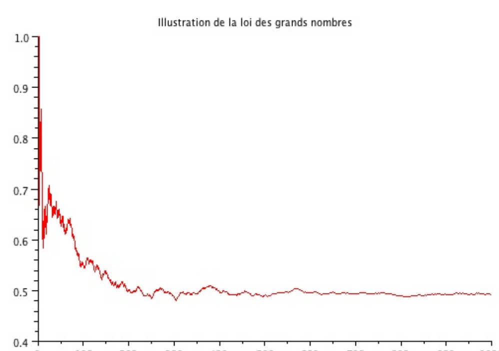
D'où le résultat. □

Théorème 1.14. (Loi forte des grands nombres)
Soit X une v.a.avec $\mathbb{E}[|X|] < \infty$.
Soient $(X_n)_{n\geqslant 0}$ une suite de variables indépendantes et de même loi que X. Alors

$$\overline{X}_n = \frac{X_1+X_2+...+X_n}{n} \underset{n\to+\infty}{\longrightarrow} \mathbb{E}[X] \quad p.s..$$

La différence entre les deux théorèmes est très subtile, il faut bien évidemment retenir la loi forte des grands nombres. L'intérêt de la loi faible est que sa démonstration est simple, alors que celle de la loi forte est très compliquée.

Exemple 1.15. *On peut retrouver le résultat par simulation. On considère une pièce bien équilibrée et on joue à* "pile ou face"*. X_1 est le résultat du premier lancer, X_2 celui du second, et ainsi de suite. On donne la valeur 1 à X_i si le résultat du $i^{ème}$ lancer est* pile *et 0 si c'est* face*. $S_n = X_1 + X_2 + \ldots + X_n$ est alors le nombre de* pile *obtenu après n lancers.*
On sait que X_i suit le loi de Bernoulli $\mathcal{B}(1/2)$ et S_n la loi binomiale $\mathcal{B}(n, 1/2)$. Le graphe ci-dessus représente l'évolution de la moyenne empirique $\overline{X}_n = \frac{X_1+X_2+...+X_n}{n}$ en fonction de n. On observe bien la convergence vers $1/2$.



Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Une application importante de la loi des grands nombres est en simulation. Par exemple on considère une variable aléatoire, une fonction f et on aimerait connaître la valeur de $\mathbb{E}[f(X)]$. Le calcul de cette valeur peut-être très compliqué. On peut cependant en obtenir une valeur approximative par simulation : On simule des variables $X_1, \ldots, X_n$ de même loi que X, on prend n assez grand et d'après la loi forte des grands nombres, on sait que $\frac{1}{n}\sum_{i=1}^{n} f(X_i)$, qui facilement calculable par ordinateur, est proche de $\mathbb{E}[f(X)]$.

1.3.5 Le théorème limite central

On se demande maintenant à quelle vitesse la vitesse de convergence dans la loi des grands nombres (à quelle vitesse la moyenne empirique converge vers l'espérance ?).

Théorème 1.16. (Théorème limite central)
Soit X une v.a. avec $\mathbb{E}[X^2] < \infty$. On pose $m = \mathbb{E}[X]$ et $\sigma^2 = Var(X)$.
Soient X_1, X_2, ..., X_n des variables indépendantes et de même loi que X. Alors

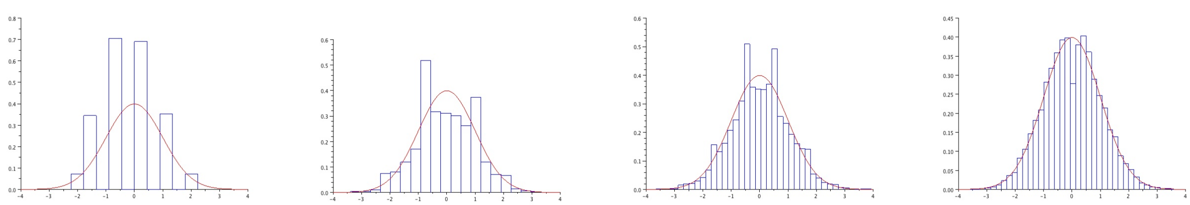
$$\frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + X_2 + ... + X_n}{n} - m\right) \underset{n\to+\infty}{\longrightarrow} Z \quad \text{en loi, avec } \mathcal{L}(Z) = \mathcal{N}(0,1).$$

Donc $\forall a, b \in \mathbb{R}$, $a < b$,

$$\mathbb{P}(a \leqslant \frac{\sqrt{n}}{\sigma}\left(\frac{X_1 + X_2 + ... + X_n}{n} - m\right) \leqslant b) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(a \leqslant Z \leqslant b) = \int_a^b e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

La vitesse de convergence de la moyenne empirique est assez lente car elle est en $\frac{\sqrt{n}}{\sigma}$ (bien sûr, plus σ est petit plus la vitesse est rapide).

Exemple 1.17. *Reprenons l'exemple précédent. On simule des variables $X_1, \ldots, X_n$ suivant la loi $\mathcal{B}(1/2)$. Les graphes ci-dessous représentent en bleu la loi empirique (i.e. l'histogramme associé à la moyenne empirique) de la variable $2\sqrt{n}\left(\frac{X_1+X_2+...+X_n}{n} - 1/2\right)$ pour différentes valeurs de n : 5, 50, 500, 1000. On observe bien la convergence vers la densité de la loi normale $\mathcal{N}(0,1)$ tracée en rouge.*



Lorsqu'on utilise le théorème limite central, on a souvent besoin besoin de connaitre les valeurs de la fonction de répartition de la loi normale. Elle n'est malheureusement pas calculable à la main, mais n'importe quel logiciel (Matlab, Scilab, R, ...) peut donner ses valeurs. On peut aussi trouver des tables de la loi normale dans les livres ou sur internet.

Exercice 1.18. *À propos du théorème central limite*
On sait déjà par la loi des grands nombres que la moyenne du nombre de pile obtenu en lançant une pièce équilibrée converge presque sûrement vers 1/2. On se demande maintenant combien de fois faut-il lancer la pièce pour que cette moyenne soit dans l'intervalle $[1/2 - 0.01, 1/2 + 0.01]$ avec une probabilité supérieure ou égale à 0.95 ?
On déterminera un tel nombre de deux façons :

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

1. en utilisant l'inégalité de Bienaymé-Chebichev ;
2. par approximation par la loi normale.

(Soit Z une variable gaussienne $\mathcal{N}(0,1)$, on a $\mathbb{P}(|Z| \leqslant 1.96) = 0.95$.)

Solution succinte : On pose X_i le résultat du $i^{\text{ème}}$ lancer. On note $X_i = 1$ si le $i^{\text{ème}}$ lancer est pile et $X_i = 1$ si le $i^{\text{ème}}$ lancer est face. Les X_i suivent la loi de Bernoulli $\mathcal{B}(1/2)$.
Le nombre de piles obtenus sur n lancers $\sum_{i=1}^n X_i$ et $\overline{X}_n$ est la moyenne des piles obtenus. On cherche par conséquent n tel que $\mathbb{P}(\overline{X}_n \in [1/2-0.01, 1/2+0.01]) \geqslant 0.95$.

1. *On remarque que $\mathbb{E}[\overline{X}_n] = \mathbb{E}[X_1] = 1/2$ et du fait de l'indépendance entre les X_i on a $var(\overline{X}_n) = var(X_1)/n = \frac{1}{4n}$. On a*

$$\mathbb{P}(\overline{X}_n \in [1/2-0.01, 1/2+0.01]) = \mathbb{P}(|\overline{X}_n - 1/2| \leqslant 0.01) = 1 - \mathbb{P}(|\overline{X}_n - 1/2| > 0.01)$$

En utilisant l'inégalité de Bienaymé-Chebichev, on a

$$\mathbb{P}(\overline{X}_n \in [1/2-0.01, 1/2+0.01]) \geqslant 1 - \frac{var(\overline{X}_n)}{(0.01)^2}.$$

Il faut donc trouver n tel que $1 - \frac{1}{4n(0.01)^2} > 0.95$.

2. *On veut utiliser le théorème limite central. Dans notre cas, $m = 1/2$ et $\sigma^2 = 1/4$. Pour n assez grand,*

$$\mathbb{P}(\overline{X}_n \in [1/2-0.01, 1/2+0.01]) = \mathbb{P}(2\sqrt{n}|\overline{X}_n - 1/2| \leqslant 0.01 \times 2\sqrt{n}) \simeq \mathbb{P}(|Z| \geqslant 0.01 \times 2\sqrt{n}),$$

où $Z \sim \mathcal{N}(0,1)$. En utilisant le résultat de l'énoncé, il suffit donc de choisir n tel que

$$0.01 \times 2\sqrt{n} \geqslant 1.96,$$

soit $n \geqslant \left(\frac{1.96}{0.02}\right)^2$.

△

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

1.4 Quelques Lois usuelles

1.4.1 Lois de probabilité discrètes

- *Loi de Bernoulli*, $\mathcal{B}(p)$, avec $p \in]0,1[$:

$P(X=1)=p \quad P(X=0)=1-p.$

- *Loi Binomiale*, $\mathcal{B}(n,p)$, avec $p \in]0,1[$, $n \in \mathbb{N}^*$: *(Loi du nombre de succès sur n essais indépendants)*

$P(X=k)=\binom{n}{k} p^k (1-p)^{n-k}$ avec $k \in \{0,1,..,n\}$.

- *Loi Géométrique*, $\mathcal{G}(p)$, avec $p \in]0,1[$: *(Loi du premier succès)*

$P(X=k)=p(1-p)^{k-1}$ avec $k \in \mathbb{N}^*$.

- *Loi de Poisson*, $\mathcal{P}(\lambda)$, avec $\lambda > 0$:

$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$ avec $k \in \mathbb{N}$.

Loi	**Notation**	**Espérance**	**Variance**	**Fonction génératrice des moments**
Bernoulli	$\mathcal{B}(p)$	p	$p(1-p)$	$1-p+pe^t$
Binomiale	$\mathcal{B}(n,p)$	np	$np(1-p)$	$(1-p+pe^t)^n$
Géométrique	$\mathcal{G}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\dfrac{pe^t}{1-(1-p)e^t}$
Poisson	$\mathcal{P}(\lambda)$	λ	λ	$e^{\lambda(e^t-1)}$

1.4.2 Lois de probabilité à densité

Loi	**Notation**	**Densité**	**Espérance**	**Variance**	**Fonct. géné. des moments**
Uniforme	$\mathcal{U}([a,b])$	$\dfrac{1}{b-a}$ pour $x \in [a,b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\dfrac{e^{tb}-e^{ta}}{(b-a)t}$
Exponentielle	$\mathcal{E}(\lambda)$, $\lambda > 0$	$\lambda e^{-\lambda x}$ pour $x \geqslant 0$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$	$\dfrac{\lambda}{\lambda-t}$, $t<\lambda$
Normale	$\mathcal{N}(m,\sigma^2)$, $\sigma^2>0$	$\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\dfrac{(x-m)^2}{2\sigma^2}}$	m	σ^2	$e^{mt+\frac{\sigma^2t^2}{2}}$
Gamma	$G(a,\lambda)$, $a,\lambda>0$	$\frac{\lambda^a}{\Gamma(a)}x^{a-1}e^{-\lambda x}$, $x>0$	$\dfrac{a}{\lambda}$	$\dfrac{a}{\lambda^2}$	$\left(\frac{\lambda}{\lambda-t}\right)^a$, $t<\lambda$
Cauchy	$\mathcal{C}$	$\dfrac{1}{\pi}\dfrac{1}{1+x^2}$	pas d'espérance	pas de variance	

Remarque : La somme de n variables exponentielles indépendantes suit une loi Gamma : si X_i sont des

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

variables indépendantes de loi $\mathcal{E}(\lambda)$, alors $\sum_{i=1}^{n} X_i$ suit la loi $G\,(n, \lambda)$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Chapitre 2

Chaînes de Markov

2.1 Introduction

Que signifie le terme processus ? Un **processus stochastique** est une famille $(X_t)_{t \in T}$ de variables aléatoires X_t. L'indice t peut représenter par exemple le temps ou une coordonnée le long d'un chemin.

Exemple 2.1. *X_t peut par exemple être*
- *la position d'un joueur au monopoly ou à Serpent-Échelle ;*
- *le nombre d'accidents de voiture à Montréal à l'instant t ;*
- *le nombre véhicule sur une autoroute au kilomètre t ;*
- *le prix d'une action à l'instant t ;*
- *le nombre d'étudiants à l'UQAM à l'instant t.*

L'ensemble des indices T du processus peut-être fini, infini-dénombrable (i.e., je peux compter chaque élément) comme $\mathbb{N}$ ou $\mathbb{Z}$ ou non dénombrable comme $\mathbb{R}^+$ ou $\mathbb{R}$. Lorsque T est fini ou dénombrable, le processus est dit à **temps discret**, alors que si T est un intervalle de $\mathbb{R}$ le processus est dit à **temps continu**.

Exemple 2.2.
- *la position d'un joueur à Serpent-Échelle est prise à des temps discrets : $T = \mathbb{N}$;*
- *le nombre d'accidents de voiture à Montréal est considéré à tout instant positif : $T = [0, +\infty[$;*
- *le nombre véhicule sur une autoroute au kilomètre t peut-être regardé juste entre le kilomètre 500 et le kilomètre 800 de l'autoroute : $T = [500, 800]$.*

Les processus les plus simples sont les suites de variables aléatoires indépendantes (par exemple, lorsque qu'on regarde les résultats successifs d'un lancer de dé). Il existe de nombreux résultats sur de telles suites (comme la loi des grands nombres ou le théorème limite central), mais dans la réalité beaucoup de choses n'évolue pas de façon indépendante. On va étudier ici un type de processus qui possède une dépendance (*contrôlée*) par rapport à son passé : les chaînes de Markov.

Il s'agit d'une famille de variables indexées par $\mathbb{N}$ et qui possède la propriété d'évoluer en fonction de son dernier emplacement et non de tout son passé, un peu comme quelqu'un de ivre qui marche dans la rue. On retrouve ce type de processus dans de nombreuses situations, notamment dans les jeux de société comme le Monopoly ou Serpent-Échelle ou en assurance dans le système de Bonus-Malus.
Regardons le jeu de Serpent-Échelle.

On joue avec un dé équilibré. S'il n'y avait ni serpent ni échelle, partant de la case i, on aurait la même probabilité, égale à $1/6$, d'atterrir au coup d'après sur l'une des 6 cases suivantes, mais du fait de la présence de serpents et d'échelles le jeu est un peu plus compliqué. Connaissant notre position actuelle, on peut déterminer la probabilité de tomber sur les cases suivantes. On note X_n la position de notre pion après le $n^{\text{ème}}$ lancer.
Fixons n. La position X_{n+1} ne dépend que de la position précédente X_n et du résultat du $n+1^{\text{ème}}$ lancer du dé. Savoir comment on est arrivé à la position X_n ne nous donne aucune information supplémentaire pour connaitre la valeur de X_{n+1}. Il s'agit d'une chaîne de Markov. On peut alors se demander quel est le temps moyen d'une partie de jeu, ou même sachant qu'on est déjà arrivé à la case i combien de temps en moyenne faudra-t-il attendre pour finir la partie ?

2.1.1 Définitions

Considérons un processus stochastique $(X_n)_{n\geqslant 0}$ indexé par $\mathbb{N}$ à valeurs dans un espace noté S. S est appelé **espace d'état**. L'espace d'état S est dénombrable, il peut être fini ou infini.
Si $X_n = i$, on dit que le processus est à l'état i à l'instant n.
L'état à l'instant 0 est l'**état initial** du processus. Il est en général aléatoire et on appelle sa loi la **loi initiale** du processus.

Définition 2.3. *Le processus $(X_n)_{n\geqslant 0}$ est une* **chaîne de Markov** *si pour tous $i, j, i_0, \dots, i_{n-1} \in S$, $n \geqslant 0$, les probabilités suivantes ne dépendent que de i, j et n*

$$\mathbb{P}(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_1 = i_1, X_0 = i_0) = \mathbb{P}(X_{n+1} = j | X_n = i).$$

On note $P_{i,j}^{n,n+1}$ cette quantité que l'on appellera **probabilité de transition** *de l'état i à l'état j entre les instants n et $n+1$.*
Si les probabilités de transition $P_{i,j}^{n,n+1}$ ne dépendent pas de n, on dit que la chaîne de Markov est **homogène**.

On ne regardera ici que des chaînes homogènes, on notera alors $P_{i,j}$ les probabilités de transition. Les calculs se compliquent beaucoup dans le cas non homogène.

Remarque 2.4. *La probabilité de transition $P_{i,j}$ représente la probabilité de passer de l'état i à l'état j en un coup. On a pour tout $i, j \in S$,*

$$0 \leqslant P_{i,j} \leqslant 1 \qquad \sum_{j \in S} P_{i,j} = 1. \tag{2.1}$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

(En effet, $\sum_{j\in S} P_{i,j} = \mathbb{P}(X_j \in S|X_n = i)$*.)*

On note en général les probabilités de transition sous forme de matrice $P = (P_{i,j})_{i,j\in S}$*, qui est appelé* **matrice de transition***. La matrice est de taille infinie si l'espace d'état* S *est infini.*

$$P = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,j} & \cdots \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,j} & \cdots \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,j} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ P_{i,0} & P_{i,1} & P_{i,2} & \cdots & P_{i,j} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots \end{bmatrix}.$$

Au croisement de la $i^{\text{ème}}$ *ligne et* $j^{\text{ème}}$ *colonne de la matrice* P *se trouve la probabilité* $P_{i,j}$ *de passer de l'état* i *à l'état* j *en un coup.*

Une matrice qui vérifie les relations (2.1) *est appelée* **matrice stochastique***.*

On peut aussi utiliser des graphes pour représenter les probabilités de transition.

En conclusion, une chaîne de Markov est un processus dont le futur ne dépend de son passé qu'à travers sa position à l'instant présent, on peut aussi l'énoncé ainsi "une chaîne de Markov est un processus où la loi conditionnelle de l'état futur X_{n+1} sachant tous les états passés $X_0, X_1, \ldots, X_{n-1}$ et l'état présent X_n est indépendante du passé et ne dépend que de l'état présent".
Autrement dit, "le passé et le futur sont indépendants conditionnellement au présent". Le processus est dit "à faible mémoire".

Exemple *Prévisions météorologiques* (Ross, p.192)
On suppose que le climat du lendemain ne dépend que du climat d'aujourd'hui et pas des jours précédents. On suppose que s'il pleut aujourd'hui il y a une probabilité p qu'il pleuve demain, par contre s'il ne pleut pas aujourd'hui il y a une probabilité q qu'il pleuve le lendemain.

L'ensemble des états est $\begin{cases} 0 & \text{il pleut}\,; \\ 1 & \text{il ne pleut pas.} \end{cases}$

Vu les hypothèses le processus $(X_n)_{n\geqslant 0}$, où X_n représente le temps le jour n, est une chaîne de Markov de matrice de transition

$$P = \begin{bmatrix} p & 1-p \\ q & 1-q \end{bmatrix}.$$

Proposition 2.5. *Soit* $(X_n)_{n\geqslant 0}$ *une chaîne de Markov sur un espace d'état* S*, de probabilités de transition* $(P_{i,j})_{i,j\in S}$ *et de loi initiale* $p_i = \mathbb{P}(X_0 = i)$*. Alors pour tout* $n \geqslant 0$ *et* $i_0, i_1, \ldots, i_n \in S$*, on a*

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = p_{i_0} \prod_{k=0}^{n-1} P_{i_k, i_{k+1}}.$$

La probabilité de suivre le chemin $\{i_0, i_1, \ldots, i_n\}$ est simplement la probabilité de partir de i_0, puis de passer de i_0 à i_1, puis de i_1 à i_2, ainsi de suite jusqu'à i_n.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Démonstration.

$$
\begin{aligned}
&\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n)\\
&\quad \overset{\text{Bayes}}{=} \mathbb{P}(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1})\\
&\quad \overset{\text{Markov}}{=} \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1})\\
&\quad = P_{i_{n-1},i_n}\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1})\\
&\quad = P_{i_{n-1},i_n}P_{i_{n-2},i_{n-1}}\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_{n-2} = i_{n-2})\\
&\qquad\qquad \vdots\\
&\quad = P_{i_{n-1},i_n}P_{i_{n-2},i_{n-1}} \dots P_{i_1,i_2}\mathbb{P}(X_0 = i_0, X_1 = i_1)\\
&\quad = P_{i_{n-1},i_n}P_{i_{n-2},i_{n-1}} \dots P_{i_0,i_1}\mathbb{P}(X_0 = i_0).
\end{aligned}
$$

□

Exemple *Prévisions météorologiques*
Reprenons l'exemple précédent et on suppose que la loi initiale est $p_0 = 0.1$ et $p_1 = 0.9$. On prend $p = 0.7$ et $q = 0.4$.

On souhaite prendre 3 jours de vacances et on se demande qu'elle est la probabilité qu'il ne pleuve pas 3 jours de suite :

$$\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 1) = 0.9 \times 0.6^2 = 0.324.$$

Maintenant, sachant qu'il ne pleut pas aujourd'hui, qu'elle est la probabilité qu'il ne pleuve pas les 2 prochains jours :

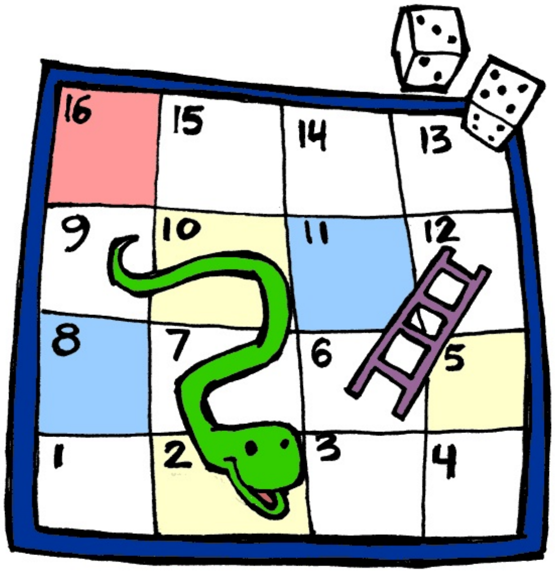
$$\mathbb{P}(X_1 = 1, X_2 = 1 | X_0 = 1) = \frac{\mathbb{P}(X_0 = 1, X_1 = 1, X_2 = 1)}{\mathbb{P}(X_0 = 1)} = 0.36.$$

2.1.2 Quelques exemples

Le jeu Serpent-Échelle

Dans le cas du jeux Serpent-Échelle, la chaîne est homogène. Considérons un jeu un peu plus simple avec seulement 16 cases :

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012



On suppose que la partie est terminée à partir du moment où on est tombé sur la dernière case ou si on l'a dépassée (contrairement au vrai jeu). Connaissant notre position actuelle, on peut déterminer la probabilité de tomber sur les cases suivantes. On note X_n la position de notre pion après le $n^{\text{ème}}$ lancer. L'espace d'état est $S = \{1, 2, 3, 4, 5, 7, 8, 10, 11, 12, 13, 14, 15, 16\}$.

$(X_n)_{n\geqslant 0}$ est bien une chaîne de Markov car la position X_{n+1} à l'instant n ne dépend que de la position du pion X_n à l'instant n et du résultat du $n+1^{\text{ème}}$ lancer de dé (les lancers de dés étant indépendants entre eux).

Calculons la matrice de transition :

$$P = \begin{pmatrix}
0 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 1/6 & 0 & 1/6 & 1/6 & 0 & 1/6 & 0 & 1/6 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 \\
0 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 1/3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 2/3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 5/6 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$

(J'ai fait le choix de prendre une matrice 16×16, même si l'espace d'état est de taille 14, afin que la ligne numéro i corresponde à la case numéro i du jeu. En fait, les colonnes 6 et 9 ne sont pas atteignables.)

