

# Simulation Numérique des lois de probabilité

## Principe général

### Définition d'une simulation numérique

Simuler numériquement une loi de probabilité  $P$  consiste à donner les valeurs  $x_1, x_2, x_3, \dots, x_N$  obtenues lors d'une réalisation de  $X_1, X_2, \dots, X_N$  un échantillon de  $X$  qui suit la loi  $P$ .

On notera  $R$ , une variable aléatoire qui suit la loi uniforme sur l'intervalle  $(0, 1)$ ; puis à partir de  $R$ , on forme une variable aléatoire  $X$  qui suit la loi  $P$  demandée.

Ainsi à partir d'une simulation  $r_1, r_2, r_3, \dots, r_N$  ( en principe connue ) de  $R$ ; on obtient les  $x_1, x_2, x_3, \dots, x_N$  cherchées, qui simulent la loi  $P$ .

## Simulation de la loi uniforme sur ( 0 , 1 )

Quelquefois ,il est utilisé des propriétés de phénomènes physiques pour obtenir les  $r_1 , r_2 , r_3 \dots , r_N$  qui simulent la loi uniforme sur ( 0 , 1 ) .

Nous allons, ici, simplement décrire, une méthode numérique pour obtenir ces valeurs  $r_1 , r_2 , r_3 \dots , r_N$  .

Considérons la formule de récurrence  $x_{n+1} = a x_n$  modulo  $M$

où  $M$  est un nombre entier le plus grand possible pour la machine considérée, (pour retarder le moment où on obtient un entier déjà obtenu) on prend  $M = 2^q$  pour une machine à  $q$  bits

$a$  est un nombre entier premier avec  $M$  (pour éviter d'obtenir 0 ) ; on prend  $a = 3^k$  avec  $k$  tel que,  $a$  soit peut différent de  $M/2$  .

Le premier nombre entier  $x_0$  est pris tel que

$0 < x_0 < M$  ( il est initialisé avec l'horloge, ou un autre moyen physique) .

On obtient ainsi une suite  $x_n$

de nombres entiers  $0 < x_n < M$

Alors les  $r_n$  cherchés sont les réels  $r_n = x_n / M$  .

On a ainsi obtenu une suite de réels t.q  $0 < r_n < 1$  .

En réalisant des tests statistiques (par exemple , test du Chi-deux , que l'on verra dans la partie :Statistique) on est amené à considérer que cette suite simule correctement la loi uniforme sur ( 0 , 1 ) .

## **Méthode de la transformation inverse**

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

**Théorème : Soit  $f$  la densité d'une loi ; et  $F$  la fonction de répartition de cette loi ;  $F$  est une bijection croissante , de l'intérieur du support de  $f$ , sur  $] 0 , 1 [$**

**donc il existe une fonction réciproque  $F^{-1}$**

**alors**

**la variable aléatoire  $X = F^{-1}(R)$  suit la loi de densité  $f$  (où  $R$ , conformément à nos notations dans V1, suit la loi uniforme sur  $] 0 , 1 [$ )**

Preuve : la fonction de répartition de  $X$  est :

$$\begin{aligned} \text{Prob}( X \leq x ) &= \text{Prob}( F^{-1}(R) \leq x ) = \text{Prob}( R \leq F(x) ) \\ &= F(x) \end{aligned}$$

## Exemples de simulation

### loi uniforme sur un intervalle ( a , b )

C'est une application de la méthode de la transformation inverse

La loi à simuler est la loi de densité  $f = \text{constante} = 1/(b-a)$  sur cet intervalle ( a , b )

sa fonction de répartition F vaut

$$F(x) = x/(b-a) - a/(b-a) =$$

$$=(x-a)/(b-a) \text{ sur l'intervalle ( a , b )}$$

donc la fonction réciproque  $F^{-1}$  vaut  $F^{-1}(r) = a + r(b-a)$  ;

Donc la suite qui simule la loi uniforme sur ( a , b ) est

$$x_n = a + r_n(b-a)$$

### Simulation de la loi exponentielle

C'est une application de la méthode de la transformation inverse

La loi exponentielle , de début 0 et de paramètre a admet pour densité la fonction  $f(x) = a \exp(-ax)$  pour x positif

$F(x) = 1 - \exp(-ax)$  est une bijection croissante de l'ensemble des réels  $> 0$  sur l'intervalle ] 0 , 1 [

sa fonction réciproque vaut  $F^{-1}(r) = -1/a \ln(1-r)$  ;

La suite qui simule la loi exponentielle voulue est

$$x_n = -1/a \ln(1-r_n)$$

Remarquer que l'on peut remplacer

la suite  $(1-r_n)$  par la suite  $r_n$  .

