Simulation par la méthode d'inversion

Prof. Mohamed El Merouani
http://elmerouani.jimdo.com
e-mail: m_merouani@yahoo.fr

1

Introduction:

- on suppose que l'on dispose d'un bon générateur de nombres pseudo-aléatoires et on se demande comment à partir d'une suite $(U_i)_{i \geq 1}$ de variables aléatoires indépendantes et identiquement distribuées suivant la loi uniforme sur [0, 1] construire une variable aléatoire de loi donnée, avec une attention particulière pour les lois usuelles continues et discrètes.

2

Plan

- Rappels sur la fonction de répartition
- Simulation des Lois de probabilités continues
- Simulation des Lois de probabilités discrètes
- Exemples sur Excel

3

Fonction de répartition (Rappels):

4

Fonction de répartition d'une v.a. discrète:

- La probabilité pour que la variable aléatoire X prenne une valeur inférieure ou égale à x est une fonction F(x).
- Cette fonction est appelée fonction de répartition de x.

  $F(x)=P(X\leq x) = \sum_{y\leq x} P(X=y)$
- La fonction F(x) est une fonction en escalier, croissante de 0 à 1.

5

Exemple de fonction de répartition:

- Pour l'expérience de lancement d'une pièce de monnaie, on a la loi de probabilité de X est résumée par le tableau suivant:

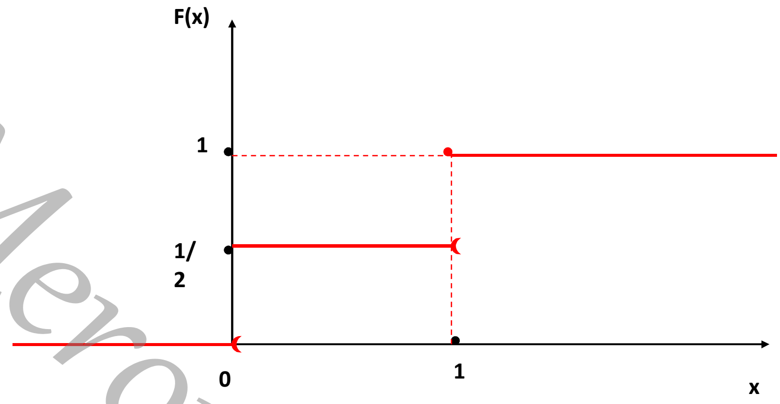
x_i	0	1	Σp_i
p_i	1/2	1/2	1

- Sa fonction de répartition sera:

$$F(x)=\begin{cases} 0 & \text{si } x<0 \\ 1/2 & \text{si } 0\leq x<1 \\ 1 & \text{si } x\geq 1 \end{cases}$$

6

Représentation graphique de F(x):



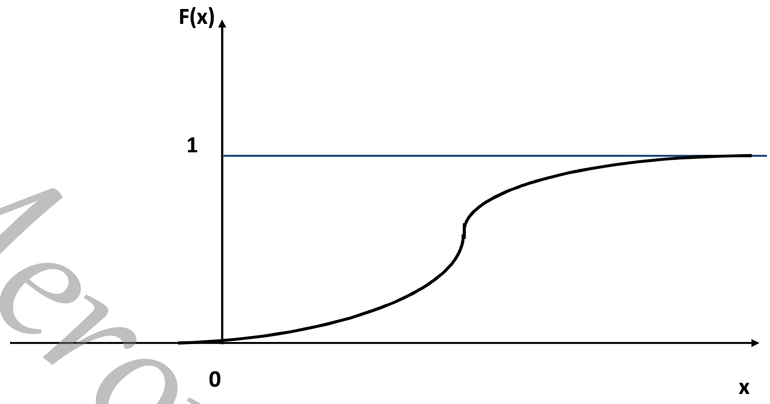
7

Fonction de répartition d'une v.a. continue:

- On définit la fonction de répartition F d'une v. a. continue X de la même manière que pour une v.a. discrète, c'est-à-dire $F(x)=P(X\leq x)$.
- La fonction continue $F(x)=\int_{-\infty}^{x} f(t)dt$ est croissante de 0 à 1 lorsque x varie de -∞ à +∞.
- Par conséquent, une v.a. continue est une v.a. dont la fonction de répartition est continue.
- La fonction de densité f d'une v.a. continue X est la dérivée de da fonction de répartition F, c'est-à-dire $F'(x)=f(x)$.