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

La marche aléatoire sur $\mathbb{Z}$

On considère une personne qui a un peu trop bu et qui souhaite rentrer chez elle à pied (elle a au moins la présence d'esprit de ne pas prendre sa voiture !). On se place en dimension 1 pour simplifier le problème (elle se déplace sur une ligne).
La personne a tellement bu qu'à l'instant n elle ne se souvient même plus où elle était auparavant. Sa position à l'instant $n+1$ ne dépend par conséquent que de sa position à l'instant n et de la décision qu'elle va prendre : d'aller à droite ou à gauche. On peut par conséquent modéliser sa position à l'aide d'une chaîne de Markov. On suppose qu'elle décide d'aller à droite avec probabilité p et d'aller à gauche avec probabilité $1-p$.
La variable X_n représente donc la position de notre buveur à l'instant n. L'espace d'état est $S=\mathbb{Z}$. Les probabilités de transition sont :

$$P_{i,i-1}=1-p,\ P_{i,i+1}=p,\ \text{ et } P_{i,j}=0 \text{ pour } j\neq i-1,i+1.$$

Si $p=1/2$ la marche aléatoire est dite **symétrique** et si $p\neq 1/2$ la marche est dite **asymétrique** (il suffit que le contenu des poches de notre buveur soit plus lourd d'un coté que de l'autre pour l'influencer).
Le but de cette personne est de rentrer chez elle, on espère qu'elle y arrivera en temps fini.

Le système de Bonus-Malus en assurance automobile

Le calcul des primes dans les modèles européens ou asiatiques de l'assurance automobile est basé sur le nombre d'accident de chaque conducteur. On va décrire un modèle (il en existe beaucoup, voir Ross p.194 pour un modèle très semblable).
Un nouveau conducteur doit payer une certaine prime d'assurance la première année. S'il n'a pas eu d'accident pendant la première année, le conducteur bénéficie d'un bonus qui lui permet de payer une prime d'assurance réduite pour la seconde année, si lors de la seconde année il n'a toujours pas d'accident, sa prime est à nouveau baissée pour la troisième année et ainsi de suite jusqu'à un certain niveau. Par contre, s'il a au moins un accident pendant une année, alors le conducteur obtient un malus qui augmente sa prime d'assurance pour l'année suivante.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

On définit différents niveaux : le niveau le plus bas 0 correspond à la prime la plus basse et plus le niveau augmente plus la prime augmente.
Si un client est au niveau i, le nombre d'accident qu'il aura pendant l'année suit une loi de Poisson de paramètre λ_i (rappel : la moyenne d'une loi de Poisson $\mathcal{P}(\lambda)$ est λ).
Un conducteur passe du niveau i au niveau $i+j$ s'il a eu j accidents au cours de l'année. Par contre, s'il n'a eu aucun accident au cours de l'année, le conducteur passe du niveau i au niveau $i-1$. Bien évidemment, si le client était déjà au niveau 0 et qu'il n'a pas eu d'accident au cours de l'année, il reste au niveau 0.
Pour $i \geqslant 0$, $j \geqslant 0$, la probabilité qu'un conducteur au niveau i ait j accidents est

$$P_{\lambda_i}(j) = \frac{\lambda_i^j}{j!} e^{-\lambda_i}.$$

On note X_n le niveau de la prime du conducteur l'année n. $(X_n)_{n\geqslant 0}$ est une chaîne de Markov car X_{n+1} ne dépend que du niveau précédent X_n et du nombre d'accident que le conducteur aura au cours de l'année (on suppose que les années sont indépendantes entre elles). Son espace d'état est

$$S = \{0, 1, 2, \ldots\}.$$

La matrice de transition est la suivante

$$P = \begin{bmatrix} P_{\lambda_0}(0) & P_{\lambda_0}(1) & P_{\lambda_0}(2) & P_{\lambda_0}(3) & P_{\lambda_0}(4) & \cdots \\ P_{\lambda_1}(0) & 0 & P_{\lambda_1}(1) & P_{\lambda_1}(2) & P_{\lambda_1}(3) & \cdots \\ 0 & P_{\lambda_2}(0) & 0 & P_{\lambda_2}(1) & P_{\lambda_2}(2) & \cdots \\ 0 & 0 & P_{\lambda_3}(0) & 0 & P_{\lambda_3}(1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

De n'importe quel niveau on peut atteindre tous les niveaux strictement supérieurs, par contre si on est un bon conducteur on ne peut atteindre que le niveau juste en dessous.
Le but de la compagnie d'assurance est d'évaluer la prime dans chaque niveau afin d'être solvable.

2.2 Les équations de chapman-Kolmogorov

Connaissant les probabilités de transition, on sait passer de l'instant n à l'instant $n+1$:

$$\begin{aligned} \mathbb{P}(X_{n+1} = j) &= \mathbb{P}(X_{n+1} = j, X_n \in S) = \sum_{i\in S} \mathbb{P}(X_{n+1} = j, X_n = i) \\ &= \sum_{i\in S} P_{i,j} \mathbb{P}(X_n = i) \quad \text{(par la formule de Bayes).} \end{aligned}$$

Les équations de Chapman-Kolmogorov vont nous fournir un moyen de passer de la loi à l'instant n à la loi du processus à l'instant $n+m$. Notons $P_{i,j}^n$ la probabilité de l'état i à l'état j en n étapes :

$$P_{i,j}^n = \mathbb{P}(X_{n+k} = j | X_k = i) \quad \text{pour } n \geqslant 0,\ i, j \in S.$$

Dans le cas homogène, ces probabilités ne dépendent pas de k.

Attention, $P_{i,j}^n$ n'est pas le produit n fois de $P_{i,j}$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Proposition 2.6. *On a pour tout $n, m \geqslant 0$ et pour tout $i, j \in S$*

$$P_{i,j}^{n+m} = \sum_{k \in S} P_{i,k}^{n} P_{k,j}^{m}.$$

Ces équations sont appelées **équations de Chapman-Kolmogorov**.

La probabilité d'aller de i à j en $n+m$ étapes est donc la probabilité d'aller de i à k en n étapes et de k à j en m étapes en sommant sur toutes les valeurs possibles de k.

Démonstration.

$$\begin{aligned} P_{i,j}^{n+m} &= \mathbb{P}(X_{n+m} = j | X_0 = i) = \sum_{k \in S} \mathbb{P}(X_{n+m} = j, X_n = k | X_0 = i) \\ &= \sum_{k \in S} \mathbb{P}(X_{n+m} = j | X_n = k, X_0 = i) \mathbb{P}(X_n = k | X_0 = i) \qquad (2.2) \\ &\overset{\text{Propriété de Markov}}{=} \sum_{k \in S} \mathbb{P}(X_{n+m} = j | X_n = k) \mathbb{P}(X_n = k | X_0 = i) = \sum_{k \in S} P_{i,k}^{n} P_{k,j}^{m}. \end{aligned}$$

Pour l'égalité (2.2), on utilise la formule de Bayes, plus précisément on utilise :

$$\mathbb{P}(A \cap B | C) = \mathbb{P}(A | B \cap C) \mathbb{P}(B | C).$$

□

On reconnaît la formule du produit matriciel. Par conséquent la matrice $(P_{ij}^{n})_{i,j \in S}$ est tout simplement le produit n-fois de la matrice de transition P :

$$P^n = \underbrace{P \times P \times \ldots \times P}_{\text{(n fois)}} = (P_{ij}^{n})_{i,j \in S}.$$

Une transition de pas n correspond à multiplier n fois la matrice de transition. D'où

$$P^{n+m} = P^n P^m.$$

Conséquence 2.7. La loi à l'instant n du processus X_n ne dépend que de la loi de X_0 et de sa matrice de transition : soit $n \geqslant 0$, $j \in S$

$$\mathbb{P}(X_n = j) = \sum_{i \in S} P_{i,j}^{n} \mathbb{P}(X_0 = i).$$

Exemple *Prévisions météorologiques* (Ross p.196)
Reprenons l'exemple sur les prévisions météorologiques avec $p = 0.7$ et $q = 0.4$. La matrice de transition est

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}.$$

Sachant qu'il pleut aujourd'hui, on souhaite connaitre la probabilité qu'il pleuve dans 2 jours. On cherche donc $\mathbb{P}(X_{n+2} = 0 | X_n = 0)$ soit $P_{0,0}^{2}$. En calculant la puissance 2 de la matrice, on obtient

$$P^2 = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix}.$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Sachant qu'il pleut aujourd'hui, on a une probabilité de 61% qu'il pleuve dans 2 jours.
Maintenant, on souhaite connaître la probabilité qu'il pleuve dans 4 jours sachant qu'il pleut aujourd'hui, soit $P^4_{0,0}$.

$$P^4 = \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}.$$

Sachant qu'il pleut aujourd'hui, on a une probabilité de $57,49\%$ qu'il pleuve dans 4 jours.
Maintenant, on suppose que le premier jour il pleut avec probabilité $p_0 = 0.1$ et qu'il ne pleut pas avec probabilité $p_1 = 0.9$. On aimerait connaître la probabilité qu'il pleuve dans 4 jours.

$$\begin{aligned} \mathbb{P}(X_4 = 0) &= \sum_{i \in S} P^4_{0,i} \mathbb{P}(X_0 = i) = P^4_{0,0} \mathbb{P}(X_0 = 0) + P^4_{1,0} \mathbb{P}(X_0 = 1) \\ &= 0.5749 \times 0.1 + 0.5668 \times 0.9 = 0.5676 \end{aligned}$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

2.3 Irréductibilité et classification des états

On souhaite étudier les chemins possible de la chaîne de Markov. Est-ce que tous les chemins sont possible entre les différents états, ou est-ce que certains sont impossibles ? Est-ce que certains chemins peuvent être emprunté un nombre infini de fois ou est-ce que certains états ne sont visités qu'un nombre fini de fois ? Voilà les questions que l'on va se poser dans cette partie. Pour cela, on va classifier les états d'une chaîne de Markov en classes et ces classes seront entièrement déterminées par la matrice transition de la chaîne.

2.3.1 irréductibilité

On notera $i \rightsquigarrow j$, si i **mène à** j (ou j est accessible à partir de i), c'est à dire qu'il existe un chemin de probabilité strictement positive permettant de passer de i à j : il existe $n \geqslant 0$ tel que $P_{ij}^n > 0$.

On remarque que s'il n'existe pas $n \geqslant 0$ tel que $P_{i,j}^n > 0$, alors

$$\begin{aligned} \mathbb{P}(\text{atteindre } j|\text{partant de } i) &= \mathbb{P}\left(\bigcup_{n=0}^{\infty}\{X_n = j\}|X_0 = i\right) \\ &\leqslant \sum_{n=0}^{\infty}\mathbb{P}(X_n = j|X_0 = i) = \sum_{n=0}^{\infty} P_{i,j}^n = 0. \end{aligned}$$

Par conséquent, $i \rightsquigarrow j$ si et seulement si il existe $n \geqslant 0$ tel que $P_{ij}^n > 0$.

On dira que deux états i et j **communiquent**, noté $i \leftrightarrow j$, si $i \rightsquigarrow j$ et $j \rightsquigarrow i$.

On remarque que tout état communique avec lui même (car $\mathbb{P}(X_0 = i|X_0 = i) = 1$) et que si $i \leftrightarrow j$ et $j \leftrightarrow k$, alors $i \leftrightarrow k$.
La relation $\leftrightarrow$ est une relation d'équivalence. On dira que deux états qui communiquent sont dans la même **classe**. Ceci permet de décomposer l'espace d'état S en différentes classes, où à l'intérieur de chacune les états communiquent entre eux.
Pour faire cette décomposition, l'utilisation des graphes peut être utile.

Exemple *Serpent-Échelle*
On a $1 \rightsquigarrow 3$, par contre $3 \not\rightsquigarrow 1$, et $2 \leftrightarrow 3$. L'espace d'état se décompose selon les classes suivantes :

$$\{1\}, \{2, 3, 4, 5, 7, 8\}, \{11\}, \{12\}, \{13\}, \{14\}, \{15\}, \{16\}.$$

Définition 2.8. *Une chaîne de Markov est dite* **irréductible** *si elle ne possède qu'une seule classe, c'est à dire si tous les états communiquent entre eux.*

Exemple 2.9. *Considérons une chaîne de Markov d'espace d'état $S = \{0, 1, 2\}$ de matrice de transition*

$$P = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Tous les états communiquent, la chaîne est irréductible.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Exemple 2.10. *Considérons une chaîne de Markov d'espace d'état $S = \{0, 1, 2, 3\}$ de matrice de transition*

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3/4 & 1/4 & 0 & 0 \\ 1/8 & 0 & 7/8 & 0 \\ 0 & 0 & 1/9 & 8/9 \end{bmatrix}.$$

On souhaite définir les communications entre les états. On a $0 \leftrightarrow 1$, $0 \not\leftrightarrow 2$ et $2 \not\leftrightarrow 3$, les états 2 et 3 sont inaccessibles des états 1 et 0, l'état 3 est inaccessible de 2. La chaîne n'est pas irréductible, on a trois classes :

$$\{0, 1\}, \{2\}, \{3\}.$$

Exemple *Marche aléatoire*
Tous les états communiquent : pour tout $i, j \in \mathbb{Z}$, il existe un chemin pour passer de i à j. Par exemple, si $i < j$, il suffit d'aller toujours à droite, si $i > j$ il suffit d'aller toujours à gauche. La marche aléatoire est donc irréductible.

Exemple *Bonus-Malus*
De même que pour la marche aléatoire, tous les états communiquent. Il est facile d'aller en une étape d'un niveau i à au niveau $j > i$ (il faut avoir $j - i$ accidents) et si $j < i$ il suffit de ne pas avoir d'accident pendant suffisamment de temps. Le chaîne de Bonus-Malus est donc irréductible.

2.3.2 Temps d'atteinte, récurrence et transience

Définition 2.11. *Notons T_i le temps d'atteinte de i pour la première fois à partir de la première transition :*

$$T_i = \inf\{n \geqslant 1 : X_n = i\}.$$

On définit $f_i = \mathbb{P}(T_i < \infty | X_0 = i)$, c'est à dire la probabilité que le processus revienne au moins une fois en i en partant de i :

$$f_i = \mathbb{P}(\textit{il existe } n \geqslant 1 \textit{ tel que } X_n = i | X_0 = i)$$

On remarque que

$$f_i = \sum_{n=1}^{\infty} \mathbb{P}(T_i = n | X_0 = i) \quad \text{et} \quad \{T_i = n\} = \{X_1 \neq i, \ldots, X_{n-1} \neq i, X_n = i\}.$$

Exemple *Serpent-Échelle*
Dans le cadre du jeu Serpent-Échelle, soit $i \in \{1, \ldots 15\}$, $\mathbb{E}[T_{16} | X_0 = i]$ correspond à la durée moyenne de la partie sachant que je pars de la case i.

Définition 2.12. *Si $f_i = 1$, alors l'état i est dit* **récurrent** *: partant de l'état i je suis certain de retourner en i.*
Si $f_i < 1$, alors l'état i est dit **transitoire** *(ou transient).*
On dit que i est **absorbant**, *si $P_{ii} = 1$ (une fois arrivé en i on n'en bouge pas).*

Si i est un état récurrent, partant de i je suis sûr de retourner en i et ainsi de suite. Par conséquent ma chaîne passera par l'état i infiniment souvent.
Par contre, si i est un état transitoire, on a une probabilité strictement positive de ne pas retourner en i (égale à $1 - f_i$). Partant de i, le nombre de fois N_i que le processus passera en i suit une loi géométrique de paramètre $(1 - f_i)$: la probabilité de passer k fois en i est égale à $f_i^{k-1}(1 - f_i)$. Le nombre moyen de passage en i est donc fini égal à $1/(1 - f_i)$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Démonstration. On a

$$N_i = \sum_{n=0}^{\infty} \mathbb{1}_{X_n=i}.$$

Regardons tout d'abord la probabilité de ne pas retourner en i :

$$\mathbb{P}(N_i = 1|X_0 = i) = \mathbb{P}(T_i = \infty|X_0 = i) = 1 - \mathbb{P}(T_i < \infty|X_0 = i) = 1 - f_i.$$

Regardons maintenant la probabilité de n'y retourner qu'une seule fois :

$$\begin{aligned}
\mathbb{P}(N_i = 2|X_0 = i) &= \mathbb{P}(\text{il existe } n \geqslant 1\, X_n = i \text{ et } \forall m \neq n\, X_m \neq i|X_0 = i)\\
&= \sum_{n\geqslant 1} \mathbb{P}(X_n = i \text{ et } \forall m \neq n\, X_m \neq i|X_0 = i)\\
&\overset{\text{Bayes}}{=} \sum_{n\geqslant 1} \mathbb{P}(\forall m > n\, X_m \neq i|X_n = i \text{ et } \forall m \in \{1,\ldots,n-1\}\, X_m \neq i \text{ et } X_0 = i)\\
&\qquad \times\mathbb{P}(X_n = i \text{ et } \forall m \in \{1,\ldots,n-1\}\, X_m \neq i|X_0 = i)\\
&\overset{\text{Markov}}{=} \sum_{n\geqslant 1} \mathbb{P}(\forall m > n\, X_m \neq i|X_n = i)\mathbb{P}(T_i = n|X_0 = i)\\
&\overset{\text{Homogène}}{=} \sum_{n\geqslant 1} \mathbb{P}(\forall m > 0\, X_m \neq i|X_0 = i)\mathbb{P}(T_i = n|X_0 = i)\\
&= \sum_{n\geqslant 1} \mathbb{P}(T_i = \infty|X_0 = i)\mathbb{P}(T_i = n|X_0 = i) = (1 - f_i)f_i.
\end{aligned}$$

On montre en utilisant les mêmes arguments que pour tout $k \geqslant 0$

$$\begin{aligned}
\mathbb{P}(N_i = k|X_0 = i) &= \mathbb{P}(\text{il existe } 1 < n_1 < n_2 < \ldots < n_{k-1} : X_{n_j} = i\\
&\qquad \text{et } \forall m \neq n_j\, X_m \neq i \text{ pour } j \in \{1,\ldots,k-1\}|X_0 = i)\\
&= \sum_{1<n_1<n_2<\ldots<n_{k-1}} \mathbb{P}(X_{n_j} = i \text{ et } \forall m \neq n_j\, X_m \neq i \text{ pour } j \in \{1,\ldots,k-1\}|X_0 = i)\\
&= \sum_{1<n_1<n_2<\ldots<n_{k-1}} \mathbb{P}(\forall m > n_{k-1}\, X_m \neq i|X_{n_{k-1}} = i)\\
&\qquad \times\mathbb{P}(X_{n_j} = i \text{ et } \forall m < n_{k-1} : m \neq n_j\, X_m \neq i \text{ pour } j \in \{1,\ldots,k-1\}|X_0 = i)\\
&= (1 - f_i)f_i^{k-1}.
\end{aligned}$$

□

Exemple *Serpent-Échelle*
Dans le cas du jeu Serpent-Échelle, tous les états sont transitoires sauf la dernière case qui est absorbante. En effet, pour $i < 16$, il existe un chemin de probabilité strictement positive qui part de la case i et arrive à la case finale sans passer par l'échelle, donc qui ne permet pas de retourner en i. Par exemple, si $i = 2$

$$\mathbb{P}(T_2 = \infty|X_0 = 2) \geqslant \mathbb{P}(\text{avoir 6 au premier lancer, 6 au second lancer et 2 au dernier lancer}) = 1/6^3 > 0.$$

Donc $f_2 = 1 - \mathbb{P}(T_2 = \infty|X_0 = 2) < 1$.

Proposition 2.13. *Considérons un état $i \in S$, alors*

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

- i est récurrent si et seulement si $\sum_{n=1}^{\infty} P_{i,i}^n = \infty$,
- i est transitoire si et seulement si $\sum_{n=1}^{\infty} P_{i,i}^n < \infty$.

Par conséquent, si l'état i est transitoire, on a $P_{i,i}^n \to 0$ quand $n \to \infty$.

Démonstration. Posons $I_n = \mathbb{1}_{\{X_n=i\}}$. Alors $\sum_{n\geqslant 0} I_n$ représente le nombre de passage en i de la chaîne $(X_n)_{n\geqslant 0}$. On a

$$\mathbb{E}[\sum_{n=0}^{\infty} I_n | X_0 = i] = \sum_{n=0}^{\infty} \mathbb{P}[X_n = i | X_0 = i] = \sum_{n=1}^{\infty} P_{i,i}^n.$$

D'où le résultat. □

Remarque 2.14. *Soient $i, j \in S$ avec $i \leftrightarrow j$ et i récurrent, alors j est récurrent.*
Par conséquent tous les éléments d'une même classe sont de même nature (soit tous récurrents, soit tous transitoires).
Le résultat est intuitif, pour une preuve plus formelle vous pouvez aller voir le livre de Ross, p. 207.

Un état transitoire n'est visité qu'un nombre fini de fois, par conséquent si l'espace d'état S est fini, les états ne peuvent pas être tous transitoire (il suffit de prendre n assez grand). On a le résultat suivant

Théorème 2.15. *On considère une chaîne de Markov sur un espace d'état fini, alors il existe au moins un état récurrent.*
Si de plus la chaîne est irréductible (une seule classe), alors tous les états sont récurrents.

Exemple 2.16. *La chaîne de Markov à espace d'état $S = \{0, 1, 2\}$ de matrice de transition*

$$P = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

est irréductible en dimension finie, donc tous les états sont récurrents.

Exemple 2.17. *On considère une chaîne de Markov à espace d'état $S = \{0, 1, 2, 3, 4\}$ et de matrice de transition*

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 0 & 1/2 \end{bmatrix}.$$

On remarque que la chaîne possède trois classes $\{0, 1\}, \{2, 3\}, \{4\}$. Les deux premières sont récurrentes et la dernière est transitoire. En effet pour tout $n \geqslant 0$, $P_{i,i}^n = 1/2$ pour $i \neq 4$ et $P_{4,4}^n = (1/2)^n$, il suffit ensuite d'utiliser la Proposition 2.13.

Exemple *Marche aléatoire*
Tous les états sont récurrents dans le cadre de la marche aléatoire symétrique et transitoire dans le cas de la marche asymétrique. Intuitivement si $p > 1/2$, on a une plus forte probabilité d'aller à droite qu'à gauche et donc la chaîne à tendance à partir vers $+\infty$.
En dimension 2 (le buveur se ballade sur $\mathbb{Z}^2$) le résultat est le même, mais en dimension 3 quelque soit la valeur de p la chaîne est toujours transitoire.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Démonstration. (Ross p.209)
Il suffit d'étudier un seul état vu que la chaîne est irréductible. Considérons $i = 0$. Soit $n \geqslant 0$. Si $n = 2k+1$ est impair $P_{00}^n = 0$ (il est impossible de revenir en 0 en un nombre impair de coups). Si $n = 2k$ est pair, il faut aller autant de fois à gauche qu'à droite pour revenir en 0 :

$$P_{00}^{2k} = \binom{2k}{k}(p(1-p))^k,$$

où $\binom{2k}{k}$ est le nombre de façon de placer les pas vers la droite parmi les $n = 2k$ pas que fait le buveur.
On dit que $u_n \sim v_n$ si $\lim_{n\to\infty} \frac{u_n}{v_n} = 1$ (les deux suites se comportent de la même façon à l'infini).
En utilisant la formule de Stirling : $n! \sim n^{n+1/2}e^{-n}\sqrt{2\pi}$, on a

$$P_{00}^{2k} \sim \frac{(4p(1-p))^k}{\sqrt{\pi k}}$$

ce qui est le terme général d'une série convergente si $p \neq 1/2$ (car $4p(1-p) < 1$) et divergente si $p = 1/2$ (car $4p(1-p) = 1$).
En effet, la série $\sum_{n\geqslant 1} \frac{a^n}{n^b}$ converge si et seulement si $a < 1$ ou $a = 1$ et $b > 1$. □

2.3.3 Point absorbant, temps d'absorption

Considérons la chaîne de Markov $(X_n)_{n\geqslant 0}$ à espace d'état $S = \{0, 1, 2\}$ et de matrice de transition

$$P = \begin{bmatrix} 1 & 0 & 0 \\ p & q & r \\ 0 & 0 & 1 \end{bmatrix}$$

avec $0 < p, q, r < 1$ et $p + q + r = 1$. On remarque que la chaîne posséde deux points absorbants 2 et 0.
Partant de l'état 1, le processus peut y rester un certain temps, mais un jour il sautera soit à l'état 0 soit à l'état 2 dont il ne pourra plus en sortir. Il sera absorbé.
On peut se demander si le processus a plus de chance d'être absorbé en 0 ou en 2 et combien de temps en moyenne il faudra attendre avant que le processus soit absorbé ?
Pour cela on va utiliser une méthode qui s'appelle analyse des premiers (on regarde les premières transitions de la chaîne).
On définit $T = \min\{n \geqslant 0 \text{ tel que } X_n = 0 \text{ ou } X_n = 2\}$ le temps d'absorption du processus et on note

$$\begin{aligned} u &= \mathbb{P}(X_T = 0 | X_0 = 1), \\ v &= \mathbb{E}[T | X_0 = 1]. \end{aligned}$$

On suppose pour le moment que l'état initial du processus est $X_0 = 1$. Alors à l'instant 1 on a
— soit $X_1 = 0$ (avec proba p) et alors $T = 1$, $X_T = 0$;
— soit $X_1 = 2$ (avec proba r) et alors $T = 1$, $X_T = 2$;
— soit $X_1 = 1$ (avec proba q) et alors $T > 1$, et on débute alors une nouvelle transition partant à nouveau de l'état 1.