## Simulation de la loi de Poisson de paramètre a

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

**Théorème :**  $Y_1, Y_2, \dots, Y_n \dots$  étant un échantillon de modèle  $Y$  qui suit la loi exponentielle de début 0 et de paramètre a ;

**X** sera le plus grand entier t.q  $\sum_{i=1}^{i=X} Y_i \leq 1$

**Alors X** suit la loi de Poisson de paramètre a .

Principe de la preuve :

$$\text{Prob}(X = k) = \text{Prob}\left(\sum_{i=1}^{i=k} Y_i \leq 1 \text{ et } \sum_{i=1}^{i=k+1} Y_i > 1\right) =$$

$$\int_A a^{k+1} \exp(-a \sum_{i=1}^{i=k+1} y_i) dy_1 \dots dy_{k+1}$$

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

(car on rappelle que la densité de la loi exponentielle de début 0 et de paramètre a est  $f(y_i) = a \exp(-a y_i)$  pour  $y_i > 0$  )

on calcule cette intégrale =

$$a^{k+1} \int_{y_i > 0 \text{ et } \sum_{i=1}^{i=k} y_i < 1} \left\{ \int_{1 - \sum_{i=1}^{i=k} y_i}^{\infty} \exp(-a y_{k+1}) dy_{k+1} \exp(-a \sum_{i=1}^{i=k} y_i) dy_1 \dots dy_k \right.$$

..dy<sub>k</sub>

$$\text{or } \left\{ \int_{1-\sum_{i=1}^{i=k} y_i}^{\infty} \exp(-a y_{k+1}) dy_{k+1} \right\} =$$

$$1/a \exp(-a (1 - \sum_{i=1}^{i=k} y_i))$$

$$\text{donc l'intégrale vaut } a^k \int_{y_i > 0 \text{ et } \sum_{i=1}^{i=k} y_i < 1} \exp(-a) dy_1 \dots dy_k$$

$$= a^k \exp(-a) \int_{y_i > 0 \text{ et } \sum_{i=1}^{i=k} y_i < 1} dy_1 \dots dy_k = a^k \exp(-a) 1/k!$$

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

### Simulation de la loi normale réduite

**Théorème : Soit  $\rho$  qui suit la loi exponentielle de début 0 et de paramètre 1, et  $\theta$  qui suit la loi uniforme sur  $(0, 2\pi)$**

**$\rho$  et  $\theta$  étant indépendantes**

**alors**

**en posant  $r = (2\rho)^{1/2}$  ;**

**$X = r \cos(\theta)$  et  $Y = r \sin(\theta)$**

**on obtient  $X$  et  $Y$  qui sont indépendantes et suivent la loi Gauss(0, 1) .**

Preuve :

le couple  $(\theta, \rho)$  admet  $f(\theta, \rho) = 1 / 2\pi \exp(-\rho)$  comme densité sur  $(0, 2\pi) / \mathbb{R}^+$ .

On considère la bijection

$\phi$

$$(\theta, \rho) \rightarrow (\theta, r = (2\rho)^{1/2})$$

$$D1 \rightarrow D2$$

$\xi$

$$(\theta, r) \rightarrow (x = r \cos(\theta), y = r \sin(\theta))$$

$$D2 \rightarrow D3$$

$$\text{Prob}((X, Y) \in D3) = \text{Prob}((\theta, \rho) \in D1)$$

$$= \text{intégrale de } 1/2\pi \exp(-\rho) \text{ sur } D1$$

(on effectue le changement de variable  $\phi$ )

$$= \text{intégrale de } 1/2\pi r \exp(-1/2 r^2) \text{ sur } D2$$

(on effectue le changement de variable  $\xi$ )

$$= \text{intégrale de } 1/2\pi \exp(-1/2 (x^2 + y^2)) \text{ sur } D3$$

donc la densité du couple  $(X, Y)$  est

$$1/2\pi \exp(-1/2 (x^2 + y^2)) =$$

$$(1/2\pi)^{1/2} \exp(-1/2 x^2) (1/2\pi)^{1/2} \exp(-1/2 y^2)$$

=

Conclusion ; d'après tout ce que l'on vient de considérer ;

pour simuler la loi de Gauss réduite :

on simule le couple  $(\theta, \rho)$  en prenant par exemple la suite

$$(\theta_n, \rho_n) = (2\pi r_n, -\ln(r'_n))$$

(remarquer, qu'on ne prend pas les mêmes  $r_n$  pour former  $\theta_n$  et  $\rho_n$  car, sinon on n'aurait pas l'indépendance de  $\theta$  et  $\rho$ )

puis on forme les suites

$$x_n = (2\rho_n)^{1/2} \cos(\theta_n) \quad \text{et} \quad y_n = (2\rho_n)^{1/2} \sin(\theta_n)$$

On déduit une simulation de la loi Gauss( $m, \sigma$ ), à partir d'une simulation de la loi réduite, du fait que :

si  $X$  suit la loi réduite,

alors  $Y = m + \sigma X$  suit la loi Gauss( $m, \sigma$ )

## Processus de Markov , discrets

### Définitions

On considère une variable aléatoire qui évolue par étapes successives , on note  $X_n$  cette variable aléatoire à l'étape numéro  $n$  ;

on dit qu'on a un processus aléatoire discret en temps .  
Si chaque  $X_n$  est une variable aléatoire discrète , on dit que le processus aléatoire est discret en valeurs .