8

Représentation graphique de F(x) continue:



9

Simulation des lois de probabilités continues

10

Méthode d'inversion (Théorème):

Proposition:

Supposons que la v.a. X a pour fonction de répartition F **continue** et **strictement croissante**, toujours que $0<F(x)<1$.

Soit U v.a.$\rightarrow \mathcal{U}(0,1)$.

Alors, la v.a. $F^{-1}(U)$ a pour fonction de répartition F.

11

Méthode d'inversion (dém):

Démonstration:

Soit G la fonction de répartition de $F^{-1}(U)$.

Alors: $G(x)=P(F^{-1}(U)\leq x)$

$=P(F(F^{-1}(U))\leq F(x))$. (par la monotonie de F)

$=P(U\leq F(x))=F(x)$. (car $U\rightarrow \mathcal{U}(0,1)$)

12

Méthode d'inversion (algorithme):

- Cette méthode suggère que pour générer des échantillons d'une v.a. X pour laquelle F^{-1} est connue, on peut générer des nombres aléatoires U uniformes sur (0,1) et faire $X=F^{-1}(U)$.
- Nous avons alors l'algorithme d'inversion suivant:

> Générer $U \rightarrow \mathcal{U}(0,1)$
> Faire $X=F^{-1}(U)$
> Sortir X

13

Méthode d'inversion (Remarques):

- Une condition minime pour l'application de cette méthode est de connaître la forme explicite de F^{-1}.
- Cela est vérifié pour plusieurs lois de probabilités, comme l'uniforme, l'exponentielle, de Weibull, de Cauchy,…
- Remarquons qu'une telle condition n'est pas suffisante, par exemple, pour la loi beta, il est possible théoriquement de la simuler par inversion, mais elle peut résulter très couteuse.
- Parfois, nous disposons d'une bonne approximation de F^{-1}, d'où on peut utiliser la méthode par approximation.

14

Méthode d'inversion-Exemples:

1. Simulation d'une v.a. $\mathcal{U}(a,b)$:

Sa fonction de répartition est:

$$F(x)=\begin{cases} 0 & si\ x<a \\ \dfrac{x-a}{b-a} & si\ a\leq x<b \\ 1 & si\ x\geq b \end{cases}$$

Dans le schémas général, il suffit de faire:

$$X=F^{-1}(U)=a+(b-a)U$$

15

Méthode d'inversion-Exemples:

2. Simulation d'une v.a. de Weibull $\mathcal{W}(\alpha,1)$:

Sa fonction de répartition est:

$$F(x)=\begin{cases} 0 & si\ x<0 \\ 1-\exp\left(-x^{\alpha}\right) & si\ x\geq 0 \end{cases}$$

Nous faisons $X=F^{-1}(U)=\left[-Ln(1-U)\right]^{\frac{1}{\alpha}}$

ou bien $X=\left[-LnU\right]^{\frac{1}{\alpha}}$ puisque $(1-U)\rightarrow\mathcal{U}(0,1)$

16

Simulation des lois de probabilités discrètes

17

Les lois de probabilités discrètes:

- On considère une variable aléatoire discrète X qui peut prendre les valeurs $x_1, x_2, \ldots, x_n$.
- Soit $P(X=x_i)=p_i$,   i=1,2,…,n

  avec $p_i \geq 0$ et $\sum p_i=1$
- Sa loi de probabilité est donnée par:

x_i	x_1	x_2	x_3	x_4	….	x_n	$\sum p_i$
p_i	p_1	p_2	p_3	p_4	….	p_n	1

18

Fonction de répartition d'une loi de probabilité discrète:

- Sa fonction de répartition $F(x_i)=F_i=P(X\leq x_i)=\sum_{j<i} p_j$

$$F(x)=\begin{cases} 0 & si\ x < x_1 \\ p_1 & si\ x_1 \leq x < x_2 \\ p_1+p_2 & si\ x_2 \leq x < x_3 \\ p_1+p_2+p_3 & si\ x_3 \leq x < x_4 \\ \vdots & \\ 1 & si\ x \geq x_n \end{cases}$$

19

Fonction de répartition (inverse) d'une loi de probabilité discrète:

- On définit la fonction $\overline{F}$

$$\overline{F}(u)=\begin{cases} x_1 & si\quad 0<u<F(x_1) \\ x_2 & si\quad F(x_1)\leq u<F(x_2) \\ x_3 & si\quad F(x_2)\leq u<F(x_3) \\ x_4 & si\quad F(x_3)\leq u<F(x_4) \\ \vdots & \\ x_n & si\quad F(x_{n-1})\leq x<F(x_n)=1 \end{cases}$$

20

Méthode d'inversion pour les lois discrètes (Théorème):

- Soit $\overline{F}(u) = \min\{x : F(x) \geq u\}$
- Si U→U([0,1]), alors la v.a. $X = \overline{F}(U)$ a pour fonction de répartition F.
- Donc, dans ces conditions $\overline{F}$ joue le rôle de F^{-1}

21

Méthode d'inversion-Dém:

Remarquons le minimum est atteint parce que F est continue à droite, alors $\overline{F}$ est bien définie.

En plus, $F\left(\overline{F}(u)\right) \geq u$

et $\overline{F}\left(F(x)\right) = \min\{y : F(y) \geq F(x)\} \leq x$

D'où l'égalité des ensembles:

$$\left\{(u,x) : \overline{F}(u) \leq x\right\} = \left\{(u,x) : u \leq F(x)\right\}$$

et des probabilités:

$$P(X \leq x) = P\left(\overline{F}(U) \leq x\right) = P\left(U \leq F(x)\right) = F(x)$$

22

Algorithme de la méthode d'inversion pour les lois discrètes:

- Pour une loi discrète générale, on a $\overline{F}(u)=i$ avec $F_{i-1}<u\leq F_i$, donc la méthode d'inversion est équivalente à chercher l'indice i convenable dans la liste des F_i.
- En général:

$$\overline{F}(u)=x_i \qquad si\ \ F(x_{i-1})=F_{i-1}\leq u<F_i=F(x_i)$$

- Algorithme:

Générer $U \rightarrow \mathcal{U}(0,1)$
Tant que $F_i \leq U$, Faire $i=i+1$
Sortir $X=i$

Simulation d'une loi de Bernoulli de paramètre p ∈ [0, 1]:

- La loi de probabilité de X suivant une loi de Bernoulli est:

x_i	0	1	Σp_i
p_i	q=1-p	p	1

- Sa fonction de répartition est:

$$F(x)=\begin{cases} 0 & si\ \ x<0 \\ q=1-p & si\ 0\leq x<1 \\ 1 & si\ x\geq 1 \end{cases}$$

Simulation d'une loi de Bernoulli de paramètre p ∈ [0, 1]:

- Sa fonction de répartition inverse sera:

$$\overline{F}(u)=\begin{cases}0 & si\ \ 0<u<q=1-p\\ 1 & si\ q=1-p\leq u<1\end{cases}$$

- D'où l'algorithme:

Générer $U \to \mathcal{U}(0,1)$
Si U≥1-p, sortir $X=1$
Autrement, Sortir $X=0$

25

Simulation d'une loi uniforme discrète:

- Une v.a. X suivant une loi de probabilité uniforme discrète prend les valeurs entiers naturelles i=1,2,…,n avec la même probabilité

$$P(X=i)=\frac{1}{n},\ \ \forall i=1,2,\ldots,n$$

- Sa fonction de répartition est:

$$F(x)=P(X\leq i)=\sum_{j\leq i} p_j=\frac{i}{n}\qquad si\ \ i\leq x<i+1$$

26

Simulation d'une loi uniforme discrète:

- Sa fonction de répartition inverse est:

$$\overline{F}(u) = X = i \quad si \;\; F(x_{i-1}) = F_{i-1} = \frac{i-1}{n} \leq u < \frac{i}{n} = F_i = F(x_i)$$

$$\Rightarrow i-1 \leq nu \leq i$$

ou encore $X = ent(nu)+1$

27

Exemples sur Excel

28

Simulation d'une loi par la méthode d'inversion sur Excel

- Avec Excel, on peut simuler la réalisation d'une variable de loi uniforme en utilisant la fonction **ALEA.**
- Par conséquent, on simule une réalisation de la variable X en appliquant la réciproque de sa fonction de répartition au résultat de la fonction ***ALEA***.

29

Exemple: La loi normale standard

- La fonction de répartition est calculée par la fonction **LOI.NORMALE.STANDARD,** sa réciproque par la fonction **LOI.NORMALE.STANDARD.INVERSE.**
- Par exemple, la formule **=LOI.NORMALE.STANDARD(1,96)** donne 0,975, probabilité d'obtenir une valeur inférieure à 1,96.
- En appliquant la fonction réciproque sur le résultat, c'est-à-dire en tapant la formule **=LOI.NORMALE.STANDARD.INVERSE(0,975),** on retrouve la valeur de départ, 1,96.

30

Exemple: La loi normale standard

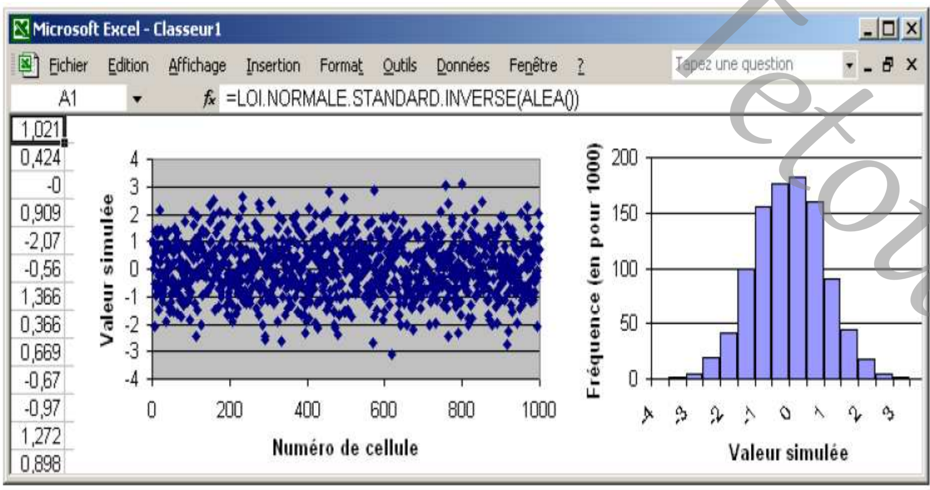
- Appliquons la réciproque de la fonction de répartition au résultat de la fonction **ALEA.**

=LOI.NORMALE.STANDARD.INVERSE(ALEA())

- Recopions la formule sur 1000 cellules et représentons la distribution des valeurs obtenues par un nuage de points ou par un histogramme.