Par conséquent, $\mathbb{P}(T = 1 | X_0 = 1) = p + r$.
Par ailleurs, vu qu'on considère une chaîne de Markov homogène, on remarque que

$$\mathbb{P}(X_T = 0 | X_1 = 1) = \mathbb{P}(X_T = 0 | X_0 = 1) = u,$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

d'où u vérifie la relation

$$\begin{aligned} u &= \mathbb{P}(X_T=0|X_0=1) = \sum_{k=0}^{2}\mathbb{P}(X_T=0, X_1=k|X_0=1) \\ &= \sum_{k=0}^{2}\mathbb{P}(X_T=0|X_1=k)\mathbb{P}(X_1=k|X_0=1) \quad \text{(Bayes + propriété de Markov)} \\ &= 1.p + u.q + 0.r \end{aligned}$$

On en déduit que $u = \frac{p}{1-q} = \frac{p}{p+r}$. Par conséquent

$$\mathbb{P}(X_T=2|X_0=1) = 1 - \mathbb{P}(X_T=0|X_0=1) = \frac{r}{p+r}.$$

Cherchons maintenant le temps moyen d'absorption. Partant de l'état 1, il faut au moins une transition pour être absorbé. Comme on l'a vu plus haut, soit on est absorbé dès le premier pas, soit on repart à l'état initial et il faudra en moyenne $v = \mathbb{E}[T|X_0=1]$ transitions supplémentaires pour être absorbé (ceci est lié à l'absence de mémoire d'une chaîne de Markov), d'où

$$\begin{aligned} v &= \mathbb{E}[T|X_0=1] \\ &= 1.\mathbb{P}(T=1|X_0=1) + (1+\mathbb{E}[T|X_0=1])\mathbb{P}(X_1=1|X_0=1) \\ &= (p+r) + (1+v)q = 1 + vq \end{aligned}$$

On en déduite que le temps moyen d'absorption est $1/(p+r)$.

Remarque 2.18. *Dans ce cas particulier, on pouvait calculer le temps moyen d'absorption d'une autre manière. En effet, l'événement $\{T=n\}$, signifie je n'ai pas été absorbé pendant $n-1$ étapes et donc je suis resté en 1 et j'ai ensuite fait la transition vers 0 ou 2. La probabilité de cet événement est : $\mathbb{P}(T=n|X_0=1) = q^{n-1}(1-q)$. T suit la loi géométrique (loi du premier succès) de paramètre $1-q=p+r$, son espérance est donc $1/(p+r)$.*
L'intérêts de l'analyse des premiers pas est que l'on peut la généraliser à des exemples plus compliqués.

Exemple 2.19. *On considère deux jeux. Le premier consiste à lancer une pièce équilibrée jusqu'à obtenir 2 faces de suite suivi d'un pile : FFP, le second consiste à obtenir face, pile, face : FPF. On aimerait comparer les temps moyen d'une partie pour chacun des jeux.*
Le premier jeu : on considère 4 états : $S_1 = \{0, F, FF, FFP\}$.
On considère une chaîne de Markov $(X_n)_{n\geqslant 0}$ définie sur S_1 telle que l'on retourne au point de départ dès qu'on sort des configurations $\{F, FF, FFP\}$. Dès qu'on a atteint la configuration FFP le jeu s'arrête, ce point est absorbant. La matrice de transition de cette chaîne est par conséquent

$$P_1 = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Le second jeu : on considère 4 états : $S_2 = \{0, F, FP, FPF\}$.
On considère une chaîne de Markov $(Y_n)_{n\geqslant 0}$ définie sur S_2 telle que l'on retourne au point de départ dès qu'on sort des configurations $\{F, FP, FPF\}$. Dès qu'on a atteint la configuration FPF le jeu s'arrête, ce

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

point est absorbant. La matrice de transition de cette chaîne est par conséquent

$$P_2 = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Calculons les temps moyens d'une partie pour chacun des jeux.
Notons v_0 *le temps moyen pour le premier jeu partant de l'état* 0, v_F *quand on part de l'état* F *et* v_{FF} *quand on part de l'état* FF. *Bien évidemment si on part de* FFP *le jeu est terminé, on a* $v_{FFP} = 0$. *On utilise l'analyse des premiers pas :*

$$\begin{cases} v_0 = \frac{1}{2}(1+v_0) + \frac{1}{2}(1+v_F) = 1 + \frac{1}{2}v_0 + \frac{1}{2}v_F \\ v_F = 1 + \frac{1}{2}v_0 + \frac{1}{2}v_{FF} \\ v_{FF} = 1 + \frac{1}{2}v_{FF} + \frac{1}{2}v_{FFP} \end{cases}$$

D'où

$$v_0 = 8 \quad v_F = 6 \quad v_{FF} = 2.$$

Pour le second jeu, notons w_0 *le temps moyen pour le premier jeu partant de l'état* 0, w_F *quand on part de l'état* F *et* w_{FP} *quand on part de l'état* FP. *Bien évidemment si on part de* FPF, *on a* $w_{FPF} = 0$. *On utilise l'analyse des premiers pas :*

$$\begin{cases} w_0 = 1 + \frac{1}{2}w_0 + \frac{1}{2}w_F \\ w_F = 1 + \frac{1}{2}w_F + \frac{1}{2}w_{FP} \\ w_{FP} = 1 + \frac{1}{2}w_0 + \frac{1}{2}w_{FPF} \end{cases}$$

D'où

$$w_0 = 10 \quad w_F = 8 \quad w_{FP} = 6.$$

Plus de lancers sont nécessaires en moyenne pour gagner dans le second jeu, alors que la probabilité d'avoir FFP ou FPF en trois coups consécutifs est la même égale à $1/8$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

2.4 Probabilité limite, loi invariante

On considère une chaîne de Markov $(X_n)_{n\geqslant 0}$ à espace d'état S, de matrice de transition P. Connaissant le loi initiale, la loi à chaque instant n est donnée par la formule (voir la Conséquence 2.7 page 24)

$$\mathbb{P}(X_n = j) = \sum_{i \in S} P^n_{i,j} \mathbb{P}(X_0 = i).$$

la loi de X_n ne dépend que de X_0 et de la puissance $n^{\text{ème}}$ de la matrice de transition.
Considérons la matrice de transition de l'exemple *Prévisions météorologiques*. On observe que

$$P = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \quad P^{10} = \begin{bmatrix} 0.5714311 & 0.4285689 \\ 0.5714252 & 0.4285748 \end{bmatrix}$$
$$P^{20} = \begin{bmatrix} 0.5714286 & 0.4285714 \\ 0.5714286 & 0.4285714 \end{bmatrix} \quad P^{50} = \begin{bmatrix} 0.5714286 & 0.4285714 \\ 0.5714286 & 0.4285714 \end{bmatrix}$$

On a l'impression que les puissances de la matrice de transition convergent vers une matrice dont toutes les lignes sont identiques.
Remarque : ceci n'est pas vrai avec toutes les matrices de transition.
Dans cette section on va étudier la loi de X_n lorsque n grandit. Pour cela on aura besoin de notions supplémentaires (encore !).

2.4.1 Périodicité

Définition 2.20. *Considérons un état $i \in S$ et notons $R(i) = \{n \geqslant 0 : P^n_{ii} > 0\}$. On définit la période d d'un état $i \in S$ comme le plus grand commun diviseur de $R(i)$: $d = pgcdR(i)$.*
L'état i est dit **apériodique** *si $d = 1$.*

On remarque que la périodicité est une propriété de classe : tous les états d'une même classe ont la même période.

Exemple *Marche aléatoire*
Dans le cas de la marche aléatoire la période est 2. Il faudra forcément un chemin de longueur un multiple de 2 pour revenir au point initial (il faut aller autant de fois à gauche qu'à droite).
Par contre, si on considère un individu qui a vraiment trop bu et qui a des moments d'absence. Il a alors une probabilité p d'aller à droite, q d'aller à gauche et r de rester sur place, avec $0 < p, q, r < 1$ et $p + q + r = 1$. La chaîne est toujours irréductible, mais l'individu pouvant rester sur place, la période est 1.

Exemple *Bonus-Malus*
Partant du niveau 0 je peux rester à ce niveau, cet état est apériodique. La chaîne modélisant le système de Bonus-Malus étant irréductible (une seule classe). Tous les états sont donc apériodiques.

2.4.2 Récurrence positive/ récurrence nulle

Définition 2.21. *On considère un état récurrent $i \in S$. On note $\mu_i = \mathbb{E}[T_i | X_0 = i]$.*
- *L'état i est dit* **récurrent positif** *si partant du point i l'espérance du temps de retour en i est finie. Autrement dit, i est récurrent positif si $\mu_i < \infty$.*
- *Si $\mu_i = \infty$, l'état i est alors dit* **récurrent nul**.
- *Un état récurrent positif et apériodique est dit* **ergodique**.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Propriété 2.22. — *La récurrence positive (ou nulle) est une propriété de classe.*
- — *Dans le cas d'un espace d'état S fini, tous les états récurrents sont forcément récurrents positifs.*
- — *Un état i récurrent est récurrent nul si $P_{ii}^n \to 0$ quand $n \to \infty$, sinon il est récurrent positif.*

Théorème 2.23. *Une chaîne de Markov irréductible à espace d'état* **fini** *est nécessairement récurrente positive (i.e., tous les états sont récurrents positifs).*

Exemple *Marche aléatoire*
La marche aléatoire symétrique ($p = 1/2$) est récurrente nulle. Le résultat est admis, mais pour ceux qui souhaite avoir une preuve la voici.

Démonstration. (Voir *Promenade aléatoire* de Michel Benaïm et Nicole El Karoui, Ed. École Polytechnique) *Non fait en cours.*
Soit $i \in \mathbb{Z}$. On veut calculer $\mathbb{E}[T_i | X_0 = i]$. On note $\varphi_i(t) = \mathbb{E}[e^{tT_i} | X_0 = i] = \sum_{n \geqslant 1} e^{nt} \mathbb{P}(T_i = n | X_0 = i)$ la fonction génératrice des moments de T_i sachant que $X_0 = i$. On sait retrouver l'espérance à l'aide $\varphi_i(t)$, on va donc calculer cette fonction. On remarque que la marche aléatoire est invariante par translation, donc $\varphi_i(t)$ ne dépend pas de i (quelque soit la valeur de i, le problème est le même et ne dépend pas de la valeur de i). Notons alors cette fonction $\varphi(t)$.
Prenons $i = 0$. On sait que $T_0 \geqslant 2$ car il est impossible de revenir au point de départ en une seule étape, on va
- — soit d'abord à droite (avec probabilité $1/2$) et alors $X_1 = 1$ et $T_0 = 1 + T_{1,0}$ où $T_{1,0}$ est le temps pour aller de 1 à 0,
- — soit d'abord à gauche (avec probabilité $1/2$) et alors $X_1 = -1$ et $T_0 = 1 + T_{-1,0}$ où $T_{-1,0}$ est le temps pour aller de -1 à 0.

Par conséquent

$$\begin{aligned} \varphi(t) &= \mathbb{E}[e^{tT_0} | X_0 = 0] \\ &= \frac{1}{2}\mathbb{E}[e^{t(1+T_{1,0})} | X_0 = 0] + \frac{1}{2}\mathbb{E}[e^{t(1+T_{-1,0})} | X_0 = 0] \\ &= \frac{1}{2}e^t \left(\mathbb{E}[e^{tT_{1,0}} | X_0 = 0] + \mathbb{E}[e^{tT_{-1,0}} | X_0 = 0] \right) \\ &= e^t \mathbb{E}[e^{tT_{1,0}} | X_0 = 0] \quad \text{(marche symétrique)} \end{aligned}$$

On recommence, pour aller de 1 à 0, on y va
- — soit en un pas avec probabilité $1/2$, d'où $X_2 = 0$ et $T_{1,0} = 1$,
- — soit on saute encore à droite avec probabilité $1/2$ et on se retrouve en 2, d'où $X_2 = 2$ et $T_{1,0} = 1 + T_{2,0}$ où $T_{2,0}$ est le temps pour aller de 2 à 0.

On a donc

$$\mathbb{E}[e^{tT_{1,0}} | X_0 = 0] = \frac{1}{2}e^t \left(1 + \mathbb{E}[e^{tT_{2,0}} | X_0 = 0]\right)$$

Pour aller de 2 à 0, il est nécessaire de passer par 1, donc $T_{2,0} = T_{2,1} + T_{1,0}$, avec $T_{2,1}$ et $T_{1,0}$ indépendantes (du à la propriété de Markov) et de même loi (invariance par translation). Donc

$$\mathbb{E}[e^{tT_{1,0}} | X_0 = 0] = \frac{1}{2}e^t \left(1 + \mathbb{E}[e^{tT_{1,0}} | X_0 = 0]^2\right).$$

On en déduit que

$$\mathbb{E}[e^{tT_{1,0}} | X_0 = 0] = e^{-t}(1 - \sqrt{1 - e^{2t}}),$$

et

$$\varphi(t) = 1 - \sqrt{1 - e^{2t}}.$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

(vérification : la variable T_0 étant positive, $\varphi(t)$ est une fonction croissante.)
Regardons la dérivée de φ en 0 : $\varphi'(t) = \frac{e^{2t}}{\sqrt{1-e^{2t}}}$, d'où pour tout $i \in \mathbb{Z}$, $\mathbb{E}[T_i|X_0 = i] = \varphi'(0) = \infty$. Tous les points sont bien récurrents nuls. □

2.4.3 Loi stationnaire

Théorème 2.24. *Si la chaîne de Markov est irréductible et ergodique (une seule classe dont les éléments sont tous ergodiques), alors* $\lim_{n\to+\infty} P_{ij}^n$ *existe et ne dépend pas de* i*.*

Notons $\pi_j = \lim_{n\to+\infty} P_{ij}^n$ *pour* $j \in S$*. Alors* $\pi = (\pi_j)_{j\in S}$ *est l'unique loi de probabilité sur* S *solution de*

$$\pi_j = \sum_{i\in S} \pi_i P_{ij} \quad \textit{pour tout } j \in \mathcal{S}.$$

π *est appelée* **loi stationnaire** *(ou loi invariante) de la chaîne de Markov.*

On sait que (voir conséquence 2.7 page 24), la loi de X_n vérifie

$$\mathbb{P}(X_n = j) = \sum_{i\in S} P_{i,j}^n \mathbb{P}(X_0 = i).$$

Sous les hypothèses du théorème, quand n converge vers $+\infty$, $\mathbb{P}(X_n = j)$ converge vers $\sum_{i\in S} \pi_j \mathbb{P}(X_0 = i) = \pi_j$. Donc $X_n \xrightarrow[n\to\infty]{Loi} \pi$.
Si on considère une chaîne irréductible et ergodique, quelque soit sa condition initiale, pour n suffisamment grand, la loi de X_n *est "proche" de* π *et reste par la suite proche de* π*.*

Remarque 2.25.
- *C'est l'irréductibilité de la chaîne qui assure l'unicité de la loi invariante. Si on n'a pas l'irréductibilité et que chaque classe est ergodique, il y a alors existence d'une infinité de lois invariantes.*
- *La convergence en loi est assurée par l'apériodicité de la chaîne.*
- *Si la chaîne est irréductible et ergodique, la suite* $(P^n)_{n\geqslant 0}$ *converge quand n tend vers l'infini vers une matrice dont toutes les lignes sont identiques, chaque ligne correspondant à la loi invariante* π*. Ceci permet en général de calculer* π *numériquement par ordinateur lorsque la dimension de la matrice n'est pas trop grande, comme on l'a fait au début de cette section pour les prévisions météorologiques.*
- *La loi invariante* π *peut être considérée comme un vecteur ligne* $\pi = (\pi_0, \pi_1, \dots)$ *et donc* π *est solution des équations*
$$\pi.P = \pi \ \textit{ et } \ \sum_{i\in S} \pi_i = 1.$$

Proposition 2.26. *Sous les hypothèses du théorème, si la loi initiale de* X_0 *est* π*, alors pour tout* $n \geqslant 0$*,* X_n *suit la loi* π*.*
D'où l'appellation de loi stationnaire.

Démonstration. En effet, pour $n = 0$ la loi de X_0 est bien π et si on suppose que X_n suit la loi π, on a

$$\mathbb{P}(X_{n+1} = j) = \sum_{i\in S} P_{i,j} \mathbb{P}(X_n = i) = \sum_{i\in S} P_{i,j} \pi_i = \pi_j.$$

Par induction, pour tout $n \geqslant 0$, la loi de X_n est π. □

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Proposition 2.27. *Soit $(X_n)_{n\geqslant 0}$ une chaîne de Markov à espace d'état S. On suppose la chaîne irréductible, récurrente* (pas forcément positive) *et apériodique, alors pour tout état $i \in S$ on a*

$$\lim_{n\to+\infty} P^n_{ii} = \frac{1}{\mu_i},$$

où μ_i est l'espérance $\mathbb{E}[T_i|X_0 = i]$ du temps de retour en i partant de l'état i.

On retrouve le dernier résultat de la propriété 2.22.

Conséquence 2.28. Si la chaîne est irréductible et ergodique et si on note π sa loi stationnaire, on a alors pour tout $i \in S$

$$\pi_i = \frac{1}{\mu_i}.$$

Trouver π revient à résoudre un système linéaire, c'est donc un moyen assez simple de calculer le temps moyen de retour dans un état partant de cet état.

Exemple *Prévisions météorologiques*
Dans cet exemple, l'espace d'état est fini et la chaîne est irréductible, donc tous les états sont récurrents positifs. Par ailleurs, l'état 0 est apériodique car $P_{0,0} > 0$. La chaîne est irréductible et ergodique, elle admet donc une unique probabilité stationnaire.
Calculons la loi invariante dans le cas des prévisions météorologies. $\pi = (\pi_0, \pi_1)$ vérifie

$$\pi.P = \pi$$

soit

$$\begin{cases} 0.7\pi_0 + 0.4\pi_1 = \pi_0 \\ 0.3\pi_0 + 0.6\pi_1 = \pi_1 \end{cases}$$

Comme π est une loi de probabilité, on a aussi $\pi_0 + \pi_1 = 1$. On doit par conséquent résoudre le système

$$\begin{cases} -0.3\pi_0 + 0.4\pi_1 = 0 \\ \pi_0 + \pi_1 = 1 \end{cases}$$

On obtient $\pi = (4/7, 3/7)$.
En utilisant la Proposition 2.27, on en déduit en particulier que le temps moyen d'attente d'un jour de pluie sachant qu'il pleut aujourd'hui est 7/4 jours.

Exemple 2.29. *On considère une chaîne de Markov à espace d'état $S = \{0, 1, 2\}$ de matrice de transition*

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

Chaque période de temps que le processus passe à l'état 0 coûte 2\$ *et les période de temps passées aux états* 1 *et* 2 *coûtent chacune* 1\$*. On peut se demander quel est coût moyen par période de temps à long terme de cette chaîne.*
En dessinant le graphe, on observe que la chaîne est irréductible à espace d'état fini. Par ailleurs, comme $P_{2,2} > 0$ l'état 2 est apériodique (et donc tous les autres états). La chaîne est par conséquent irréductible et ergodique, alors la loi istationnaire π existe et est unique. Calculons la :

$$\begin{cases} \frac{1}{2}\pi_1 + \frac{1}{3}\pi_2 = \pi_0 \\ \pi_0 + \frac{1}{3}\pi_2 = \pi_1 \\ \frac{1}{2}\pi_1 + \frac{1}{3}\pi_2 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases} \Leftrightarrow \begin{cases} -\pi_0 + \frac{1}{2}\pi_1 + \frac{1}{3}\pi_2 = 0 \\ \pi_0 - \pi_1 + \frac{1}{3}\pi_2 = 0 \\ \frac{1}{2}\pi_1 + \frac{2}{3}\pi_2 = 0 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{cases}$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

D'où $\pi = (\frac{3}{10}, \frac{2}{5}, \frac{3}{10})$. *À long terme la loi de* X_n *est proche de* π*, donc le coût moyen à long terme est*

$$C = 2.\frac{3}{10} + 1.\frac{2}{5} + 1.\frac{3}{10} = 1,3\$.$$

Exemple *Bonus-Malus*
Revenons sur l'exemple des Bonus-Malus. Considérons un cas plus simple où l'espace d'état est fini $S = \{0,1,2,3,4,5,6\}$ (correspond au système brésilien, voir *Bonus-Malus systems : The european and asian approach to merit-rating* de Jean Lemaire, dans North American Actuarial Journal (1998)). Le principe est le même mais on ne peut pas dépasser le niveau 6. Par exemple, si on était dans la classe 2 et qu'on a eu plus de 4 accidents au cours de l'année, alors notre nouvelle prime sera celle du niveau 6. On choisit le même paramètre λ pour chacun des niveaux : la probabilité d'avoir k accidents est donc

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}$$

et la matrice de transition est

$$P = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & \sum_{k \geqslant 6} p_k \\ p_0 & 0 & p_1 & p_2 & p_3 & p_4 & \sum_{k \geqslant 5} p_k \\ 0 & p_0 & 0 & p_1 & p_2 & p_3 & \sum_{k \geqslant 4} p_k \\ 0 & 0 & p_0 & 0 & p_1 & p_2 & \sum_{k \geqslant 3} p_k \\ 0 & 0 & 0 & p_0 & 0 & p_1 & \sum_{k \geqslant 2} p_k \\ 0 & 0 & 0 & 0 & p_0 & 0 & \sum_{k \geqslant 1} p_k \\ 0 & 0 & 0 & 0 & 0 & p_0 & \sum_{k \geqslant 1} p_k \end{bmatrix}.$$

On remarque que, comme $\sum_{k=0}^{\infty} p_k = 1$, on a

$$\sum_{k \geqslant n} p_k = 1 - \sum_{k=0}^{n-1} p_k.$$

La matrice de transition s'écrit donc

$$P = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & p_4 & p_5 & 1 - \sum_{k=0}^{5} p_k \\ p_0 & 0 & p_1 & p_2 & p_3 & p_4 & 1 - \sum_{k=0}^{4} p_k \\ 0 & p_0 & 0 & p_1 & p_2 & p_3 & 1 - \sum_{k=0}^{3} p_k \\ 0 & 0 & p_0 & 0 & p_1 & p_2 & 1 - p_0 - p_1 - p_2 \\ 0 & 0 & 0 & p_0 & 0 & p_1 & 1 - p_0 - p_1 \\ 0 & 0 & 0 & 0 & p_0 & 0 & 1 - p_0 \\ 0 & 0 & 0 & 0 & 0 & p_0 & 1 - p_0 \end{bmatrix}.$$

La chaîne est irréductible, apériodique (l'état 0 est apériodique car $P_{0,0} = p_0 > 0$), donc ergodique. Il existe une unique probabilité stationnaire π. Pour $\lambda = 0.1$, on obtient

$$p_0 = 0.9048374, p_1 = 0.0904837, p_2 = 0.0045242, p_3 = 0.0001508, p_4 = 3.8.10^{-6}, p_5 = 7.54.10^{-8}$$

et donc

$$\pi = (0.88948, 0.09355, 0.01444, 0.00215, 0.00032, 0.00005, 0.00001).$$

Par conséquent, en régime "stationnaire" (en "temps long"), 89% des assurés sont dans l'état 0 de plus bas niveau et partant de l'état 0 il faut en moyenne attendre 1/0.88948 années pour revenir à l'état 0, soit 1 an 1 mois et 15 jours.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

La valeur de π permet à l'assureur d'ajuster la valeur de la prime dans chaque niveau afin d'être solvable et compétitif. La valeur de λ (qui correspond au nombre moyen d'accident par assuré) est estimée à l'aide de méthodes statistiques en fonction des données que l'assureur possède sur ses assurés.
Supposons que $\lambda = 0.1$, la compagnie d'assurance décide d'augmenter le montant de la prime de 20% à chaque fois qu'on passe au niveau supérieur, c'est à dire

— le montant de la prime dans l'état 0 est $a_0 = a$,
— le montant de la prime dans l'état 1 est $a_1 = a_0 + a_0 * 20\% = 1.2a$,
— le montant de la prime dans l'état 2 est $a_2 = a_1 + a_1 * 20\% = (1.2)^2 a$,
$\vdots \qquad \vdots$
— le montant de la prime dans l'état 6 est $a_6 = a_5 + a_5 * 20\% = (1.2)^6 a$.