Un processus aléatoire discret en temps et en valeurs sera simplement appelé processus aléatoire discret .

Nous noterons  $e_k$   $k = 0, 1, \dots$  les valeurs possibles des variables aléatoires  $X_n$

**Un processus aléatoire discret sera appelé processus aléatoire discret de Markov ssi il vérifie :**

**pour tout  $m < n$**

$$\text{Prob}(X_n=e_n, X_{m+1}=e_{m+1} \mid X_m=e_m, X_{m-1}=e_{m-1}, X_0=e_0)$$

$$\text{Prob}(X_n=e_n, X_{m+1}=e_{m+1} \mid X_m=e_m)$$

**Les états futurs du processus aléatoire ne dépendent que du présent , mais pas du passé .**

Exemple : les jeux qui se jouent en plusieurs parties (sans mémoire du passé) sont des processus discrets de Markov .

On appelle **transitions** : les  $\text{Prob}( X_k = e_k \mid X_{k-1} = e_{k-1} )$

Un processus aléatoire discret de Markov est appelé **chaîne de Markov** ssi il est invariant par translation dans le temps ; c.à.d :

$\text{Prob}( X_n = e' \mid X_0 = e ) = \text{Prob}( X_{n+k} = e' \mid X_k = e )$   
ceci pour tout  $k, n$  et toutes valeurs  $e$  et  $e'$  .

Exemples :

Les jeux qui se jouent en plusieurs parties (sans mémoire du passé) et sans changer la règle du jeu au cours du temps sont des chaînes de Markov .

En considérant que les valeurs  $e_k$  sont les sommets d'un graphe, une chaîne de Markov correspond à un parcours aléatoire du graphe, ceci constitue le modèle le plus général de chaîne de Markov à un nombre fini de valeurs.

## **Matrice de transition**

Dans le cas où les valeurs possibles sont en nombre fini on notera simplement celles ci :  $1, 2, \dots, N$   
(au lieu de :  $e_1, e_2, \dots, e_N$  )

**on notera :  $G_{i,j}$  les transitions  $\text{Prob}(X_1 = j \mid X_0 = i)$   
la matrice  $G$  de coefficients  $G_{i,j}$  est appelée matrice de transition .**

Remarque : une matrice de transition  $G$  vérifie :

$$\sum_j G_{i,j} = 1, \text{ pour tout } i$$

Propriété : une matrice de transition  $G$  admet 1 comme valeur propre , ses autres valeurs propres sont de module  $\leq 1$

L'évolution du processus est décrite par la matrice  $G^{(k)}$  dont les coefficients sont les  $G^{(k)}_{i,j} = \text{Prob}(X_k = j \mid X_0 = i)$

**Théorème : pour tout  $k$  ; on a :  $G^{(k)} = G^k$**

Preuve par récurrence sur  $k$  ; vraie pour  $k=1$  ;  
supposons cette propriété vérifiée jusqu'à l'indice  $k$  :

$$\text{alors } G^{(k+1)}_{i,j} = \text{Prob}(X_{k+1} = j \mid X_0 = i) = \sum_n \text{Prob}(X_{k+1} = j \mid X_k = n) \text{Prob}(X_k = n \mid X_0 = i)$$

(c'est la formule des probabilités totales)

mais  $\text{Prob}(X_{k+1} = j \mid X_k = n) = G_{n,j}$  "chaîne de Markov"  
et  $\text{Prob}(X_k = n \mid X_0 = i) = G^{(k)}_{i,n} = G^k_{i,n}$

d'après l'hypothèse de récurrence.

$$\text{on a donc } G^{(k+1)}_{i,j} = \sum_n G^k_{i,n} G_{n,j} = G^{k+1}_{i,j}$$

### **Exemple de la marche aléatoire**

A chaque étape un 'promeneur' peut se déplacer soit vers la gauche (avec une probabilité  $p$ ) soit vers la droite (avec une probabilité  $q$ ) soit rester à la même place (avec une probabilité  $= 1 - p - q$ ) ; il a un nombre fini de positions possibles ; c.à.d. qu'il y a des Bords (absorbants ou réfléchissants, selon la nature du problème) .

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

Considérons le problème avec 4 positions possibles et Bords absorbants avec  $p = q = 1/2$ .