La prime moyenne payée par assuré est donc de

$$\begin{aligned} \text{prime moyenne} &= a_0\pi_0 + a_1\pi_1 + \ldots + a_6\pi_6 \\ &= a(\pi_0 + 1.2\pi_1 + \ldots + (1.2)^6\pi_6) \\ &= a \times 1.02704 \end{aligned}$$

Afin d'être solvable et compétitive, la compagnie d'assurance souhaite que la prime moyenne payée par assuré soit de 100\$, il faut donc que $a \times 1.02704 = 100$, soit $a = 97.36$\$. On obtient ainsi le montant de la prime dans chaque niveau

Niveau	0	1	2	3	4	5	6
Montant de la prime	97.36$	116.83$	140.2$	168.25$	201.9$	242.27$	290.73$

2.4.4 Théorème ergodique

Cette partie est hors-programme, juste pour la culture et pour les gens qui devront un jour faire des simulations avec des chaînes de Markov. On peut aller voir le livre de Norris, *Markov Chain*, Cambridge University Press (en partie disponible sur sa page internet : http ://www.statslab.cam.ac.uk/ james/Markov/) ou l'article de Galin L. Jones pour le théorème central limite pour les chaînes de Markov, *On the Markov chain central limit theorem*, Probab. Surveys, Volume 1 (2004), pages 299-320 (disponible sur http ://projecteuclid.org/).
Le théorème ergodique est l'équivalent de la loi des grands nombres pour les chaînes de Markov.
La loi des grands nombre stipule que si on a une suite $(X_n)_{n\geqslant 0}$ de variables indépendantes, de même loi et d'espérance fini, alors avec probabilité 1 on a

$$\frac{1}{n}\sum_{k=1}^{n} X_k \underset{n\to+\infty}{\longrightarrow} \mathbb{E}[X_1].$$

On peut aussi le réécrire, pour toute fonction f bornée, on a avec probabilité 1

$$\frac{1}{n}\sum_{k=1}^{n} f(X_k) \underset{n\to+\infty}{\longrightarrow} \mathbb{E}[f(X_1)].$$

Ce théorème est très utilisé en simulation pour simuler $\mathbb{E}[f(X_1)]$ qui n'est pas toujours facile à calculer. Pour cela, on simule n variables indépendantes de même loi et d'après la loi des grands nombres $\frac{1}{n}\sum_{k=1}^{n} f(X_k)$ est proche de la valeur cherchée. Lorsqu'on travaille avec des variables qui ne sont pas indépendantes la loi des grands nombres ne s'applique plus.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Théorème 2.30. *Soit $(X_n)_{n\geqslant 0}$ une chaîne de Markov à espace d'état S. On suppose la chaîne irréductible récurrente positive et on note π sa loi invariante. Alors pour toute fonction f bornée, alors quelque soit la loi initiale, avec probabilité 1,*

$$\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \underset{n\to+\infty}{\longrightarrow} \sum_{i\in S} f(i)\pi_i.$$

Un second théorème ergodique peut s'apparenter au théorème central limite. On peut écrire le théorème central limite de la manière suivante : si $(X_n)_{n\geqslant 0}$ est une suite de variables indépendantes, de même loi et d'espérance m et de variance σ^2 finie, alors

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=1}^{n} X_k - m\right) \underset{n\to+\infty}{\overset{Loi}{\longrightarrow}} \mathcal{N}(0,\sigma^2).$$

Théorème 2.31. *Soit $(X_n)_{n\geqslant 0}$ une chaîne de Markov irréductible à espace d'état S fini. On considère une fonction f et on note*

$$m = \sum_{i\in S} f(i)\pi_i.$$

Alors, sous certaines hypohèses, il existe σ dépendant de f tel que

$$\sqrt{n}\left(\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) - m\right) \underset{n\to+\infty}{\overset{Loi}{\longrightarrow}} \mathcal{N}(0,\sigma^2).$$

2.5 Marche aléatoire absorbée : problème de la ruine du joueur

On considère un joueur qui joue soit jusqu'à ce qu'il soit ruiné, soit jusqu'à ce qu'il ait atteint son objectif de repartir avec un capital de N\$. À chaque fois qu'il joue, il gagne 1\$ ou perd 1\$. On suppose que les jeux successifs sont indépendants entre eux et qu'à chaque fois le joueur a une probabilité p de gagner et une probabilité $q = 1-p$ de perdre.
On souhaite connaître la probabilité que le joueur ressorte du casino riche plutôt que ruiné en fonction de son capital initial.

Notons X_n la richesse du joueur à l'instant n. $(X_n)_{n\geqslant 0}$ est une chaîne de Markov à valeur dans $\{0,1,\ldots,N\}$. À chaque instant le joueur gagne ou perd juste 1\$, par conséquent les probabilités de transitions sont pour $i \in \{1,2,\ldots,N-1\}$

$$\begin{aligned} P_{0,0} &= P_{N,N} = 1 \\ P_{i,i+1} &= 1 - P_{i,i-1} = p \end{aligned}$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

et les autres probabilités de transitions sont nulles.
Cette chaîne a deux états absorbants et on peut décomposer l'espace d'état en trois classes : deux récurrente $\{0\}, \{N\}$ et une transitoire $\{1, 2, \dots, N-1\}$.
Vu que les états transitoires ne sont visités qu'un nombre fini de fois, on est sur que le temps de jeux est fini (on arrivera forcément dans un des états absorbants). On note T la durée du jeu :

$$T = \min\{n \geqslant 0 : X_n = 0 \text{ ou } X_n = N\}.$$

À l'instant T le joueur est soit ruiné soit riche, d'où $X_T = 0$ ou N. On note $P_i = \mathbb{P}(X_T = N | X_0 = i)$ la probabilité de sortir du jeu riche sachant que le capital initial du joueur était i.
Alors en utilisant l'analyse des premiers pas,

$$\begin{aligned} P_i &= \mathbb{P}(X_T = N, X_1 = i+1 | X_0 = i) + \mathbb{P}(X_T = N, X_1 = i-1 | X_0 = i) \\ &= \mathbb{P}(X_T = N | X_1 = i+1) P_{i,i+1} + \mathbb{P}(X_T = N | X_1 = i-1) P_{i,i-1} \end{aligned}$$

La probabilité $\mathbb{P}(X_T = N | X_1 = i+1)$ représente la probabilité d'être riche sachant qu'on avait un capital de $i+1$, par conséquent P_{i+1}. Le suite (P_i), vérifie donc

$$P_i = pP_{i+1} + qP_{i-1}.$$

Comme $p + q = 1$, si on note $u_i = P_{i+1} - P_i$, on a $pu_i = qu_{i-1}$. D'où $u_i = \left(\frac{q}{p}\right)^i u_0$.
Comme $P_0 = 0$, on a donc

$$\begin{aligned} P_2 - P_1 &= \frac{q}{p} P_1 \\ P_3 - P_2 &= \frac{q}{p}(P_2 - P_1) = \left(\frac{q}{p}\right)^2 P_1 \\ &\vdots \\ P_i - P_{i-1} &= \left(\frac{q}{p}\right)^{i-1} P_1 \\ &\vdots \\ P_N - P_{N-1} &= \left(\frac{q}{p}\right)^{N-1} P_1 \end{aligned}$$

En sommant les i premières équations, on obtient

$$P_i - P_1 = \sum_{k=1}^{i-1} \left(\frac{q}{p}\right)^k P_1.$$

On distingue deux cas, si $\frac{q}{p} = 1$, alors $P_i = iP_1$ et si $\frac{q}{p} \neq 1$, $P_i = \dfrac{1 - \left(\frac{q}{p}\right)^i}{1 - \frac{q}{p}} P_1$.

(En effet, de manière générale, si $a \neq 1$ on a $\displaystyle\sum_{k=0}^{n} a^k = \frac{1 - a^{n+1}}{1 - a}$.)

Vu que $P_N = 1$, on en déduit que $P_1 = 1/N$ si $\frac{q}{p} = 1$ et $P_1 = \dfrac{1 - \frac{q}{p}}{1 - \left(\frac{q}{p}\right)^N}$ sinon.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Par ailleurs $\frac{q}{p} = 1$ si et seulement si $p = q = 1/2$. Par conséquent, on a

$$\begin{cases} P_i = \dfrac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^N} & \text{si } p \neq 1/2; \\ P_i = \dfrac{i}{N} & \text{si } p = 1/2. \end{cases}$$

On a $q/p < 1$ si et seulement si $p > 1/2$. Si on considère que la richesse de la banque augmente $N \to +\infty$, on a

$$P_i \to \begin{cases} 1 - \left(\frac{q}{p}\right)^i & \text{si } p > 1/2; \\ 0 & \text{si } p \leqslant 1/2. \end{cases}$$

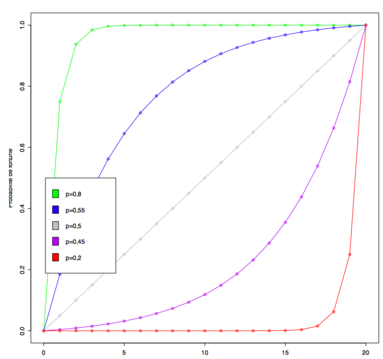
Quand $N \to +\infty$ le joueur est par conséquent sur d'être ruiné si le jeu lui est défavorable ($p < 1/2$), ce qui n'est pas étonnant, mais c'est aussi le cas même si le jeu est équitable. Par contre, si le jeu lui est favorable $p > 1/2$, le joueur a une probabilité strictement positive (mais strictement inférieur à 1) que sa richesse croisse indéfiniment.

Proposition 2.32. Ruine du joueur
On considère une marche aléatoire entre deux barrières absorbantes 0 *et* N*, d'espace d'état* $S = \{0, 1, \ldots, N\}$*. On suppose que les probabilités de transitions sont données par*

$$P_{i,i+1} = p = 1 - P_{i,i-1} \textit{ pour } i = 1, \ldots N-1 \textit{ et } P_{0,0} = P_{N,N} = 1.$$

Alors la probabilité d'absorption en N *en partant de l'état* i *est*

$$P_i = \begin{cases} \dfrac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^N} & \textit{si } p \neq 1/2, \\ \frac{i}{N} & \textit{si } p = 1/2. \end{cases}$$



Probabilité de sortir gagnant en fonction du capital initial pour différentes valeurs de p.

Exemple 2.33. *On considère Max et Patty qui jouent à pile ou face. Patty a* 5$ *en poche et Max a* 10$ *en poche. Chacun parie* 1$ *à chaque fois. Les deux amis jouent jusqu'à la ruine d'un des joueurs. On cherche la probabilité de gagner pour Patty.*
Il suffit d'imaginer un boulier avec 5 *boules du coté gauche et* 10 *boules du coté droit. A chaque fois que Patty gagne on déplace une boule de droite à gauche et à chaque fois que Max gagne on déplace une boule de gauche à droite. Le jeu se termine quand il n'y a plus de boule d'un coté.*

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

On note X_n le nombre de boules du coté de Patty (la fortune de Patty). Il y a par conséquent $15 - X_n$ boules du coté de Max. X_n est à valeur dans $\{0, 1, \dots, 15\}$. On a $X_{n+1} = X_n + 1$ si Patty a gagné au jeu de pile ou face et sinon on a $X_{n+1} = X_n - 1$. On a bien une chaîne de Markov.
D'après les résultats sur la ruine du joueur, la pièce étant supposée équilibrée ($p = q = 1/2$), la probabilité que Patty gagne est $P_5 = \mathbb{P}(X_T = 15 | X_0 = 5)$:

$$P_5 = \frac{5}{15} = \frac{1}{3}.$$

Par conséquent, Max gagne avec une probabilité $2/3$.

Exercice 2.34. Supposons qu'à chaque partie d'un jeu, Alexandre gagne 1\$ avec probabilité p, ou perd 1\$ avec probabilité $q = 1 - p$. Alexandre continue de jouer jusqu'à temps qu'il gagne N\$ ou perde M\$. Quelle est la probabilité qu'Alexandre quitte le jeu en tant que gagnant ?

Solution succinte : En se ramenant au problème de la ruine du joueur, on cherche la probabilité qu'un joueur partant avec M\$ atteigne $(N + M)$\$ avant d'être ruiné. Donc la probabilité qu'Alexandre gagne est :

$$\frac{1 - \left(\frac{q}{p}\right)^M}{1 - \left(\frac{q}{p}\right)^{M+N}}.$$

△

Exercice 2.35. Pour le problème de la ruine du joueur, notons M_i le nombre moyen de parties avant que le joueur soit ruiné ou gagne une fortune de N, sachant qu'il part d'un capital i ($i \in \{0, 1, ..., N\}$). Donnez un ensemble d'équations linéaires afin de trouver M_i ($i \in \{0, 1, ..., N\}$).

Solution succinte : Par une analyse des premiers pas, on a $M_0 = 0 = M_N$ et pour $i \in \{1, \dots, N - 1\}$

$$M_i = 1 + pM_{i+1} + qM_{i-1}.$$

△

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Chapitre 3

Processus de Poisson

3.1 Introduction

On considère une compagnie d'assurance. Elle souhaite évaluer le montant des primes que doivent verser ses assurés afin qu'elle soit solvable, tout en restant concurrentielle par rapport aux autres compagnies d'assurance. Chaque client verse une certaine prime chaque année. En retour, si le client a un accident la compagnie doit lui verser des indemnités. Pour évaluer le montant de la prime la compagnie d'assurance devra par conséquent évaluer le nombre d'indemnités qu'elle devra verser chaque année et leur montant.

On va se placer sous les hypothèses suivantes : les clients se comportent tous de la même manière et de façon indépendante.

Pour répondre au problème de la compagnie d'assurance, il va falloir modéliser les sinistres (les instants où il y a des sinistres et leur montant) et ensuite évaluer la probabilité de ruine de la compagnie d'assurance en fonction de la prime fixée.

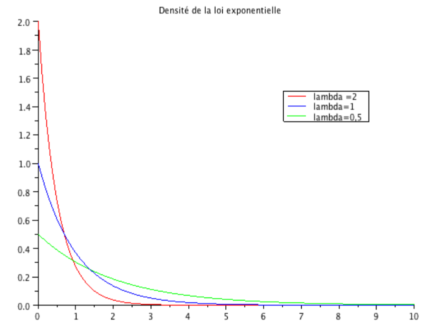
3.2 Notions sur la loi exponentielle et la loi de Poisson

3.2.1 Loi exponentielle

Définition 3.1. *Une variable aléatoire X de loi exponentielle de paramètre (ou de taux) $\lambda > 0$ est une variable continue à valeurs positives de densité*

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{x \geqslant 0}.$$

La loi exponentielle de paramètre λ est notée $\mathcal{E}(\lambda)$.



Propriété 3.2. *Soit X une variable aléatoire de loi exponentielle $\mathcal{E}(\lambda)$.*

1. *Sa fonction de répartition est*

$$F(t) = \begin{cases} 1 - e^{-\lambda t} & si\ t \geqslant 0 \\ 0 & si\ t < 0. \end{cases}$$

2. *Sa fonction génératrice des moments est*

$$\varphi(t) = \mathbb{E}[e^{tX}] = \begin{cases} \infty & si\ t \geqslant \lambda \\ \frac{\lambda}{\lambda - t} & si\ t < \lambda. \end{cases}$$

3. *Sa moyenne et sa variance sont*

$$\mathbb{E}[X] = \frac{1}{\lambda} \quad Var(X) = \frac{1}{\lambda^2}.$$

Preuve.

1. $F(t) = \mathbb{P}(X \leqslant t) = 0$ si $t < 0$ car X est une variable positive et si $t \geqslant 0$, on a

$$F(t) = \int_0^t f_X(x)dx = \int_0^t \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x}\right]_0^t = 1 - e^{-\lambda t}.$$

2. Sa fonction génératrice des moments vérifie

$$\varphi(t) = \mathbb{E}[e^{tX}] = \int_0^{+\infty} \lambda e^{tx} e^{-\lambda x} dx = \int_0^{+\infty} \lambda e^{(t-\lambda)x} dx = \left[\frac{\lambda}{t-\lambda}^{(t-\lambda)x}\right]_0^{+\infty}.$$

D'où le résultat.

3. Calculons les dérivées de la fonction génératrice, on a

$$\varphi'(t) = \frac{\lambda}{(\lambda - t)^2} \quad \varphi''(t) = \frac{2\lambda}{(\lambda - t)^3}.$$

D'où

$$\mathbb{E}[X] = \varphi'(0) = \frac{1}{\lambda} \text{ et } Var(X) = \varphi''(0) - \mathbb{E}[X]^2 = \frac{1}{\lambda^2}.$$

$\square$

La loi exponentielle possède une propriété dite d'absence de mémoire. Elle est en fait la seule loi continue à vérifier cette propriété :

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Propriété 3.3. Absence de mémoire
Soit X une variable aléatoire de loi exponentielle, alors pour tout $t, s \geqslant 0$,

$$\mathbb{P}(X > s+t | X > t) = \mathbb{P}(X > s).$$

Preuve. En effet,

$$\mathbb{P}(X > s+t | X > t) = \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = \mathbb{P}(X > s).$$

□

Du fait de sa propriété d'absence de mémoire, la loi exponentielle est souvent utilisée pour modéliser des durées de vie pour un système sans usure ou des temps d'attente. Par exemple, si on modélise la durée de vie d'une lampe par la loi exponentielle, l'absence de mémoire revient à dire "sachant que la lampe fonctionne encore après t heures de service, la probabilité qu'elle fonctionne encore pendant au moins s heures est égale à la probabilité qu'elle fonctionne au moins de s heures". C'est comme si la lampe avait oublié qu'elle avait déjà fonctionné t heures, voilà pourquoi on utilise cette loi pour des systèmes sans usure.
On peut réécrire la propriété d'absence de mémoire de la façon suivante

$$\mathbb{P}(X > s+t) = \mathbb{P}(X > s)\mathbb{P}(X > t).$$

Exemple 3.4. *On suppose que le temps d'attente T à la poste suit une loi exponentielle de moyenne 10 minutes ($\lambda = 1/10$). La probabilité qu'un client attende plus de 15 minutes est*

$$\mathbb{P}(T > 15) = e^{-15/10} = e^{-3/2} \simeq 0.22.$$

Sachant que le client a déjà attendu au moins 10 minutes, le probabilité qu'il attende au moins 15 minutes est

$$\mathbb{P}(T > 15 | T > 10) = \mathbb{P}(T > 5) = e^{-5/10} = e^{-1/2} \simeq 0.604.$$

Proposition 3.5. *Le minimum de variables de loi exponentielle indépendantes suit encore une loi exponentielle : soient $X \sim \mathcal{E}(\lambda)$ et $Y \sim \mathcal{E}(\mu)$ deux variables indépendantes, alors*

$$\min(X, Y) \sim \mathcal{E}(\lambda + \mu).$$

De plus

$$\mathbb{P}(\min(X, Y) = X) = \mathbb{P}(X \leqslant Y) = \frac{\lambda}{\lambda + \mu}.$$

Preuve. On pose $Z = \min(X, Y)$. Z est une variable positive. Calculons sa fonction de survie : pour $t \geqslant 0$ on a

$$\mathbb{P}(Z > t) = \mathbb{P}(X > t, Y > t) \overset{\text{Indépendance}}{=} \mathbb{P}(X > t)\mathbb{P}(Y > t) = e^{-(\lambda+\mu)t}.$$

Par ailleurs, comme les variables X et Y sont indépendantes, la loi du couple (X, Y) est continue de densité $f_{X,Y}(x, y) = \lambda e^{-\lambda x} \mu e^{-\mu y} \mathbb{1}_{x \geqslant 0} \mathbb{1}_{y \geqslant 0}$, d'où

$$\mathbb{P}(Z = X) = \mathbb{P}(X \leqslant Y) = \int_0^{+\infty} \lambda e^{-\lambda x} \left(\int_x^{+\infty} \mu e^{-\mu y} dy \right) dx = \lambda \int_0^{+\infty} e^{-(\lambda+\mu)x} dx = \frac{\lambda}{\lambda + \mu}.$$

□

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Exercice 3.6. Une usine fabrique des lampes dont la durée de vie T en heures vérifie : $P(T > t) = e^{-\lambda t}$ pour $t > 0$, avec $\lambda > 0$. T suit donc la loi exponentielle de paramètre λ. La durée de vie moyenne d'une lampe est $1/\lambda$.
On considère un échantillon de n lampes dont les durées de vie $T_1, \ldots, T_n$ sont supposées indépendantes, de même loi que T. On note $U = \min(T_1, \ldots, T_n)$ le premier instant où au moins une des lampes cesse de fonctionner et $V = \max(T_1, \ldots, T_n)$ le premier instant où toutes les lampes ont cessées de fonctionner. Quelles sont les lois de U et de V ?

Solution succinte : On va calculer les fonctions de répartition de U et V pour trouver leur loi. Les deux variables sont à valeurs positives. On a pour $t \geqslant 0$

$$F_U(t) = 1 - \mathbb{P}(U > t) = 1 - \mathbb{P}(T_1 > t, \ldots, T_n > t) = 1 - e^{-n\lambda t} \text{ par indépendance.}$$

Par conséquent, U suit le loi exponentielle $\mathcal{E}(n\lambda)$.

$$F_V(t) = \mathbb{P}(V \leqslant t) = \mathbb{P}(T_1 \leqslant t, \ldots, T_n \leqslant t) = \left(1 - e^{-\lambda t}\right)^n \text{ par indépendance.}$$

V est une variable de densité

$$f_V(t) = n\lambda e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{n-1} \mathbb{1}_{t \geqslant 0}.$$

△

Proposition 3.7. *La somme de n variables indépendantes de loi exponentielle de même paramètre suit la loi Gama : soient $X_1, \ldots, X_n$ des variables indépendantes de loi $\mathcal{E}(\lambda)$, alors*

$$\sum_{i=1}^n X_i \sim G(n, \lambda)$$

où $G(n, \lambda)$ est la loi Gamma de paramètre n et λ de densité

$$\frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \mathbb{1}_{x \geqslant 0}.$$

Preuve. On prouve le résultat par induction. Pour $n = 1$, on retrouve bien la loi exponentielle. On suppose que le résultat est vrai pour n, cherchons la loi de

$$\sum_{i=1}^{n+1} X_i = \sum_{i=1}^{n} X_i + X_{n+1}.$$

Les variables $\sum_{i=1}^n X_i$ et X_{n+1} sont indépendantes, donc la densité vérifie

$$\begin{aligned} f_{\sum_{i=1}^{n+1} X_i}(x) &= \int_0^{+\infty} f_{X_{n+1}}(x-y) f_{\sum_{i=1}^n X_i}(y) dy \\ &= \int_0^x \lambda e^{-\lambda(x-y)} \frac{\lambda^n}{(n-1)!} y^{n-1} e^{-\lambda y} dy \\ &= \frac{\lambda^{n+1}}{n!} e^{-\lambda x} \int_0^x n y^{n-1} dy = \frac{\lambda^{n+1}}{n!} e^{-\lambda x} x^n. \end{aligned}$$

Ce qui prouve le résultat. □

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Conséquence 3.8. On déduit de la proposition précédente l'espérance, la variance et la fonction génératrice des moments d'une variable Y de loi Gamma $G(n, \lambda)$:

$$\mathbb{E}[Y] = \frac{n}{\lambda},\ Var(Y) = \frac{n}{\lambda},\ \varphi_Y(t) = \begin{cases} \infty & \text{si } t \geqslant \lambda \\ \left(\frac{\lambda}{\lambda - t}\right)^n & \text{si } t < \lambda. \end{cases}$$

Preuve. En effet, comme les X_i sont indépendantes et de même loi, on a

$$\mathbb{E}[Y] = n\mathbb{E}[X],\ Var(Y) = nVar(X) \text{ et } \varphi_Y(t) = (\varphi_X(t))^n.$$

□

3.2.2 Loi de Poisson

Définition 3.9. *Une variable aléatoire X de loi Poisson de paramètre $\lambda > 0$ est une variable discrète à valeur dans $\mathbb{N}$ de probabilité*

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

La loi de Poisson de paramètre λ est notée $\mathcal{P}(\lambda)$.

Cette loi est en général utilisée des événements rares comme le nombre d'accidents de voiture, le nombre de mutations génétiques fixées dans l'ADN,…En fait, la loi de Poisson a été introduite en 1838 par Siméon-Denis Poisson dans son ouvrage Recherches sur la probabilité des jugements en matière criminelle et en matière civile.

Propriété 3.10. *Soit X une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda)$.*

— *Sa fonction génératrice des moments est*

$$\varphi(t) = \mathbb{E}[e^{tX}] = e^{\lambda(e^t - 1)}.$$

— *Sa moyenne et sa variance sont*

$$\mathbb{E}[X] = \lambda \quad Var(X) = \lambda.$$

Preuve.

$$\varphi(t) = \sum_{k=0}^{\infty} e^{tk}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!} = e^{-\lambda}e^{\lambda e^t}.$$

On calcule les dérivées de la fonction génératrice :

$$\varphi'(t) = \lambda e^t e^{\lambda(e^t-1)} \quad \varphi''(t) = (1 + \lambda e^t)\lambda e^t e^{\lambda(e^t-1)}$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

D'où
$$\mathbb{E}[X] = \varphi'(0) = \lambda \text{ et } Var(X) = \varphi''(0) - \mathbb{E}[X]^2 = \lambda.$$

□

Proposition 3.11. *La somme de n variables indépendantes de loi Poisson suit encore une loi de Poisson : soient $X_1, \ldots, X_n$ des variables indépendantes de loi respective $\mathcal{P}(\lambda_i)$, alors*
$$\sum_{i=1}^{n} X_i \sim \mathcal{P}(\sum_{i=1}^{n} \lambda_i)$$

Preuve. On calcule la fonction génératrice de $Z = \sum_{i=1}^{n} X_i$ et on utilise le fait que la fonction génératrice caractérise la loi : par indépendance des X_i on a
$$\varphi_Z(t) = \varphi_{X_1}(t) \times \varphi_{X_2}(t) \times \ldots \times \varphi_{X_n}(t) = e^{\lambda_1(e^t-1)}.e^{\lambda_2(e^t-1)} \ldots e^{\lambda_n(e^t-1)} = e^{(e^t-1)\sum_{i=1}^{n}\lambda_i}.$$

□

Exercice 3.12. On considère une compagnie d'assurance habitation. On note X les sinistres causés par une personne de manière volontaire (vols, feux volontaires, ...) et Y les sinistres accidentels. On suppose que X et Y sont indépendantes et suivent des loi de Poisson de paramètre respectif λ et μ. Montrer que la loi conditionnelle de X sachant le nombre total de sinistres au cours de l'année est une loi binomiale : la loi conditionnelle de X sachant $X + Y = n$ est la loi binomiale $\mathcal{B}\left(n, \frac{\lambda}{\lambda+\mu}\right)$.

Solution succinte : On sait que $X+Y$ suit la loi de Poisson de paramètre $\lambda+\mu$. Par ailleurs, si $\{X+Y = n\}$ alors forcément $X \leqslant n$. On a donc pour $k > n$ $\mathbb{P}(X = k|X + Y = n) = 0$ et pour $k \leqslant n$

$$\begin{aligned}
\mathbb{P}(X = k|X + Y = n) &= \frac{\mathbb{P}(X = k, X + Y = n)}{\mathbb{P}(X + Y = n)} = \frac{\mathbb{P}(X = k, Y = n - k)}{\mathbb{P}(X + Y = n)} \\
&= \frac{\mathbb{P}(X = k)\mathbb{P}(Y = n - k)}{\mathbb{P}(X + Y = n)} \text{ par indépendance de } X \text{ et } Y \\
&= \frac{n!}{k!(n-k)!}.\frac{\lambda^k e^{-\lambda}\mu^{n-k}e^{-\mu}}{(\lambda+\mu)^n e^{-(\lambda+\mu)}} \\
&= \begin{pmatrix} n \\ k \end{pmatrix} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right)^{n-k} = \begin{pmatrix} n \\ k \end{pmatrix} \left(\frac{\lambda}{\lambda+\mu}\right)^k \left(1 - \frac{\lambda}{\lambda+\mu}\right)^{n-k}.
\end{aligned}$$

△

3.3 Calcul d'espérance par conditionnement

Les lois conditionnelles dans les cas discrets et continus sont abordées dans le Livre de Ross, chapitre 3, page 97.
Considérons deux variables aléatoires X et Y. On appelle **espérance conditionnelle** de X sachant Y, noté $\mathbb{E}[X|Y]$, la variable qui prend la valeur $\mathbb{E}[X|Y = y]$ en y.
Cette quantité est une fonction de Y (et ne dépend surtout pas de X !).
Si X et Y sont des variables indépendantes alors l'espérance conditionnelle est constante : $\mathbb{E}[X|Y] = \mathbb{E}[X]$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

L'espérance conditionnelle vérifie la relation

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]].$$

Si Y est discrète, l'expression précédente se réécrit

$$\mathbb{E}[X] = \sum_y \mathbb{E}[X|Y = y]\mathbb{P}(Y = y),$$

et si Y est une variable continue de densité $f_Y(y)$ on a

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} \mathbb{E}[X|Y = y] f_Y(y).$$

Exemple 3.13. *Un relecteur travaille sur deux livres. Le nombre de coquilles dans le premier livre suit une loi de Poisson de moyenne* 2 *et le nombre de coquilles dans le second livre suit une loi de Poisson de paramètre* 5*. Le relecteur choisit aléatoirement de façon équitable le livre sur lequel il va travailler. On note* X *le nombre de coquilles que le relecteur va relever. On note* Y *le choix du livre :* $Y = 1$ *s'il relit le livre* 1 *et* $Y = 2$ *s'il relit le livre* 2*.*
Alors le nombre moyen de coquilles vaut

$$\mathbb{E}[X] = \mathbb{E}[X|Y = 1]\mathbb{P}(Y = 1) + \mathbb{E}[X|Y = 2]\mathbb{P}(Y = 2) = 2 \times 1/2 + 5 \times 1/2 = 7/2.$$

Exemple *Somme d'un nombre aléatoire de variables aléatoires*
On pose N le nombre d'accidents de voitures pendant une année. Pour le $i^{\text{ème}}$ accident, on note Z_i le montant des indemnités que la compagnie d'assurance verse au conducteur. On suppose que les Z_i sont indépendants et de même loi et indépendants de N.
Le montant total que la compagnie devra débourser au cours d'une année est

$$X = \sum_{i=1}^{N} Z_i.$$

On souhaite calculer sa valeur moyenne. Comme X s'écrit à l'aide de N, il est naturel de conditionner par rapport à cette valeur. On a

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[\sum_{i=1}^{N} Z_i|N]\right]$$

Par ailleurs, on a

$$\begin{aligned}
\mathbb{E}[\sum_{i=1}^{N} Z_i|N = n] &= \mathbb{E}[\sum_{i=1}^{n} Z_i|N = n] = \sum_{i=1}^{n} \mathbb{E}[Z_i|N = n] \\
&= \sum_{i=1}^{n} \mathbb{E}[Z_i] \quad \text{par indépendance entre les } Z_i \text{ et } N \\
&= n\mathbb{E}[Z] \quad \text{car les } Z_i \text{ sont de même loi.}
\end{aligned}$$

Par conséquent $\mathbb{E}[X|N] = N\mathbb{E}[Z]$, on en déduit que

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[\sum_{i=1}^{N} Z_i|N]\right] = \mathbb{E}[N\mathbb{E}[Z]] = \mathbb{E}[N]\mathbb{E}[Z].$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

On peut généraliser le calcul d'espérance au calcul de l'espérance de n'importe quelle fonction h de X : si Y est discrète, on a

$$\mathbb{E}[h(X)] = \sum_y \mathbb{E}[h(X)|Y = y]\mathbb{P}(Y = y),$$

et si Y est une variable continue de densité $f_Y(y)$ on a

$$\mathbb{E}[h(X)] = \int_{-\infty}^{+\infty} \mathbb{E}[h(X)|Y = y]f_Y(y).$$

Ceci est utile pour calculer la variance ou la fonction génératrice d'une variable aléatoire. Notamment pour la variance $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$, on obtient la formule suivante

Proposition 3.14. (Formule de décomposition de la variance)

$$Var(X) = \mathbb{E}[Var(X|Y)] + Var(\mathbb{E}[X|Y]).$$

Preuve. On a

$$\begin{aligned}
\mathbb{E}[Var(X|Y)] &= \mathbb{E}\left[\mathbb{E}[X^2|Y] - \mathbb{E}[X|Y]^2\right] \\
&= \mathbb{E}\left[\mathbb{E}[X^2|Y]\right] - \mathbb{E}\left[\mathbb{E}[X|Y]^2\right] = \mathbb{E}[X^2] - \mathbb{E}\left[\mathbb{E}[X|Y]^2\right]
\end{aligned}$$

et

$$Var(\mathbb{E}[X|Y]) = \mathbb{E}\left[\mathbb{E}[X|Y]^2\right] - \mathbb{E}\left[\mathbb{E}[X|Y]\right]^2 = \mathbb{E}\left[\mathbb{E}[X|Y]^2\right] - \mathbb{E}[X]^2.$$

En sommant les deux expressions, on obtient le résultat. □

Exemple *Somme d'un nombre aléatoire de variables aléatoires (appelé modèle agrégé)*
On reprend l'exemple précédent et on veut calculer la variance des remboursements au cours d'une année. On va utiliser la formule de décomposition de la variance.
On a $\mathbb{E}[X|N] = N\mathbb{E}[Z]$, par conséquent

$$Var(\mathbb{E}[X|N]) = Var(N\mathbb{E}[Z]) = \mathbb{E}[Z]^2 Var(N).$$

D'un autre coté, on a

$$\begin{aligned}
Var(X|N = n) &= Var(\sum_{i=1}^n Z_i|N = n) \\
&= \sum_{i=1}^n Var(Z_i|N = n) \quad \text{car les } Z_i \text{ sont indépendants} \\
&= \sum_{i=1}^n Var(Z_i) \quad \text{car les } Z_i \text{ et } N \text{ sont indépendants} \\
&= nVar(Z) \quad \text{car les } Z_i \text{ sont de même loi.}
\end{aligned}$$

D'où $Var(X|N) = NVar(Z)$ et donc $\mathbb{E}[Var(X|N)] = \mathbb{E}[N]Var(Z)$. Par la formule de décomposition de la variance, on en déduit que

$$Var(X) = \mathbb{E}[Z]^2 Var(N) + Var(Z)\mathbb{E}[N].$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Notamment, si on suppose que le nombre de sinistres suit une loi de Poisson $\mathcal{P}(\lambda)$, on a $\mathbb{E}[N] = Var(N) = \lambda$ et donc

$$Var(X) = \lambda(\mathbb{E}[Z]^2 + Var(Z)) = \lambda\mathbb{E}[Z^2].$$

Des exemples précédents on obtient la proposition suivante :

Proposition 3.15. (Espérance et variance d'une somme aléatoire de variables aléatoires)
On considère une suite $(Z_i)_{i\geqslant 1}$ des variables indépendantes et de même loi que Z et N une variable aléatoire. On suppose que N et les Z_i sont indépendants, alors

$$\mathbb{E}\left[\sum_{i=1}^{N} Z_i\right] = \mathbb{E}[N]\mathbb{E}[Z],\ Var\left(\sum_{i=1}^{N} Z_i\right) = \mathbb{E}[Z]^2 Var(N) + Var(Z)\mathbb{E}[N].$$

Exemple 3.16. Modèle de Galton-Watson *(Ross p. 245)*
En 1873, Galton s'est inquiété de la disparition des noms de familles nobles en Angleterre. Avec l'aide du mathématicien Watson ils ont étudié ce problème. Ce modèle est encore très utilisé en dynamique des populations et en génétique.

On suppose qu'à l'instant initial, $t = 0$ on a un seul ancêtre. On suppose que les générations ne se chevauchent pas et que chaque individu à la génération n donne naissance à des enfants de façon indépendante et selon la même loi $p = (p_0, p_1, \ldots)$ pour tout le monde, loi que l'on connaît, avec $p_i < 1$ pour tout $i \geqslant 0$. On note $\mu = \sum_{i\geqslant 0} ip_i$ le nombre moyen d'enfant par individu et σ^2 la variance de la loi p.

Le nombre d'individu X_{n+1} à la génération $n+1$ est égal à la somme des enfants de chacun des individus de la génération n, soit

$$X_{n+1} = \sum_{k=1}^{X_n} Z_k,$$

où Z_k est le nombre d'enfant du $i^{\text{ème}}$ individu de la génération n. Les variables Z_k sont indépendantes de X_n.
Ce type de processus est aussi appelé **processus de branchement** *ou* **de ramification**.
La suite $(X_n)_{n\geqslant 0}$ est une chaîne de Markov à espace d'état $S = \mathbb{N}$ et l'état 0 est absorbant.
Si $p_0 > 0$, alors tous les états non nuls sont transitoires, car on a un probabilité non nulle de passer de l'état i à l'état 0 égale à p_0^i (les i individus n'ont aucun enfant) et de cet état on ne peut pas retourner en i (l'état 0 est absorbant).

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Calculons la taille moyenne de la population et sa variance. X_n s'exprimant comme la somme aléatoire de variables aléatoires, d'après la proposition 3.15, on a

$$\mathbb{E}[X_n] = \mathbb{E}[Z]\mathbb{E}\left[X_{n-1}\right] = \mu\mathbb{E}\left[X_{n-1}\right].$$

Par conséquent, comme $X_0 = 1$,

$$\mathbb{E}[X_n] = \mu^n.$$

De même, en utilisant la proposition 3.15 on a

$$Var(X_n) = \mu^2 Var(X_{n-1}) + \sigma^2\mathbb{E}[X_{n-1}] = \mu^2 Var(X_{n-1}) + \sigma^2\mu^{n-1}.$$

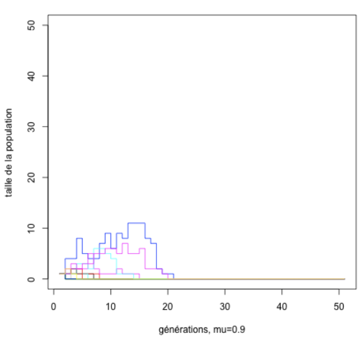
Par conséquent, comme $X_0 = 1$

$$\begin{array}{rcl}
Var(X_1) & = & \sigma^2 \\
Var(X_2) & = & \mu^2\sigma^2 + \sigma^2\mu = \sigma^2\mu(\mu+1) \\
Var(X_3) & = & \mu^2\sigma^3(\mu+1) + \sigma^2\mu^2 = \sigma^2\mu^2(\mu^2+\mu+1) \\
& \vdots & \\
Var(X_n) & = & \sigma^2\mu^{n-1}(\mu^{n-1} + \mu^{n-2} + \ldots + 1) = \sigma^2\mu^{n-1}\displaystyle\sum_{k=0}^{n-1}\mu^k
\end{array}$$

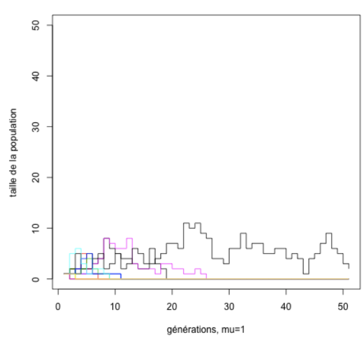
On a donc

$$Var(X_n) = \begin{cases} \sigma^2\mu^{n-1}\frac{1-\mu^n}{1-\mu} & si\ \mu \neq 1 \\ n\sigma^2 & si\ \mu = 1 \end{cases}$$

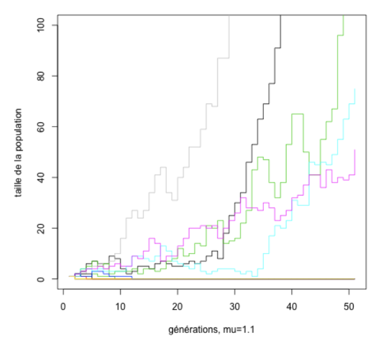
Dans les graphes ci-dessous on a tracé plusieurs trajectoires de X_n en fonction de n pour différentes valeur de μ (μ étant le nombre moyen d'enfant par individu) :



$\mu = 0.9$



$\mu = 1$



$\mu = 1.1$

Si $\mu < 1$, le nombre moyen d'individu dans la population et sa variance tendent vers 0 quand $n \to \infty$.
Si $\mu = 1$, le nombre moyen d'individu dans la population reste constant, mais sa variance part à l'infini quand $n \to \infty$.
Si $\mu > 1$, le nombre moyen d'individu dans la population et sa variance tendent vers ∞ quand $n \to \infty$.
En fait, on peut montrer que si $\mu \leqslant 1$, la population s'éteint de façon certaine et si $\mu > 1$ la probabilité d'extinction de la population est strictement inférieure à 1.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

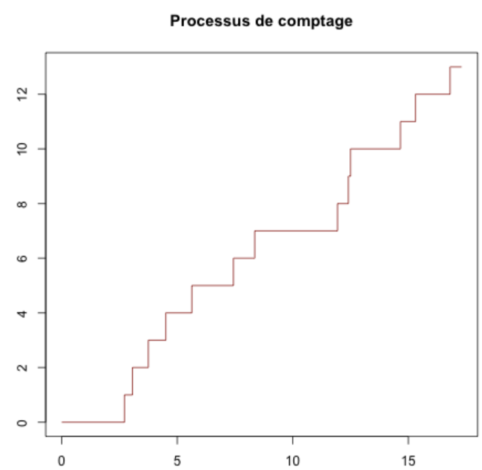
3.4 Processus Ponctuel de Poisson

Définition 3.17. *Un processus stochastique* $(N(t))_{t\geqslant 0}$ *est appelé* **processus de comptage** *si* $N(t)$ *représente le nombre total de* sauts *(ou* événements*) qui sont arrivés avant l'instant* t*.*

Exemple 3.18.
- *Dans le cas de la compagnie d'assurance,* $N(t)$ *représente le nombre de sinistres intervenus sur l'intervalle de temps* $[0, t]$*. Chaque* saut *correspond à l'arrivée d'un nouveau sinistre.*
- *On peut compter le nombre de visiteurs dans un musée qui sont arrivés avant l'instant* t*. Chaque* saut *correspond à l'arrivée d'un nouveau visiteur. D'ailleurs si vous allez au musée des beaux-arts de Montréal en entrant vous croiserez une personne muni d'un compteur.*

Un processus de comptage est forcément positif, à valeurs entières et croissant.

Propriété 3.19. *Soit* $(N(t))_{t\geqslant 0}$ *un processus de comptage. On a*
- $N(t) \geqslant 0$,
- $N(t) \in \mathbb{N}$,
- *Si* $s < t$*, on a* $N(s) \leqslant N(t)$*,*
- *Pour* $s < t$*,* $N(t) - N(s)$ *représente le nombre de sauts intervenus dans l'intervalle de temps* $(s, t]$*.*



On va étudier ici un processus de comptage particulier : le processus de Poisson.

3.4.1 Première définition

Définition 3.20. Première définition du Processus de Poisson *Le processus de comptage* $(N(t))_{t\geqslant 0}$ *est appelé* **processus de Poisson** *d'intensité* λ*, avec* $\lambda > 0$ *si*

(i) $N(0) = 0$,

(ii) le processus a des **accroissements indépendants***, c'est à dire pour* $t > s$*, le nombre de sauts* $N(t) - N(s)$ *intervenus sur* $(s, t]$ *est indépendant du nombre de sauts* $N(s)$ *intervenus avant l'instant* s*.*

(iii) le nombre d'événements sur un intervalle de longueur t *est distribué selon une loi de Poisson* $\mathcal{P}(\lambda t)$*, c'est à dite pour tout* $s, t \geqslant 0$

$$\mathbb{P}(N(t+s) - N(s) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Les accroissements de même longueur du processus de Poisson ont tous la même loi, on dit que le processus a des **accroissements stationnaires***.*

Le processus de Poisson est un processus de Markov : les accroissements étant indépendants, la valeur du processus après l'instant t ne dépend que de la valeur du processus à l'instant t mais pas de tout ce qui s'est passé avant.
Comme $N(t) \sim \mathcal{P}(\lambda t)$ pour $t > 0$, on remarque que le nombre moyen de saut par intervalle de temps est $\mathbb{E}[N(t)]/t = \lambda$.