Alors la matrice  $G$  de transition est

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

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

$$G^n = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 \end{pmatrix} + (1/2)^{n+1} \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + 2(-1/2)^{n+1} \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1/3 & 1 & -1 & 1/3 \\ 1/3 & -1 & 1 & -1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

## **Etude asymptotique**

On veut étudier l'évolution du processus lorsque  $n$  le nombre d'étapes tend vers l'infini .

Dans l'exemple précédent  $G^n$  tend vers la matrice

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2/3 & 0 & 0 & 1/3 \\ 1/3 & 0 & 0 & 2/3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Réfléchir sur l'interprétation de ce résultat .

## Ergodicité

Notons  $q^k$  le vecteur dont la composante n°i est  $q_i^k = \text{Prob}(X_k = i)$  c'est le vecteur d'état de la chaîne de Markov  $X_k$   $k = 1, \dots, N$

Remarque :  $\|q^k\|_1 = 1$  pour tout k

Notons P la matrice transposée de la matrice de transition G .

**Propriété :  $P q^k = q^{k+1}$  .**

Preuve :  $(P q^k)_i = \sum_j P_{i,j} q_j^k = \sum_j G_{j,i} q_j^k = \sum_j \text{Prob}(X_{k+1} = i \mid X_k = j) \text{Prob}(X_k = j) = \text{Prob}(X_{k+1} = i) = q_i^{k+1}$   
( car pour une chaîne de Markov  $G_{j,i} = \text{Prob}(X_{k+1} = i \mid X_k = j)$  )

**Corollaire :  $q^k = P^k q^0$  pour tout k**

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

On appelle vecteur d'état asymptotique : la limite (si elle existe) des vecteurs d'états  $q^k$ , lorsque  $k$  tend vers l'infini

on la notera  $q^\infty$ .

Remarque : Comme, pour tout  $k$  on a :  $\|q^k\|_1 = 1$   
il s'ensuit que  $\|q^\infty\|_1 = 1$

Définition : On dit que la chaîne de Markov est Ergodique ssi  $q^\infty$  existe et ne dépend pas de l'état initial  $q^0$ .

**Théorème :** Si 1 est valeur propre simple de G (la matrice de transition) et les autres valeurs propres sont de module  $< 1$  alors la chaîne de Markov est Ergodique .

**Preuve :** G et P ont les mêmes valeurs propres .

Supposons (pour simplifier les calculs) que P soit diagonalisable (sinon , il faut remplacer la diagonalisation , par un réduction de Jordan ; ce qui ensuite donne des calculs un peu plus techniques) .

Considérons  $( e_1 , e_2 , \dots , e_N )$  une base de vecteurs propres pour P telle que :

$e_1$  est un vecteur propre associé à la valeur propre 1 et chaque  $\| e_i \|_1 = 1$

On décompose  $q^0$  sur cette base

$$q^0 = a_1 e_1 + \sum_{i=2}^N a_i e_i$$

$$q^k = P^k q^0 = a_1 e_1 + \sum_{i=2}^N a_i (\lambda_i)^k e_i$$

Comme  $|\lambda_i| < 1$  pour tout  $i > 1$  ,on peut passer à la limite pour k tend vers  $\infty$

Cela donne  $q^\infty = a_1 e_1$  .

Comme les  $\| \cdot \|_1$  de  $q^\infty$  et de  $e_1$  valent 1, il s'ensuit que  $a_1 = 1$  ou  $a_1 = -1$

Comme les  $q_i^k$  sont positifs pour tout k et i , les  $q_i^\infty$  sont positifs pour tout i ,

donc  $q^\infty$  est **le** vecteur propre de la matrice P , unitaire pour la norme  $\| \cdot \|_1$  , associé à la valeur propre 1 qui a toutes ses composantes positives (il ne dépend donc, pas de  $q^0$  ).

### Exemple

Considérons le cas de la marche aléatoire (vue en V2d) avec 3 positions possibles, des Bords réfléchissants et  $p = q = 1/3$

$$\text{la matrice } G = \begin{pmatrix} 0 & 1 & 0 \\ 1/3 & 1/3 & 1/3 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{donc } P = \begin{pmatrix} 0 & 1/3 & 0 \\ 1 & 1/3 & 1 \\ 0 & 1/3 & 0 \end{pmatrix}$$

les valeurs propres de  $P$  sont : 1 , 0 et  $-2/3$  donc les hypothèses du théorème précédent, sont vérifiées , donc cette chaîne de Markov est Ergodique .

Le vecteur d'état asymptotique  $q^\infty$  est le vecteur propre de la matrice  $P$  , unitaire pour la norme  $\| \cdot \|_1$  , associé à la valeur propre 1 , qui a toutes ses composantes positives

c'est le vecteur de composantes :  $1/5$  ,  $3/5$  ,  $1/5$  .

Réfléchir sur l'interprétation de ce résultat.