Exemple 3.21. *On suppose que les accidents qu'aura un conducteur se produisent selon un processus de Poisson $(N(t))_{t\geqslant 0}$ d'intensité 0.01 par mois.*
La probabilité que le conducteur n'ait aucun accident au cours de sa première année de conduite est donc

$$\mathbb{P}(N(12) = 0) = e^{-0.01\times 12} \simeq 88,7\%,$$

car $N(12)$ suit la loi de Poisson $\mathcal{P}(0.01 \times 12)$.
La probabilité que le conducteur ait exactement un accident entre sa seconde année de conduite et sa troisième année de conduite est

$$\mathbb{P}(N(36) - N(24) = 1) = \mathbb{P}(N(12) - N(0) = 1) = \mathbb{P}(N(12) = 1) = 0.01 \times 12.e^{-0.01\times 12} \simeq 10,6\%,$$

car les accroissements sont stationnaires.
La probabilité qu'il ait au moins deux accidents avant la fin de sa deuxième année de conduite sachant qu'il a eut exactement un accident au cours de sa première année est

$$\begin{aligned}
\mathbb{P}(N(24) \geqslant 2|N(12) = 1) &= \mathbb{P}(N(24) - N(12) \geqslant 1|N(12) = 1)\\
&= \mathbb{P}(N(24) - N(12) \geqslant 1) \quad \textit{car les accroissements sont indépendants}\\
&= \mathbb{P}(N(12) \geqslant 1) \quad \textit{car les accroissements sont stationnaires}\\
&= 1 - \mathbb{P}(N(12) = 0) \simeq 11,3\%.
\end{aligned}$$

Remarque 3.22. *Soit $(N(t))_{t\geqslant 0}$ un processus de Poisson d'intensite λ. Le nombre de saut sur un petit intervalle de temps ne dépasse généralement pas 1. Plus précisément, soit $h > 0$, pour h suffisamment petit on a $\lambda h < 1$ et donc*

$$\mathbb{P}(N(h) = 0) = e^{-\lambda h} \quad \mathbb{P}(N(h) = 1) = \lambda h e^{-\lambda h} \quad \mathbb{P}(N(h) \geqslant 2) \leqslant \frac{(\lambda h)^2}{2}$$

et quand h temps vers 0, on a

$$\frac{\mathbb{P}(N(h) = 1)}{h} \to \lambda \quad \textit{et} \quad \frac{\mathbb{P}(N(h) \geqslant 2)}{h} \to 0.$$

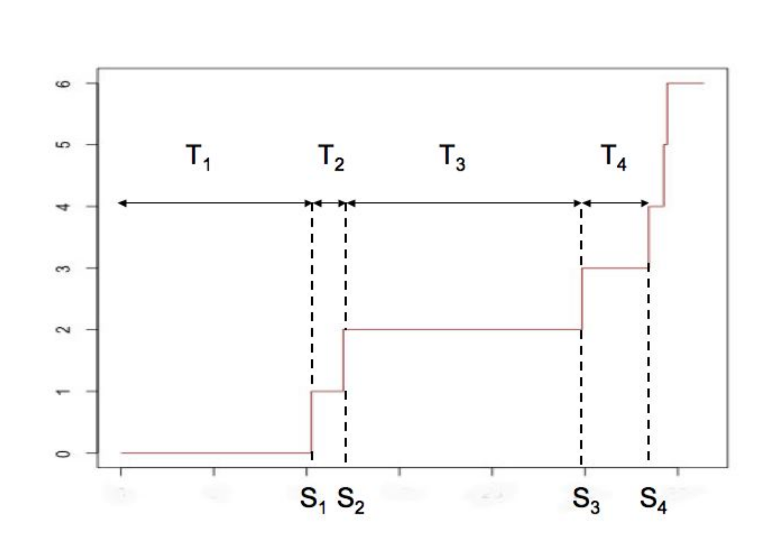
3.4.2 Distribution des temps d'attente et des inter-arrivées

Un processus de Poisson sert à compter des sauts (ou des événements). On peut se demander à quels moments arrivent ces sauts, et quelle distance il y a entre deux sauts. Les sauts arrivent de façon aléatoire, on va donc essayer de trouver la loi d'attente entre deux sauts : la loi des instants inter-arrivées.
Considérons un processus de Poisson $(N(t))_{t\geqslant 0}$. Le premier saut arrive à un instant aléatoire T_1, puis il faudra attendre un temps T_2 avant que le second saut survienne, puis un temps T_3 ainsi de suite.... On note T_n le temps écoulé entre le $(n-1)^{\text{ème}}$ saut et le $n^{\text{ème}}$ saut.
La suite des instants $(T_n)_{n\geqslant 1}$ est appelée suite des **instants inter-arrivées**.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

D'autres quantités intéressantes sont les instants de chaque saut :

$$S_n = \sum_{i=1}^{n} T_i.$$



Cherchons la loi de T_n et de S_n.

Théorème 3.23. *La suite des instants inter-arrivées $(T_n)_{n\geqslant 1}$ est une suite de variables indépendantes identiquement distribuées de loi exponentielle $\mathcal{E}(\lambda)$.*
Par conséquent, la durée moyenne d'un intervalle de temps entre deux sauts est $1/\lambda$.

Preuve. Étudions le premier instant de saut T_1. Si à l'instant t on a $N(t) = 0$ ceci signifie que l'on n'a pas encore sauté et donc $T_1 > t$. Par conséquent

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N(t) = 0) = e^{-\lambda t}.$$

Ceci implique que T_1 suit la loi exponentielle de paramètre λ.
Maintenant étudions la loi de T_2 en conditionnant par rapport à T_1 : soit $t, s > 0$

$$\mathbb{P}(T_2 > t | T_1 = s) = \mathbb{P}(\text{pas de saut sur l'intervalle }]s, s+t] | T_1 = s)$$

Le nombre de saut sur l'intervalle $]s, s+t]$ suit une loi de Poisson $\mathcal{P}(\lambda t)$ et est indépendant de ce qui s'est passé avant, par conséquent

$$\mathbb{P}(T_2 > t | T_1 = s) = \mathbb{P}(\text{pas de saut sur l'intervalle }]s, s+t]) = e^{-\lambda t}.$$

Par conséquent $T_2 \sim \mathcal{E}(\lambda)$ et les variables T_2 et T_1 sont indépendantes. En itérant ce raisonnement, on prouve le théorème. $\square$

Corollaire 3.24. *L'instant $S_n = \sum_{i=1}^{n} T_i$ est la somme de n variables indépendantes de loi exponentielle, elle suit donc la loi Gamma $G(n, \lambda)$.*
On remarque que

$$\{N(t) \leqslant n\} = \{S_n \geqslant t\}.$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Exemple 3.25. *On suppose que les nouveaux arrivants au Québec arrivent selon un processus de Poisson d'intensité $\lambda = 1$ par jour.*
Le temps moyen avant de voir arriver le $10^{\text{ème}}$ immigrant est égal à

$$\mathbb{E}[S_{10}] = \frac{10}{\lambda} = 10 \text{ jours.}$$

Si maintenant on regarde la probabilité que le temps d'attente entre le $10^{\text{ème}}$ immigrant et le $11^{\text{ème}}$ immigrant soit supérieur à 2 jours :

$$\mathbb{P}(T_{11} > 2) = e^{-2\lambda} = e^{-2} \simeq 0.133.$$

3.4.3 Deuxième définition du processus de Poisson

Une autre façon de définir un processus de Poisson est de le définir à partir de ses instants de sauts :

Définition 3.26. Définition du processus de Poisson à l'aide des temps inter-arrivées
On considère $(T_n)_{n\geqslant 1}$ une suite de variables indépendantes, de loi exponentielle de paramètre λ. On définit $S_0 = 0$ et on pose

$$S_n = T_1 + T_2 + \ldots + T_n.$$

Alors le processus $(N(t))_{t\geqslant 0}$ définit par

$$N(t) = \sum_{n=1}^{\infty} \mathbb{1}_{S_n \leqslant t} = \max\{n \geqslant 0 : S_n \leqslant t\}$$

est un processus de Poisson d'intensité λ.

Les deux définitions sont équivalentes.

3.5 Propriétés du processus de Poisson

3.5.1 Somme de deux processus de Poisson indépendants

On considère deux processus de Poisson indépendants et on cherche la loi de la somme de ces deux processus. On sait déjà que la somme de variables de Poisson indépendante suit une loi de Poisson. Il en est de même pour les processus de Poisson

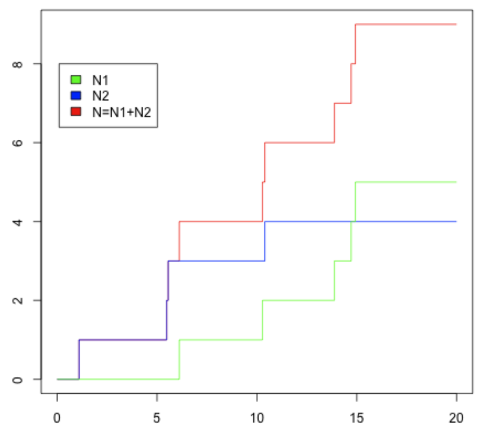
Proposition 3.27. Somme de deux processus de Poisson
Considérons deux processus de Poisson $(N^1(t))_{t\geqslant 0}$ et $(N^2(t))_{t\geqslant 0}$ indépendants d'intensité respective λ_1 et λ_2. Alors le processus $N(t) = N^1(t) + N^2(t)$ est un processus de Poisson d'intensité $\lambda_1 + \lambda_2$.
La probabilité que le premier processus saute avant le second est égale à

$$\mathbb{P}(T_1^1 < T_1^2) = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

où T_1^1 est le premier instant de saut de N^1 et T_1^2 le premier instant de saut de N^2.

La probabilité que les deux processus sautent en même temps est nulle car leurs temps de saut sont indépendants et suivent des lois gammas (loi continue !).

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012



Preuve. On considère le processus $N(t) = N^1(t) + N^2(t)$. On a $N(0) = 0$. Par ailleurs, si $t, s > 0$, alors

$$N(t+s) - N(s) = \left(N^1(t+s) - N^1(s)\right) + \left(N^2(t+s) - N^2(s)\right).$$

N^1 est un processus de Poisson, par conséquent $N^1(t+s) - N^1(s)$ suit la loi $\mathcal{P}(\lambda_1 t)$ et est indépendant de $N^1(s)$. De même $N^2(t+s) - N^2(s)$ suit la loi $\mathcal{P}(\lambda_2 t)$ et est indépendant de $N^2(s)$. Les processus N^1 et N^2 étant indépendants, obtient que (voir la proposition 3.11) $N(t+s) - N(s)$ suit une loi de Poisson $\mathcal{P}(\lambda_1 t + \lambda_2 t)$ et est indépendant de $N(s) = N^1(s) + N^2(s)$.

Par ailleurs, le premier instant de saut de N a lieu dès que le premier processus entre N^1 et N^2 saute, par conséquent N saute pour la première fois en $T_1 = \min(T_1^1, T_1^2)$. Les instants T_1^1 et T_1^2 étant indépendants, on utilise alors les résultats sur le minimum de variables exponentielles indépendantes (voir la proposition 3.5 de ce chapitre) pour conclure. □

Exemple 3.28. *Considérons une compagnie d'assurance. Les sinistres pour un assuré arrivent selon un processus de Poisson de paramètre λ. La compagnie a n clients. On suppose que les assurés ont des comportements indépendants. Alors le nombre de sinistres que doit gérer la compagnie d'assurance est un processus de Poisson d'intensité $n\lambda$.*
Supposons par exemple que le nombre moyen de sinistre par individu sur une année est de 0.01. On peut modéliser les instants d'arrivée des sinistres par un processus de Poisson d'intensité $\lambda = 0.01/365$ par jour. On considère que la compagnie d'assurance a 5000 clients. Alors le premier sinistre de l'année pour la compagnie arrive à l'instant T_1 qui suit une loi exponentielle de paramètre $5000\lambda = 50/365$. En moyenne le premier sinistre que devra prendre en charge la compagnie arrive au bout de $365/50 \simeq 7,3$ jours.

Exemple 3.29. *On considère une compagnie d'assurance qui a une branche assurance automobile et une branche assurance habitation. On peut modéliser les instants d'arrivée des sinistres à l'aide de processus de Poisson indépendants pour chacune des branches, d'intensité λ_a pour la branche automobile et d'intensité λ_h pour la branche habitation. Au final, pour la compagnie les instants d'arrivée d'un sinistre (quel qu'il soit) sont les temps d'arrivée d'un processus de Poisson d'intensité $\lambda_a + \lambda_h$.*

3.5.2 Décomposition d'un processus de Poisson

On va maintenant voir que si on décompose un processus de Poisson selon des classes, on obtient alors plusieurs processus de Poisson.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

On considère un processus de Poisson $(N(t))_{t\geqslant 0}$ de paramètre λ permettant de comptabiliser des événements liés à une population. Les individus de cette population peuvent être soit du type I soit du type II. Par exemple : soit de sexe masculin, soit de sexe féminin si on regarde les visiteurs dans un musée, ou soit des sinistres “habitation” soit des sinistres “automobile” si on considère une compagnie d'assurance. On suppose que la proportion d'individus de type I est égale à p (la proportion d'individus de type II est par conséquent $1-p$).
On note $N^1(t)$ le nombre de sauts de type I intervenus dans l'intervalle de temps $[0,t]$ et $N^2(t)$ le nombre de sauts de type II intervenus dans l'intervalle $[0,t]$.
On a par conséquent $N(t) = N^1(t) + N^2(t)$

Proposition 3.30. Décomposition d'un processus de Poisson selon des classes
Soit $(N(t))_{t\geqslant 0}$ un processus de Poisson de paramètre λ permettant de comptabiliser une population divisée en deux classes. La proportion d'individus dans la première classe est égale à p et la proportion d'individus dans la seconde classes est $1-p$.
Les processus $(N^1(t))_{t\geqslant 0}$ et $(N^2(t))_{t\geqslant 0}$ obtenu en séparant les sauts par rapport à chaque classe sont des processus de Poisson indépendants d'intensité respective $p\lambda$ et $(1-p)\lambda$.

Cette proposition se généralise facilement lorsque qu'on découpe la population en k sous groupes qui sont distribués selon les proportions $p_1, \ldots, p_k$ ($\sum_{i=1}^k p_i = 1$).
Preuve. Montrons que N^1 et N^2 sont des processus de Poisson au sens de la première définition. On a de façon évidente que $N^1(0) = N^2(0) = 0$, car $N(0) = 0$.
Cherchons la loi jointe de $(N^1(t), N^2(t))$: soit $n, k \in \mathbb{N}$ et $t \geqslant 0$ fixés, on a

$$\begin{aligned}
\mathbb{P}(N^1(t) = n, N^2(t) = k) &= \mathbb{P}(N(t) = n+k \text{ dont } n \text{ sauts sont de type I et } k \text{ de type II}) \\
&= \binom{n+k}{n} p^n (1-p)^k \frac{\lambda^{n+k}}{(n+k)!} e^{-\lambda} \\
&= \frac{(p\lambda)^n}{(n)!} e^{-p\lambda} \times \frac{((1-p)\lambda)^k}{k!} e^{-(1-p)\lambda}.
\end{aligned}$$

Par conséquent $N^1(t)$ et $N^2(t)$ sont indépendants et suivent respectivement les lois de Poisson $\mathcal{P}(p\lambda)$ et $\mathcal{P}((1-p)\lambda)$.
Étudions les accroissements de N^1 (le raisonnement est identique pour N^2). Soient $t, s \geqslant 0$, $N^1(t+s) - N^1(s)$ correspond au nombre de sauts du type I du processus N. Comme N a des accroissements indépendants, les sauts de type I de N sur $(s, t+s]$ sont indépendants de tous les sauts intervenus avant l'instant s, et donc indépendant de ceux de type I avant l'instant s. Donc N^1 a des accroissements indépendants. On montre que les accroissements sont stationnaires de la même façon que précédemment en calculant $\mathbb{P}(N^1(t+s) - N^1(s) = n, N^2(t+s) - N^2(s) = k)$ pour tout $s, t \geqslant 0$ et $k, n \in \mathbb{N}$. ☐

Exemple 3.31. *On considère une compagnie d'assurance qui s'occupe d'assurance habitation et assurance automobile. On suppose que sa proportion de contrat automobile est égal à $3/4$. On suppose que les sinistres arrivent selon un processus de Poisson d'intensité 10 par mois. Le probabilité de ne pas avoir de sinistre habitation pendant 3 mois peut être calculé de la manière suivante.*
Le nombre de sinistre habitation est un processus de Poisson de paramètre $1/4 \times 10 = 2.5$ par mois. Par conséquent la probabilité de ne pas avoir de sinistre habitation pendant trois mois est égale à

$$\mathbb{P}(\textit{pas de sinistre habitation pendant 3 mois}) = \mathbb{P}(T_1^h > 3) = e^{-3\times 2.5} \simeq 5.53.10^{-4}.$$

où T_1^h représente le premier instant où un sinistre habitation intervient.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Exemple 3.32. *Supposons que vous vendez aux enchères un objet sur un site internet payant. Les propositions arrivent selon un processus de Poisson d'intensité λ.*
A chaque proposition, soit vous acceptez l'offre et la vente est finie, soit vous la refusez et vous attendez une nouvelle proposition.
Le site internet sur lequel vous avez déposé votre offre est payant. On suppose que ça vous coûte c \$ par unité de temps jusqu'à ce que vous vendiez votre objet.
Vous souhaitez avoir un gain maximal. Votre gain correspond au prix de vente de votre objet moins le coût d'utilisation du site internet.
On suppose que vous vous fixez un prix minimal y pour votre objet et vous acceptez la première offre qui est supérieur à y. Quelle est la meilleure valeur pour y ?
On suppose que les propositions suivent une loi uniforme sur $[0, 500]$. Votre prix minimal étant y, la probabilité d'avoir une proposition X supérieure à y est

$$\mathbb{P}(X \geqslant y) = \frac{500 - y}{500}.$$

Par conséquent la proportion d'offre supérieur à y arrivent selon un processus de Poisson d'intensité $\lambda(500 - y)/500$ et le temps d'attente T pour avoir une telle offre suit une loi exponentielle de paramètre $\lambda(500 - y)/500$.
Le gain moyen sachant qu'on ne considère que des offres supérieures à y est par conséquent :

$$\begin{aligned} G(y) &= \mathbb{E}[\textit{montant de l'offre accepté}] - c\mathbb{E}[\textit{le temps de vente}] \\ &= \mathbb{E}[X|X > y] - c\mathbb{E}[T] \\ &= \int_0^\infty x f_{X|X\geqslant y}(x)dx - c\frac{500}{\lambda(500-y)} \\ &= \int_y^{500} \frac{x}{500-y}dx - \frac{500c}{\lambda(500-y)} = \frac{500+y}{2} - \frac{500c}{\lambda(500-y)} \end{aligned}$$

On veut trouver la meilleure valeur de y afin d'optimiser ce gain. Calculons donc la dérivée du gain :

$$G'(y) = \frac{1}{2} - \frac{500c}{\lambda(500-y)^2}.$$

La dérivée est nulle si et seulement si

$$y = 500 - \sqrt{1000c/\lambda}.$$

Par conséquent, si $\lambda = 0.05$ par minute (soit 3 offres par heure) et $c = 0.1$ \$ par minute, on trouve $y \simeq 455$ \$.
Si jamais le coût du site internet est trop élevé et que les offres n'arrivent pas assez rapidement (par exemple $c = 3$ \$ et $\lambda = 0.01$) , il n'y a pas de solution optimale comprise entre 0 et 500. Il est alors optimal d'accepter la première offre venue.

Exemple 3.33. *Considérons le système de Bonus-Malus en assurance automobile. On suppose qu'il y a juste K niveaux de prime. La niveau 0 correspondant à la prime la plus basse et plus le conducteur a eu d'accident au cours de la précédente année plus son niveau de prime augmente. La prime d'assurance payée chaque année par un conducteur est une chaîne de Markov de matrice de transition $(P_{i,j})_{0\leqslant i,j\leqslant K}$. On considère l'ensemble des assurés de la compagnie d'assurance et on suppose que chacun se comporte de façon indépendante des autres.*

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

On suppose que initialement le nombre d'assurés dans chacun des niveaux $0, 1, \ldots, K$ de primes est distribué selon la loi de Poisson de paramètre respectif $\lambda_0, \lambda_1, \ldots, \lambda_K$. On souhaite connaître la distribution des assurés à l'instant n.
Pour i fixé, on note $N_j(i)$ le nombre d'assurés qui sont passés du niveau i à l'instant initial au niveau j à l'instant n. Chaque personne initialement dans le niveau i se comporte de façon indépendante et a une probabilité $P^n_{i,j}$ d'être au niveau j à l'instant n. Par conséquent le nombre $N_j(i)$ d'individus passant du niveau i au niveau j suit une loi de Poisson de paramètre $\lambda_i P^n_{i,j}$.
Le nombre d'individu dans l'état j à l'instant n est donc $\sum_{i=0}^K N_j(i)$. Une somme de variables de Poisson indépendantes suit encore une loi de Poisson de paramètre la somme des paramètre. On en déduit que le nombre d'individu dans l'état j à l'instant n suit la loi de Poisson de paramètre $\sum_{i=0}^K \lambda_i P^n_{i,j}$.

3.5.3 Loi conditionnelle en fonction du nombre de sauts

Soit $(N(t))_{t\geqslant 0}$ un processus de Poisson d'intensité λ.
On le considère sur l'intervalle $[0,t]$. Sur cet intervalle, le nombre de saut suit une loi de Poisson de paramètre λt. Connaissant le nombre de sauts, que peut-on dire de la loi des instants de sauts.
Par exemple, sachant qu'il n'y a eu qu'un saut sur l'intervalle de temps $[0,t]$, quelle est la loi du premier instant de saut T_1. Cet instant est forcément inférieur à t. On a pour $s \leqslant t$

$$\begin{aligned}
\mathbb{P}(T_1 < s | N(t) = 1) &= \frac{\mathbb{P}(T_1 < s, N(t) = 1)}{\mathbb{P}(N(t) = 1)} = \frac{\mathbb{P}(\text{avoir 1 saut sur } [0,s) \text{ et aucun saut sur } [s,t])}{\lambda t e^{-\lambda t}} \\
&= \frac{\mathbb{P}(\text{avoir 1 saut sur } [0,s))\mathbb{P}(\text{avoir aucun saut sur } [s,t])}{\lambda t e^{-\lambda t}} \text{ par indépendance des accroissements} \\
&= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)})}{\lambda t e^{-\lambda t}} = \frac{s}{t}
\end{aligned}$$

Par conséquent la loi de T_1 sachant que $N(t) = 1$ est la loi uniforme sur $[0,t]$.

Définition 3.34. *Considérons $Y_1, \ldots, Y_n$ des variables aléatoires. On définit la statistique d'ordre $Y_{(1)}, \ldots, Y_{(n)}$ comme étant les variables réordonnées par ordre croissant : $Y_{(k)}$ est le $k^{\text{ème}}$ plus petite valeur parmi $Y_1, \ldots, Y_n$.*
On a $Y_{(1)} \leqslant Y_{(2)} \leqslant \ldots \leqslant Y_{(n)}$.

Théorème 3.35. *Sachant que $N(t) = k$, les instants de saut $S_1, S_2, \ldots, S_k$ ont la même loi que la statistique d'ordre d'un k-échantillon de loi uniforme sur $[0,t]$: $(U_{(1)}, \ldots, U_{(k)})$.*

Preuve. Voir Ross page 327. □

Ce résultat est très utile en simulation. En effet quand on veut simuler un processus de Poisson jusqu'à l'instant t. Le nombre de sauts sur cet intervalle de temps suit une loi de Poisson $\mathcal{P}(\lambda t)$. On simule par conséquent $N \sim \mathcal{P}(\lambda t)$, puis on simule N variables uniforme sur $[0,t]$ que l'on ordonne par ordre croissant pour avoir les instants de sauts de notre processus.
En R, cela donne le programme suivant :

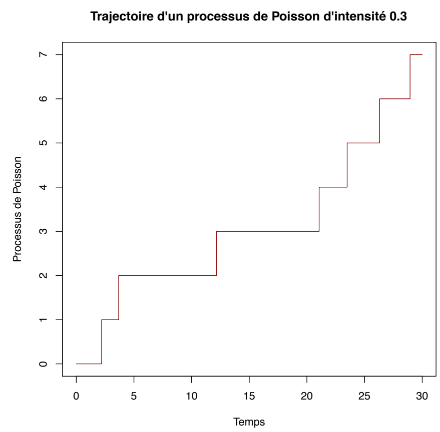
```
pdf("PP03.pdf")
# Longueur de l'intervalle de temps
T=30
# Intensite du parametre de Poisson
lambda=0.3
```


Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

```
# Simulation du nombre de sauts N sur l'intervalle [0,T]
N=rpois(1,lambda*T)
# Simulation de N variables uniforme sur [0,T] indépendantes
U=runif(N,min=0,max=T)
# Ordonnement des variables uniformes par ordre croissant
X=sort(U)

# Dessin d'une realisation du processus de Poisson sur l'intervalle [0,T]
plot(c(0,X,T),c(0:N,N),type="s",col="dark red",
xlab="Temps",ylab="Processus de Poisson",
main="Trajectoire d'un processus de Poisson d'intensite 0.3")
dev.off()
```

Voici le résultat :



3.6 Processus de Poisson généralisés

3.6.1 Processus de Poisson non homogène

Un processus de Poisson non homogène est un processus de Poisson dont le taux de saut n'est plus constant mais dépend du temps : $\lambda(t)$.
Soit $(N(t))_{t\geqslant 0}$ un processus de Poisson non homogène d'intensité $(\lambda(t))_{t\geqslant 0}$. On a

— $N(0) = 0$.
— N a des accroissements indépendants, mais ils ne sont plus stationnaires.
— Si on regarde les sauts sur un petit accroissement de temps, on a

$$\lim_{h\to 0}\frac{1}{h}\mathbb{P}(N(t+h)-N(t)=1)=\lambda(t) \qquad \lim_{h\to 0}\frac{1}{h}\mathbb{P}(N(t+h)-N(t)\geqslant 2)=0.$$

Si $(N^1(t))_{t\geqslant 0}$ et $(N^2(t))_{t\geqslant 0}$ sont deux processus de Poisson non homogènes d'intensité respective $(\lambda^1(t))_{t\geqslant 0}$ et $(\lambda^2(t))_{t\geqslant 0}$, alors $(N(t))_{t\geqslant 0}$ avec $N(t) = N^1(t) + N^2(t)$ est un processus de Poisson non homogène d'intensité $(\lambda^1(t) + \lambda^2(t))_{t\geqslant 0}$.

Exemple 3.36. *Un magasin ouvre à 8h00. De 8h00 à 10h00, les clients arrivent selon un processus de Poisson d'intensité* 4 *par heure. De 10h00 à midi, l'intensité est de* 8 *par heure. De midi à 14h00, l'intensité croît de* 8 *par heure (à midi) jusqu'à* 10 *par heure (à 14h00). De 14h00 à 17h00, l'intensité*

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

décroît de 10 par heure (à 14h00) jusqu'à 2 par heure (à 17h00). On souhaite connaitre la moyenne du nombre de clients qui entrent dans le magasin ce jour.
On note N_i les processus de Poisson associé à chaque période. Le nombre $N(t)$ de clients entrant avant l'instant t est un processus de Poisson non homogène. On souhaite évaluer

$$\mathbb{E}[N(17)-N(8)] = \mathbb{E}[N_1(10)-N_1(8)] + \mathbb{E}[N_2(12)-N_2(10)] + \mathbb{E}[N_3(14)-N_3(12)] + \mathbb{E}[N_2(17)-N_2(14)].$$

N^1 est un PP d'intensité 4, donc $N_1(10)-N_1(8) \sim \mathcal{P}(4\times 2)$ et $\mathbb{E}[N_1(10)-N_1(8)] = 8$.
N^2 est un PP d'intensité 8, donc $N_2(12)-N_2(10) \sim \mathcal{P}(8\times 2)$ et $\mathbb{E}[N_2(12)-N_2(10)] = 16$.
N^3 est un PP d'intensité $8+(t-12) = t-4$. L'intensité correspondant au taux de sauts par unité de temps on a donc

$$\mathbb{E}[N_3(14)-N_3(12)] = \int_{12}^{14}(t-4)dt = 18.$$

N^4 est un PP d'intensité $10-\frac{8}{3}(t-14) = \frac{142}{3}-\frac{8}{3}t$ et donc

$$\mathbb{E}[N_3(14)-N_3(12)] = \frac{1}{3}\int_{14}^{17}(142-8t)dt = 18.$$

Le nombre moyen de clients sur une journée est $8+16+18+18 = 60$.

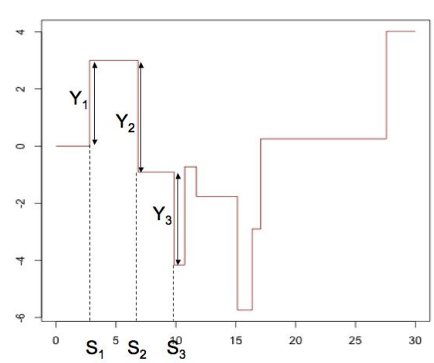
3.6.2 Processus de Poisson composé

On revient sur notre problème initial, on souhaite estimer la probabilité de ruine d'une compagnie d'assurance. Les instants des sinistres peuvent être modélisé à l'aide d'un processus de Poisson mais la hauteur des sauts dépend du montant que la compagnie devra débourser pour couvrir le sinistre, il n'est pas conséquent pas constant égal à 1.
On va donc maintenant introduire la notion de processus de Poisson composé où l'amplitude des sauts est aléatoire.

Définition 3.37. *Un processus $(X_t)_{t\geqslant 0}$ est appelé processus de Poisson composé s'il peut s'écrire pour $t \geqslant 0$*

$$X_t = \sum_{i=1}^{N(t)} Y_i$$

où $(N(t))_{t\geqslant 0}$ de paramètre λ et $(Y_i)_{i\geqslant 1}$ est une suite de variables indépendantes, de même loi que Y et indépendantes de $(N(t))_{t\geqslant 0}$.



Une trajectoire du processus de Poisson composé d'intensité $\lambda = 0.3$
dont l'amplitude des sauts suit la loi uniforme sur $[-5,5]$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Propriété 3.38. *On a*

$$\mathbb{E}[X_t] = \lambda t \mathbb{E}[Y] \quad Var(X_t) = \lambda t \mathbb{E}[Y^2].$$

Preuve. Pour calculer ces quantités on va conditionner par rapport à $N(t)$. Comme les Y_i sont indépendantes, de même loi et indépendants de N, en utilisant les résultats sur les sommes aléatoires (voir Proposition 3.15), on a

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[\sum_{i=1}^{N(t)} Y_i | N(t)]] = \mathbb{E}[N(t)]\mathbb{E}[Y] = \lambda t \mathbb{E}[Y]$$

et

$$Var(X(t)) = \mathbb{E}[Y]^2 Var(N(t)) + Var(Y)\mathbb{E}[N(t)] = \lambda t(\mathbb{E}[Y]^2 + Var(Y)) = \lambda t \mathbb{E}[Y^2].$$

□

Exemple 3.39. *On suppose que des familles immigrent au Québec selon un processus de Poisson de taux* 2 *par semaine. Si le nombre de personne dans chaque famille est indépendant et suit la loi*

$$p(0) = 1/6,\ p(1) = 1/3,\ p(2) = 1/3,\ p(4) = 1/6.$$

Par conséquent, le nombre moyen de personnes immigrant au Québec sur une période de 50 *semaines est*

$$\mathbb{E}[X(50)] = 2 \times 50 \times \mathbb{E}[Y] = 250,$$

où $\mathbb{E}[Y] = 1.1/6 + 2.1/3 + 3.1/3 + 4.1/6 = 5/2$ *et la variance du nombre moyen de personnes immigrant au Québec sur une période de* 5 *semaines est*

$$Var(X(50)) = 2 \times 50 \times \mathbb{E}[Y^2] = 2150/3,$$

car $\mathbb{E}[Y^2] = 43/6$.
Maintenant si on souhaite évaluer la probabilité qu'au moins 240 *personnes immigrent au Québec dans les prochaines* 50 *semaines. Comme* $X(50)$ *est la somme de variables indépendantes et de même loi ayant une variance finie (on imagine* $N(50)$ *grand), on va utiliser l'approximation par la loi normale (voir Théorème central limite, Chapitre 0) :*

$$\begin{aligned}
\mathbb{P}(X(50) \geqslant 240) &= \mathbb{P}\left(\frac{X(50) - 250}{\sqrt{2150/3}} \geqslant \frac{240 - 250}{\sqrt{2150/3}}\right) \\
&= 1 - \Phi(-0.3735) \qquad \textit{où } \Phi \textit{ est la fonction de répartition de la loi } \mathcal{N}(0,1) \\
&= \Phi(0.3735) = 0.646.
\end{aligned}$$

3.7 Probabilité de ruine d'une compagnie d'assurance

On considère une compagnie d'assurance qui souhaite évaluer sa probabilité de ruine en fonction du capital initial investi et des primes récoltées. On suppose qu'elle possède un capital initial c et on note R_t la réserve de la compagnie à l'instant t.
Cette compagnie d'assurance

- perçoit des cotisations de ses clients que l'on supposera mensualisées et uniformément répartie sur l'année : les recettes de la compagnie pendant un temps t sont égales à pt où p est le taux de cotisations par unité de temps.
- verse des indemnités à ses assurés sinistrés en fonction du dommage qu'ils subissent.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

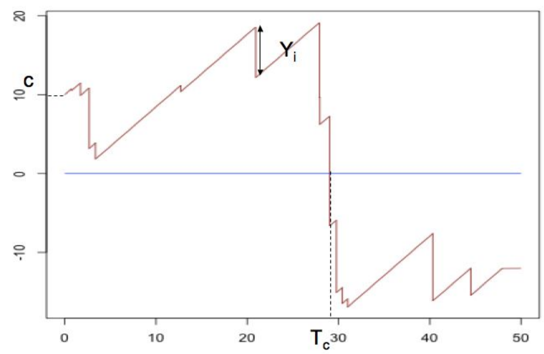
On modélise l'apparition et les coûts des sinistres de la manière suivante :

— les coûts de sinistres $(Y_k)_{k\geqslant 1}$ sont des variables indépendantes de même loi et d'espérance μ. Par facilité on supposera ici qu'ils suivent la loi $\mathcal{E}(1/\mu)$.

— les instants des sinistres sont modélisés comme les instants de sauts d'un processus de Poisson $N(t)_{t\geqslant 0}$ d'intensité λ.

Par conséquent la réserve de la compagnie d'assurance à l'instant t est :

$$R_t = c + pt - \sum_{k=1}^{N_t} Y_k.$$



Trajectoire de R_t avec $\lambda = 0.3$, $\mu = 3$, $p = 1$ et $c = 10$.

La compagnie est ruinée dès que sa réserve R_t descend sous 0. La probabilité de ruine de la compagnie est par conséquent

$$p_c = \mathbb{P}(\min_{t\geqslant 0} R_t < 0 | R_0 = c)$$

et l'instant où la ruine intervient est

$$T_c = \min\{t \geqslant 0 : R_t < 0\}.$$

La réserve moyenne et sa variance valent à l'instant t :

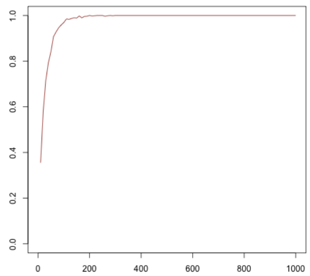
$$\mathbb{E}[R_t] = c + (p - \lambda\mu)t \quad Var(R_t) = 2\lambda\mu^2 t.$$

Si $p < \lambda\mu$, on remarque que $\mathbb{E}[R_t] \to -\infty$. Par contre, si $p = \lambda\mu$ la réserve moyenne reste constante, mais la variance $\mathbb{E}[Var(R_t)] \to +\infty$. On peut en fait montrer que dans ces deux cas que la compagnie a une probabilité 1 d'être ruinée.

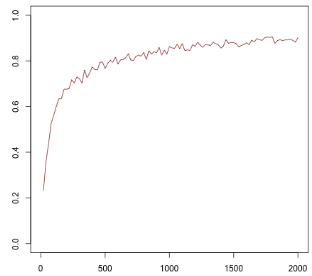
Si $p > \lambda\mu$, on a $\mathbb{E}[R_t] \to +\infty$ et $\mathbb{E}[Var(R_t)] \to +\infty$, on ne peut rien en conclure.

Voici quelques simulations de la probabilité de ruine. On a pris un capital initial égale à $c = 10$, une prime $p = 1$ par unité de temps et on a tracé en fonction du temps la probabilité de ruine avant l'instant t (calculée par simulations) dans les trois différents cas :

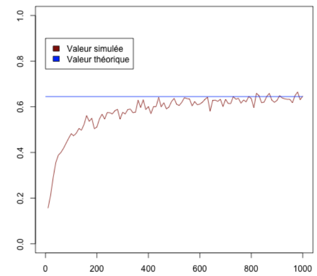
Copyright © Hélène Guérin. Université du Québec à Montréal. 2012



cas $p < \lambda\mu$
avec $\lambda = 0.5$ et $\mu = 3$



cas $p = \lambda\mu$
avec $\lambda = 0.5$ et $\mu = 2$



cas $p > \lambda\mu$
avec $\lambda = 0.3$ et $\mu = 3$.

Dans le dernier cas ($p > \lambda\mu$), on peut calculer explicitement par des techniques utilisant la propriété de Markov la probabilité de ruine de la compagnie d'assurance qui est dans ce cas strictement inférieur à 1 et égale à

$$p_c = \frac{\lambda\mu}{p} e^{-\frac{p-\lambda\mu}{p\mu}c}.$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Chapitre 4

Le Mouvement Brownien

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Table des matières

4.1 Un peu d'histoire

Le nom de mouvement Brownien vient du botaniste Robert Brown. Brown n'a pas découvert le mouvement brownien, car n'importe qui regarde dans un miscroscope peut voir le mouvement rapide et irrégulier des particules de pollen en suspension dans de l'eau. Cependant, avant lui on pensait que les particules étaient vivantes. Une autre théorie expliquait que le mouvement des particules était dû à la différence de température entre l'eau et le milieu ambiant provoquant l'évaporation de l'eau, ou qu'il était une conséquence des courants d'air. Brown (1828) réfuta ces théories et établit que les particules étaient inanimées. Il expliqua que la matière était composée de petites particules, appelées molécules actives, qui montrent un mouvement rapide et irrégulier, dont l'origine vient des particules. Puis au début des années 1900, le mouvement brownien fut caractérisé de la façon suivante :

- Le mouvement est très irrégulier, composé de translations et de rotations, la trajectoire ne semble pas avoir de tangentes.
- Deux particules semblent bouger de façon indépendantes, même si elles sont très proches.
- Le mouvement est d'autant plus actif que les particules sont petites.
- La composition et la densité des particules n'ont pas d'influence.
- Le mouvement est d'autant plus actif que le fluide n'est pas trop visqueux.
- Le mouvement est plus actif en température haute.
- Le mouvement est sans fin.

En 1905, la théorie de la physique cinétique, qui explique que le mouvement brownien des particules microscopiques est dû au bombardements des molécules du fluide, semble la plus plausible.

La mise en évidence du mouvement brownien comme processus stochastique est dû indépendamment au mathématicien français Louis Bachelier (1900) et Albert Einstein en (1905). Bachelier dans sa thèse *Théorie de la spéculation* avait pour but la modélisation de la dynamique des actifs boursiers et son application au calcul de prix d'option. Il obtient la loi du mouvement brownien à un instant donné. Il met surtout en évidence le caractère markovien du mouvement brownien : le déplacement d'une particule brownienne après un instant t ne dépend que de l'endroit où elle était à l'instant t et mais pas de comment elle est arrivée à ce point. Par ailleurs, Einstein, qui ignorait l'existence de ce débat, formula une théorie quantitative en 1905 du mouvement brownien. Einstein réussi à expliquer la nature du mouvement et donna la valeur du coefficient de diffusion (sous certaines hypothèses). La méthode d'Einstein repose sur des considérations de mécanique statistique qui le conduit à l'équation de la chaleur puis à la densité gaussienne, solution fondamentale de cette équation.

L'existence du mouvement brownien en tant que processus stochastique a été établie de façon rigoureuse par Wiener en 1923. La loi du mouvement brownien est appelée mesure de Wiener.

En finance, on utilise le mouvement brownien pour modéliser le cours d'une action. Les particules d'eau sont remplacées par les achats et ventes des traders qui font évoluer l'action.

4.2 Quelques notions sur la loi gaussienne

Une variable aléatoire X à valeurs dans $\mathbb{R}$ est dite gaussienne centrée réduite si elle suit la loi $\mathcal{N}(0,1)$ de densité

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

Il n'est pas possible de calculer explicitement la fonction de répartition, notée ϕ, de la loi gaussienne, par contre on peut estimer la queue gaussienne : pour $x \geqslant 1$,

$$\begin{aligned} 1-\phi(x) = \mathbb{P}(X \geqslant x) &= \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-t^2/2} dt \leqslant \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} t e^{-t^2/2} dt \\ &\leqslant \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \end{aligned}$$

Les queues ont une décroissance très rapide.
La fonction génératrice de la loi gaussienne est donnée par : pour $z \in \mathbb{R}$

$$\mathbb{E}[e^{zX}] = \exp(z^2/2).$$

Pour $m \in \mathbb{R}$ and $\sigma > 0$, $Y = m + \sigma X$ suit une loi gaussienne $\mathcal{N}(m, \sigma)$, de densité $\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$ et de fonction génératrice $\mathbb{E}[e^{zY}] = \exp(zm + \sigma^2 z^2/2)$.

Remarque 4.1. *Si X et Y sont des variables indépendantes de loi respectives $\mathcal{N}(m, \sigma^2)$ et $\mathcal{N}(m', \sigma'^2)$, alors $X+Y$ suit la loi $\mathcal{N}(m+m', \sigma^2+\sigma'^2)$.*

Démonstration. Il suffit d'utiliser les fonctions génératrices. Par indépendance, on a

$$\begin{aligned} \mathbb{E}[e^{z(X+Y}] &= \mathbb{E}[e^{zX}]\mathbb{E}[e^{zY}] = \exp(zm + \sigma^2 z^2/2)\exp(zm' + \sigma'^2 z^2/2) \\ &= \exp(z(m+m') + (\sigma^2+\sigma'^2)z^2/2). \end{aligned}$$

□

Propriété 4.2. *Soient (X,Y) un couple gaussien, i.e. pour tout $a, b \in \mathbb{R}$, $aX+bY$ suit une loi gaussienne. Alors X et Y sont indépendantes si et seulement si $Cov(X,Y) = 0$.*

Théorème 4.3. (Théorème limite central)
Soit X une v.a. avec $\mathbb{E}[X^2] < \infty$. On pose $m = \mathbb{E}[X]$ et $\sigma^2 = Var(X)$.
Soient X_1, X_2, ..., X_n des variables indépendantes et de même loi que X. Alors

$$\frac{\sqrt{n}}{\sigma}\left(\frac{X_1+X_2+...+X_n}{n} - m\right) \underset{n\to+\infty}{\overset{en\ loi}{\longrightarrow}} \mathcal{N}(0,1).$$

Donc $\forall a, b \in \mathbb{R}$, $a < b$,

$$\mathbb{P}(a \leqslant \frac{1}{\sigma\sqrt{n}}\left(\sum_{i=1}^n X_i - m \times n\right) \leqslant b) \underset{n\to+\infty}{\longrightarrow} \phi(b) - \phi(a),$$

où ϕ est la fonction de répartition de la loi $\mathcal{N}(0,1)$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

4.3 Vecteurs Gaussiens

À RAJOUTER...

4.4 Loi et espérance conditionnelle dans le cas continu

(Ross p.102)
Considérons une couple (X, Y) dont la loi est à densité $f_{X,Y}(x, y)$. Alors les variables X et Y sont des variables à densité de densité respective

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dy \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y)dx.$$

La loi conditionnelle de X sachant que $Y = y$ est bien définie pour les y tels que $f_Y(y) > 0$ par :

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

L'espérance conditionnelle de X sachant que $Y = y$ est alors

$$\mathbb{E}[X|Y = y] = \int_{-\infty}^{+\infty} x f_{X|Y}(x|y)dx.$$

Exemple 4.4. *Considérons un couple de variables aléatoires (X, Y) de loi jointe*

$$f(x, y) = 6xy(2 - x - y)\mathbb{1}_{(x,y)\in]0,1[^2}.$$

Alors la densité conditionnelle de X sachant $Y = y$, avec $y \in]0, 1[$ est pour $x \in]0, 1[$

$$f_{X|Y}(x|y) = \frac{6xy(2 - x - y)}{\int_0^1 6xy(2 - x - y)dx} = \frac{6x(2 - x - y)}{4 - 3y}.$$

Par conséquent,

$$\mathbb{E}[X|Y = y] = \int_0^1 \frac{6x^2(2 - x - y)}{4 - 3y}dx = \frac{5 - 4y}{8 - 6y}.$$

4.5 Le mouvement brownien

4.5.1 Construction du mouvement brownien

On va considérer une marche aléatoire symétrique avec des pas de plus en plus petits sur des intervalles de temps de plus en plus courts.
Sur un petit intervalle de temps Δt, on va faire un petit pas de taille Δx soit vers le haut avec probabilité $1/2$ soit vers le bas avec probabilité $1/2$: au lieu de considérer notre buveur classique, on considère une fourmi qui a bu une goutte d'alcool et qui est surexcitée ! On regarde la position X_t de notre fourmi à l'instant t :

$$X_t = \Delta x(Z_1 + Z_2 + \ldots + Z_{[t/\Delta t]})$$

où $[t/\Delta t]$ est le plus grand entier inférieur à $t/\Delta t$ (appelé *partie entière*) et où les Z_i représentent les différents pas (indépendants entre eux) :

$$Z_i = \begin{cases} +1 & \text{si le } i^{\text{ème}} \text{ pas est vers le haut (avec proba } 1/2), \\ -1 & \text{si le } i^{\text{ème}} \text{ pas est vers le bas (avec proba } 1/2), \end{cases}$$

Les accroissements de X_t sont indépendants et stationnaires : pour $t, s \geqslant 0$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

$X_{t+s} - X_t = \Delta x(Z_{[t/\Delta t]+1} + \ldots + Z_{[(t+s)/\Delta t]})$ est indépendant de X_t par indépendance des Z_i et de même loi que X_s.

On remarque que

$$\mathbb{E}[X_t] = 0 \text{ et } Var(X_t) = (\Delta x)^2 \left[\frac{t}{\Delta t}\right].$$

On prend $\Delta x = \sqrt{\Delta t}$ et on fait converger Δt vers 0 . On a

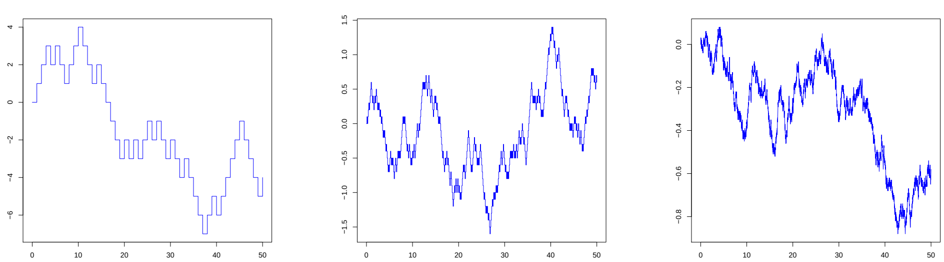
$$\mathbb{E}[X_t] = 0 \text{ et } Var(X_t) \to t.$$

En effet, $\frac{t}{\Delta t} - 1 \leqslant \left[\frac{t}{\Delta t}\right] \leqslant \frac{t}{\Delta t}$. D'après le Théorème central limite (imaginer que $\Delta t = 1/n$), on obtient

$$X_t \xrightarrow[n\to\infty]{Loi} \mathcal{N}(0,t).$$

En fait, le processus limite X_t, quand Δt tend vers 0, est un mouvement brownien.

Trajectoires de la marche aléatoire pour différentes valeurs de Δt



$\Delta t = \Delta x = 1$ $\qquad$ $\Delta t = \Delta x = 0.1$ $\qquad$ $\Delta t = \Delta x = 0.01$

4.5.2 Définitions

Définition 4.5. — *Un processus X est dit* **stationnaires**, *si $\forall t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$ et $s \geqslant 0$,*

$$(X_{t_1}, \ldots, X_{t_n}) \overset{Loi}{=} (X_{t_1+s}, \ldots, X_{t_n+s}).$$

- *Un processus X est dit* **à accroissement indépendants**, *si $\forall t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$, les variables $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.*
- *Un processus X est dit* **continu** *si la fonction $t \mapsto X_t$ est continue.*
- *Un processus X est dit* **gaussien** *si pour tout $n \geqslant 0$, $t_1 \leqslant t_2 \leqslant \ldots \leqslant t_n$, $a_1, \ldots, a_n \in \mathbb{R}$, la variable*

$$a_1 X_{t_1} + a_2 X_{t_2} + \ldots + a_n X_{t_n}$$

suit une loi gaussienne.

Définition 4.6. *Soit $B = (B_t, t \geqslant 0)$ une famille de variables indexées par le temps. On dit que B est un* mouvement brownien *si c'est un processus continu à accroissements indépendants tel que*

(i) $B_0 = 0$,

(ii) $\forall t \geqslant 0$, $B_t \sim \mathcal{N}(0,t)$.

La densité de B_t est

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

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Conséquence 4.7. Cherchons la loi d'un accroissement $B_{t+s} - B_t$, pour $t, s \geqslant 0$:
On remarque que $B_{t+s} = B_t + B_{t+s} - B_t$, avec $B_{t+s} \sim \mathcal{N}(0, t+s)$ et $B_t \sim \mathcal{N}(0, t)$, donc par indépendance, on a $B_{t+s} - B_t \sim \mathcal{N}(0, s)$.
Le mouvement brownien a des accroissements indépendants et stationnaires de loi gaussienne, On en déduit que le mouvement brownien est un processus gaussien..

4.6 Propriétés du mouvement brownien

Propriété 4.8. Covariance du brownien.
Soit B un mouvement Brownien, on a pour tout $t, s \geqslant 0$

$$\mathbb{E}[B_t B_s] = \min(t, s).$$

Démonstration. Si $s \leqslant t$, $\mathbb{E}[B_t B_s] = \mathbb{E}[(B_s + (B_t - B_s))B_s] = s + 0$. De même, si $t \leqslant s$, $\mathbb{E}[B_t B_s] = t$. □

Définition 4.9. *(équivalente)*
Un mouvement brownien est un processus continu gaussien centré de covariance $\min(t, s)$.

Dans la littérature, $\min(t, s)$ est souvent noté $t \wedge s$.

Propriété 4.10. Symétrie
Soit B un mouvement brownien. Alors $-B$ est un mouvement brownien.

Démonstration. Il est facile de vérifier que $-B$ est un processus sont gaussien, centré, continu et à accroissements indépendants. □

Propriété 4.11. Propriété de Markov.
Soit $t \geqslant 0$ fixé. On pose $B'_s = B_{t+s}$, $\forall s \geqslant 0$ ("valeurs prises après l'instant t").
Conditionnellement à $B_t = x$, le processus B' est indépendant du passé, i.e. $(B_u)_{0 \leqslant u \leqslant t}$, et a la même loi qu'un mouvement brownien partant de x : "$x + B$".

Considérons maintenant le temps d'atteinte d'un certain niveau.

Proposition 4.12. Temps d'atteinte
Soit $a > 0$, on considère

$$T_a = \inf\{t \geqslant 0 : B_t = a\}$$

le premier instant où le mouvement brownien atteint le niveau a. On note S_t le maximum de B avant l'instant t :

$$S_t = \max_{0 \leqslant s \leqslant t} B_s.$$

On a alors, pour tout $a > 0$

$$\mathbb{P}(S_t \geqslant a) = \mathbb{P}(T_a \leqslant t) = 2\mathbb{P}(B_t \geqslant a) = \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^{\infty} e^{-x^2/2} dx.$$

Si $a < 0$, par symétrie du mouvement brownien, la loi de T_a est la même que celle de T_{-a}. Par conséquent, pour tout $a \in \mathbb{R}$, on a

$$\mathbb{P}(T_a \leqslant t) = \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-x^2/2} dx.$$

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Démonstration. On remarque que $\{B_t \geqslant a\} \subset \{T_a \leqslant t\}$. En effet, le mouvement brownien part de 0 et est continu, donc si à l'instant t il est au dessus de a c'est qu'il a forcément atteint le niveau a avant l'instant t. Par conséquent

$$\mathbb{P}(B_t \geqslant a) = \mathbb{P}(B_t \geqslant a, T_a \leqslant t) = \mathbb{P}(B_t \geqslant a | T_a \leqslant t)\mathbb{P}(T_a \leqslant t).$$

Par ailleurs, par symétrie on a la même probabilité d'être au dessus de a ou en dessous à l'instant t, donc

$$\mathbb{P}(B_t \geqslant a | T_a \leqslant t) = \mathbb{P}(B_t \leqslant a | T_a \leqslant t) = 1/2.$$

On en déduit que

$$\mathbb{P}(T_a \leqslant t) = 2\mathbb{P}(B_t \geqslant a) = \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^{\infty} e^{-x^2/2} dx.$$

□

Soit $0 \leqslant a \leqslant b$, par continuité du mouvement brownien, on a

$$T_a \leqslant T_b.$$

Considérons maitnenant deux niveaux $a > 0$ et $b > 0$, on peut se demander avec quelle probabilité le brownien atteindra le niveau a avant le niveau $-b$, i.e.

$$\mathbb{P}(T_a < T_{-b}) = ?$$

Pour résoudre ce problème on utilise la construction du mouvement brownien à l'aide de la marche aléatoire. Partant de 0, en utilisant les résultats sur le problème de la ruine du joueur, la probabilité qu'une marche aléatoire faisant des pas de taille Δx atteigne a avant d'atteindre $-b$ vaut $\frac{b\Delta x}{(a+b)\Delta x} = \frac{b}{(a+b)}$. Par conséquent, à la limite, la probabilité est la même pour le mouvement brownien et $\mathbb{P}(T_a < T_{-b}) = \frac{b}{(a+b)}$. Une autre façon de répondre à la question est d'introduire

$$T_{a,-b} = \inf\{t \geqslant 0 : B_t \in \{a, -b\}\}.$$

On remarque que $T_{a,-b} = \min(T_a, T_{-b})$. Par ailleurs,

$$\begin{aligned}
\mathbb{E}[B_{T_{a,-b}}] &= a\mathbb{P}(T_a < T_{-b}) - b\mathbb{P}(T_a > T_{-b}) = (a+b)\mathbb{P}(T_a < T_{-b}) - b \\
\mathbb{E}[B_{T_{a,-b}}] &= \mathbb{E}[\mathbb{E}[B_{T_{a,-b}} | T_{a,-b}]] = 0
\end{aligned}$$

D'où

$$\mathbb{P}(T_a < T_{-b}) = \frac{b}{(a+b)}.$$

De même, on a

$$\begin{aligned}
\mathbb{E}[B^2_{T_{a,-b}}] &= a^2\mathbb{P}(T_a < T_{-b}) + b^2\mathbb{P}(T_a > T_{-b}) = a^2\frac{b}{(a+b)} + b^2\frac{a}{(a+b)} \\
\mathbb{E}[B^2_{T_{a,-b}}] &= \mathbb{E}[\mathbb{E}[B^2_{T_{a,-b}} | T_{a,-b}]] = \mathbb{E}[T_{a,-b}]
\end{aligned}$$

D'où

$$\mathbb{E}[T_{a,-b}] = \mathbb{E}[\min(T_a, T_{-b})] = ab.$$

PROPRIÉTÉ À AJOUTER : Loi conditionelle de $B_s | B_t = x$ avec $s \leqslant t$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

4.7 Généralisation du mouvement Brownien

4.7.1 Mouvement brownien avec dérive

Définition 4.13. *Un processus* $(Y_t)_{t\geqslant 0}$ *est appelé* **mouvement brownien avec dérive** μ **et variance** σ^2 *si*

(i) $Y_0 = 0$

(ii) Y *a des accroissements indépendants et stationnaires*

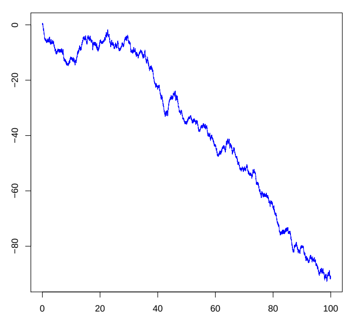
(iii) Y_t *suit la loi* $\mathcal{N}(\mu t, \sigma^2 t)$.

Si B est un mouvement brownien *standard*, alors

$$Y_t = \mu t + \sigma B_t$$

est un mouvement brownien avec dérive μ et variance σ^2.

Trajectoire d'un mouvement brownien avec dérive -1 et variance 4



4.7.2 Mouvement brownien géométrique

Définition 4.14. *Soit* $(Y_t)_{t\geqslant 0}$ *est appelé mouvement brownien avec dérive* μ *et variance* σ^2*, alors le processus* $(X_t)_{t\geqslant 0}$ *défini par*

$$X_t = e^{Y_t}$$

est appelé **mouvement brownien géométrique**.

Le mouvement brownien géométrique est très utilisé en finance. En effet, supposons que X_t soit le cours d'une action à l'instant t. Il est usuel de supposer que les rendements $Y_t = X_t/X_{t-1}$, $t \geqslant 1$ sont indépendants et identiquement distribués. Par conséquent

$$X_t = Y_t Y_{t-1} \dots Y_1 X_0$$

et

$$\ln(X_t) = \sum_{i=1}^{t} \ln(Y_i) + \ln(X_0)$$

Par le théorème central limite, en normalisant correctement $\ln(X_t)$ ressemble à un mouvement brownien avec dérive et donc X_t à un brownien géométrique.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Connaissant l'historique du processus jusqu'à l'instant s, on souhaite prédire la valeur de l'action à un instant $t > s$. Pour cela, on calcule l'espérance conditionnelle du cours à l'instant t sachant l'historique (ça nous donnera la valeur moyenne du cours à l'instant t sachant l'historique).
$(X_t)_{t\geqslant 0}$ un mouvement brownien géométrique. En utilisant la propriété d'accroissements indépendants du mouvement brownien, on a

$$\begin{aligned}\mathbb{E}[X_t|X_u, 0 \leqslant u \leqslant s] &= \mathbb{E}[e^{Y_t}|Y_u, 0 \leqslant u \leqslant s]\\ &= e^{Y_s}\mathbb{E}[e^{Y_t-Y_s}|Y_u, 0 \leqslant u \leqslant s]\\ &= X_s\mathbb{E}[e^{Y_t-Y_s}]\end{aligned}$$

Par ailleurs, $Y_t - Y_s = \mu(t-s) + \sigma(B_t - B_s)$ suit une loi gaussienne $\mathcal{N}(\mu(t-s), \sigma^2(t-s))$. D'après les rappels sur les variables gaussiennes on a donc

$$\mathbb{E}[X_t|X_u, 0 \leqslant u \leqslant s] = X_s e^{(\mu+\sigma^2/2)(t-s)}.$$

Définition 4.15. *Une variable aléatoire X est dite suivre la loi log-normale de paramètre μ et σ^2 si la variable $Y = \ln(X)$ suit le loi normale $\mathcal{N}(\mu, \sigma^2)$.*
La densité de X est alors

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}}e^{\frac{-(\ln(x)-\mu)^2}{2\sigma^2}}$$

et en utilisant les résultats sur la fonction générative d'une loi normale, on a

$$\mathbb{E}[X] = e^{\mu}e^{\frac{\sigma^2}{2}} \quad Var(X) = (e^{\sigma^2}-1)e^{2\mu+\sigma^2}.$$

Par conséquent, si $(X_t)_{t\geqslant 0}$ est un mouvement brownien géométrique, pour tout $t \geqslant 0$, la densité de X_t est

$$f_t(x) = \frac{1}{x\sigma\sqrt{2\pi t}}e^{\frac{-(\ln(x)-\mu)^2}{2t\sigma^2}}.$$

4.8 Application en finance : modèle de Black-Scholes

4.8.1 Un exemple de pricing d'option

Avoir une somme ν d'argent à l'instant t n'est pas la même chose que d'avoir cette somme à l'instant présent et de la garder jusqu'à l'instant t. En effet, si on avait dès maintenant cette somme, on pourrait la placer et donc grâce aux taux d'intérêts sa valeur à l'instant t ne serait plus la même (actif financier sans risque). Par conséquent on introduit un facteur d'actualisation α et une somme d'argent qui vaut ν à l'instant t, vaut finalement $e^{-\alpha t}\nu$ à l'instant présent. On parle de valeur actualisée. Tous les montants que l'on considérera par la suite seront actualisés.
Supposons le prix actuel d'une action (actif financier risqué) est de 100\$ et que l'on sait qu'après une unité de temps cette action vaudra soit 200\$ soit 50\$. *Les montants sont données en prix actualisé ce qui signifie qu'à l'instant 1 le prix réel sera soit $200e^{\alpha}$\$ soit $50e^{\alpha}$\$.*
Supposons que pour tout y, au prix de cy, vous pouvez vous procurer à l'instant 0 l'option d'acheter à l'instant 1 y parts de l'action au prix de 150\$ par part. Par conséquent, si vous achetez cette option et que le montant de l'action monte à 200\$, vous exercerez votre option à l'instant 1 et vous réaliserez un gain de $200 - 150 = 50$\$ pour chacune de vos y parts. Si par ailleurs, le prix à l'instant 1 est 50\$, alors vous n'exercerez pas votre option. Quelque soit le montant de l'action au temps 1, vous aurez de toute façon dépensé le coût de l'option.
Enfin, à l'instant 0 vous pouvez acheter x part de l'action au prix de $100x$.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Les quantités x et y peuvent être positives (achat) ou négatives (vente).
Le but est de trouver la bonne valeur c du coût d'une option. Si par exemple, $c \neq 50/3$, alors il existe une stratégie d'achat/vente qui assure un gain positif.
En effet, supposons qu'à l'instant 0 on achète x actions et y options, alors à l'instant 1 notre portefeuille vaut soit $200x + 50y$ si le prix de l'action monte à 200\$, soit $50x$ si le prix de l'action descend à 50\$ (dans les deux cas nous avons déboursé cy\$ pour le prix des y options). Choisissons y tel que la valeur de votre portefeuille à l'instant 1 ne dépende pas de la valeur de l'action à l'instant 1, i.e. tel que

$$200x + 50y = 50x \Longrightarrow y = -3x.$$

Ceci signifie, dans le cas où x est positif, qu'à l'instant initial pour une action achetée, il y a eu 3 options qui ont été vendues. A l'inverse, dans le cas où x est négatif, pour une action vendue à l'instant initial, il y a eu 3 options achetées.
Dans le cas où $y = -3x$, la valeur de votre portefeuille à l'instant 1 est

$$\text{valeur } = 50x.$$

Par ailleurs, vous aviez "acheté" à l'instant initial, x action et $-3x$ option, votre dépense est donc de

$$\text{dépense } = 100x - 3xc$$

Votre gain final est par conséquent de

$$\text{gain } = 50x - (100x - 3xc) = x(3c - 50).$$

Dans le cas où le prix d'une option est de $c = 50/3$, votre gain est nul quelque soit votre stratégie, par contre vous vous garantissez des gains positifs en prenant $x > 0$ si $c > 50/3$ et en prenant $x < 0$ si $c < 50/3$.
Être dans une situation où on est sûr de gagner est une situation dite "avec arbitrage". Le seul prix de l'option qui ne permet pas un arbitrage est $c = 50/3$.

4.8.2 Théorème d'arbitrage

Considérons un marché contenant d actifs financiers.
On modélise l'évolution aléatoire des actifs à l'aide par exemple d'un dé truqué à m faces : les seuls résultats possibles sont $S = \{1, 2, \ldots, m\}$. Ce dé sera lancé à chaque pas de temps et en fonction du résultat on obtient la nouvelle valeur de chaque actifs financiers.
En effet, si on achète x parts de l'actif i à l'instant 0, alors le gain est $xr_i(j)$ à l'instant 1 si le dé tombe sur la face j. Le montant x peut-être positif (il s'agit alors d'un achat) ou négatif (il s'agit alors d'une vente).
Une stratégie de portefeuille est un vecteur $\mathtt{x} = (x_1, \ldots, x_d)$ où x_i est le nombre de parts associés à l'actif i. Si le dé tombe sur la face j, alors à l'instant 1 le gain est

$$\text{gain } = \sum_{i=1}^{d} x_i r_i(j).$$

Le théorème suivant établit l'existence d'une probabilité $p = (p_1, \ldots, p_m)$ sur S (i.e. sur l'évolution des prix des actifs) telle que le gain moyen est toujours nul et telle que si l'expérience n'est pas distribuée selon cette probabilité p il existe alors une stratégie gagnante à tous les coups.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Théorème 4.16. Théorème d'arbitrage
De façon exclusive, soit il existe une probabilité p sur S telle que

$$\sum_{i=1}^{m} p_j r_i(j) = 0 \quad \textit{pour tout } i = 1, \ldots, d$$

soit il existe une stratégie $\mathtt{x} = (x_1, \ldots, x_d)$ *telle que*

$$\sum_{i=1}^{d} x_i r_i(j) > 0 \quad \textit{pour tout } j = 1, \ldots, m$$

En d'autres termes, si X est le résultat du dé qu'on lance, soit il existe une distribution p de X telle que

$$\mathbb{E}_p[r_i(X)] = 0 \quad \text{pour tout } i = 1, \ldots, d$$

soit il existe une stratégie gagnante à coup sur, quelque soit la valeur de X.

Exemple 4.17. *Un modèle envisageable est d'associer à l'actif i le gain g_i si le résultat de l'expérience est i et -1 sinon :*

$$r_i(j) = \begin{cases} g_i & \textit{si } j = i \\ -1 & \textit{sinon} \end{cases}$$

Pour qu'il n'y ait pas d'arbitrage, il faut qu'il existe une probabilité p sur $S = \{1, \ldots, d\}$ telle que

$$0 = \mathbb{E}_p[r_i(X)] = g_i p_i - (1 - p_i)$$

soit $p_i = \frac{1}{1+g_i}$ et $\sum_{i=1}^d p_i = 1$. Par conséquent, il n'y a pas d'arbitrage si

$$\sum_{i=1}^{d} \frac{1}{1+g_i} = 1$$

sinon il existe une stratégie permettant de gagner à coup sur.
Par exemple, si $d = 3$ et $g_1 = 1, g_2 = 2, g_3 = 3$, il existe une stratégie gagnante à coup sur car $1/2 + 1/3 + 1/4 > 1$. Considérons la stratégie suivante

- *vendre une part de l'actif 1 : $x_1 = -1$ (on gagne 1 si le résultat n'est pas 1 et on perd 1 si le résultat est 1),*
- *vendre 0.7 part de l'actif 2 : $x_2 = -0.7$ (on gagne 0.7 si le résultat n'est pas 2 et on perd 1.4 si le résultat est 2),*
- *vendre 0.5 de l'actif 3 : $x_3 = -0.5$ (on gagne 0.5 si le résultat n'est pas 3 et on perd 1.5 si le résultat est 1).*

Par conséquent, si le résultat de l'expérience est 1, notre gain est

$$\textit{gain } = -1 + 0.7 + 0.5 = 0.2,$$

si le résultat de l'expérience est 2 notre gain est

$$\textit{gain } = 1 - 1.4 + 0.5 = 0.1,$$

si le résultat de l'expérience est 3 notre gain est

$$\textit{gain } = 1 + 0.7 - 1.5 = 0.2.$$

Dans tous les cas, on est gagnant.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

4.8.3 La formule de Black-Scholes

Supposons que le prix d'une action à l'instant présent est $X_0 = x_0$. On note X_t son prix à l'instant t. On se fixe un instant final T. En notant α le facteur d'actualisation, la valeur actualisée de l'action à l'instant t est $e^{-\alpha t}X_t$.
On regarde le prix d'une action à travers le temps X_t, $0 \leqslant t \leqslant T$.
À chaque instant $s < t$, on regarde le processus et on choisit d'acheter (ou vendre) des actions au prix X_s pour ensuite de revendre (ou racheter) ces actions à l'instant t au prix X_t.
On suppose par ailleurs que l'on peut acheter une option à l'instant 0 de coût c par unité qui permet d'acheter des actions au prix K à l'instant t.
On cherche le montant c de façon qu'il ne puisse pas y avoir arbitrage, i.e. on cherche c de telle façon qu'il existe une probabilité sur l'ensemble des prix possibles de façon que le gain moyen soit nul.
Soit $\mathbb{P}$ une probabilité sur l'ensemble des prix. Considérons la stratégie qui consiste à acheter (ou vendre) une action à l'instant s dans l'intention de la vendre (ou acheter) à l'instant t, $0 \leqslant s < t \leqslant T$. Le prix actualisé de l'action à l'instant s est $X_s e^{-\alpha s}$ et celui de l'action à l'instant t est $X_t e^{-\alpha t}$. Par conséquent, pour avoir un gain moyen nul, la loi $\mathbb{P}$ de $(X_t)_{0\leqslant t\leqslant T}$ doit vérifier

$$\mathbb{E}_{\mathbb{P}}[X_t e^{-\alpha t} | X_u, 0 \leqslant u \leqslant s] = X_s e^{-\alpha s}.$$

Considérons maintenant la stratégie d'acheter une option. Supposons qu'une option nous donne le droit d'acheter une action à l'instant t au prix K. A l'instant t la valeur de cette option est

$$\text{valeur de l'option à l'instant } t = (X_t - K)^+ = \begin{cases} X_t - K & \text{si } X_t \geqslant K \\ 0 & \text{si } X_t < K \end{cases}$$

($x^+ = \max(x, 0)$). La valeur actualisée de l'option est par conséquent

$$(X_t - K)^+ e^{-\alpha t}$$

Si le coût de cette option à l'instant 0 est c, pour avoir un gain moyen nul il faut donc que

$$c = \mathbb{E}_{\mathbb{P}}[(X_t - K)^+ e^{-\alpha t}]$$

S'il existe une probabilité $\mathbb{P}$ qui vérifie (4.8.3) et si on choisit c tel que (4.8.3) il n'y a pas d'arbitrage possible.
Cherchons une probabilité $\mathbb{P}$ qui vérifie (4.8.3). Imaginons que

$$X_t = x_0 e^{Y_t}$$

où $(Y_t)_{t\geqslant 0}$ est un mouvement brownien avec coefficient de dérive μ et variance σ^2. On a alors

$$\mathbb{E}[X_t | X_u, 0 \leqslant u \leqslant s] = X_s e^{(\mu+\sigma^2/2)(t-s)}$$

Par conséquent, si on choisit μ et σ^2 tels que

$$\mu + \sigma^2/2 = \alpha$$

alors l'équation (4.8.3) est satisfaite.
La loi $\mathbb{P}$ d'un mouvement brownien géométrique avec $\mu + \sigma^2/2 = \alpha$ satisfait l'équation (4.8.3). Il suffit alors de prendre c tel que

$$c = \mathbb{E}_{\mathbb{P}}[(X_t - K)^+ e^{-\alpha t}]$$

pour qu'aucun arbitrage ne soit possible. On trouve

$$c = x_0 \phi(\sigma\sqrt{t} + b) - K e^{-\alpha t} \phi(b)$$

avec $b = (\alpha t - \sigma^2 t/2 - \ln(K/x_0))/(\sigma\sqrt{t})$. Cette formule est appelée *formule de Black et Scholes.*

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012

Démonstration.

$$\begin{aligned}
c = \mathbb{E}_{\mathbb{P}}[(X_t - K)^+ e^{-\alpha t}] &= \int_{\ln(K/x_0)}^{\infty} (x_0 e^y - K) e^{-\alpha t} \frac{1}{\sqrt{2\pi t}\sigma} e^{\frac{-(y-\mu t)^2}{2\sigma^2 t}} dy \\
&= x_0 e^{-\alpha t} e^{\frac{(\mu t + \sigma^2 t)^2 - (\mu t)^2}{2\sigma^2 t}} \frac{1}{\sqrt{2\pi t}\sigma} \int_{\ln(K/x_0)}^{\infty} e^{\frac{-(y-\mu t - \sigma^2 t)^2}{2\sigma^2 t}} dy - K e^{-\alpha t} \frac{1}{\sqrt{2\pi t}\sigma} \int_{\ln(K/x_0)}^{\infty} e^{\frac{-(y-\mu t)^2}{2\sigma^2 t}} dy \\
&= x_0 \frac{1}{\sqrt{2\pi}} \int_{(\ln(K/x_0) - \alpha t - \sigma^2 t/2)/\sigma\sqrt{t}}^{\infty} e^{\frac{-y^2}{2}} dy - K e^{-\alpha t} \frac{1}{\sqrt{2\pi}} \int_{(\ln(K/x_0) - \alpha t + \sigma^2 t/2)/\sigma\sqrt{t}}^{\infty} e^{\frac{-y^2}{2}} dy
\end{aligned}$$

□

Remarque 4.18. *Un processus $(M_t)_{t\geqslant 0}$ qui vérifie la relation (4.8.3), i.e. pour tout valeur de $s \leqslant t$ on a*

$$\mathbb{E}_{\mathbb{P}}[M_t | M_u, 0 \leqslant u \leqslant s] = M_s$$

est appelé martingale. Le mouvement brownien géométrique n'est pas le seul processus à résoudre notre problème d'arbitrage.
En fait, en considérant $(M_t)_{t\geqslant 0}$ une martingale et

$$c = \mathbb{E}_{\mathbb{P}}[(M_t - K e^{-\alpha t})^+]$$

il n'y a pas d'arbitrage.

Copyright © Hélène Guérin. Université du Québec à Montréal. 2012