31

Exemple: La loi normale standard



32

Lois discrètes: Méthode générale

- On fait 2 lancers de pièce de monnaie et on note le nombre de fois où on a retourné le côté face.
- La variable correspondante est distribuée de la manière suivante :

x_i	0	1	2
P_i	0,25	0,5	0,25

33

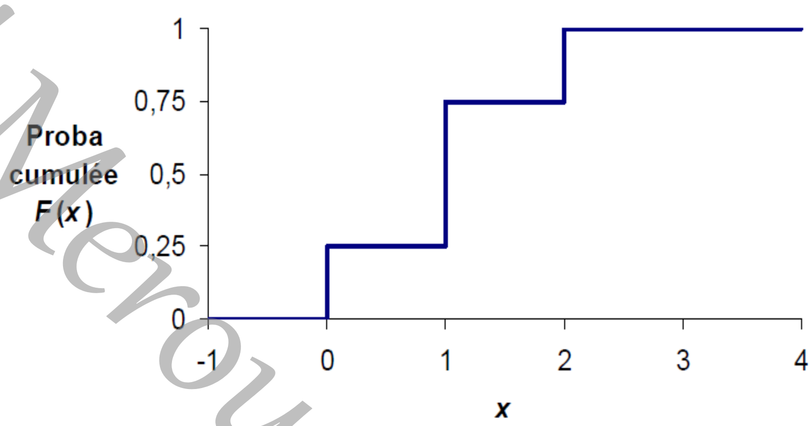
Lois discrètes: Méthode générale

- Sa fonction de répartition est:

$$F(x)=\begin{cases} 0 & si\ \ x<0 \\ 0{,}25 & si\ \ 0\leq x<1 \\ 0{,}75 & si\ 1\leq x<2 \\ 1 & si\ x\geq 2 \end{cases}$$

34

Lois discrètes: Méthode générale



35

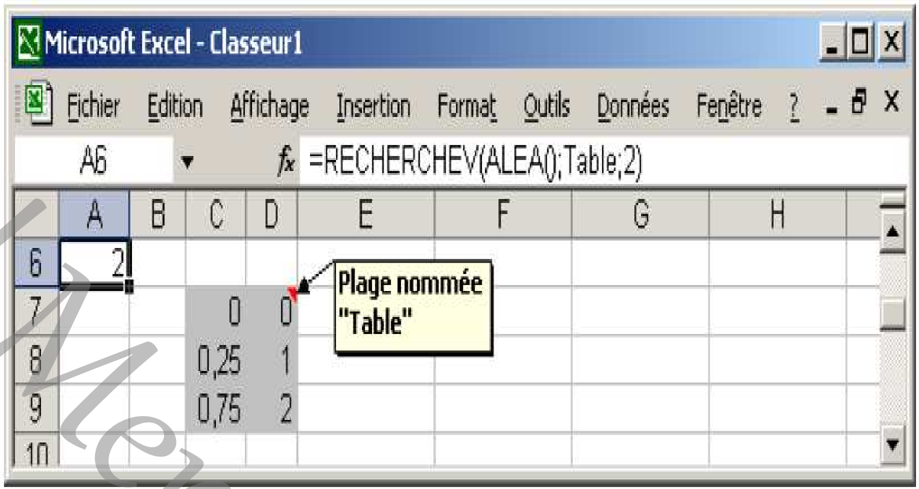
Lois discrètes: Méthode générale

- La fonction de répartition inverse sera:

$$\overline{F}(u) = \begin{cases} 0 & si \ u < 0{,}25 \\ 1 & si \ 0{,}25 \le u < 0{,}75 \\ 2 & si \ 0{,}75 \le u < 1 \end{cases}$$

- Cette fonction peut être réalisée sur Excel avec la fonction **RECHERCHEV.**

36



- Vérifiez que les valeurs de la variable apparaissent avec des fréquences proches des probabilités.

37

La loi discrète uniforme

- On veut simuler le point marqué par un dé.
- Pour faire correspondre aux valeurs de la fonction **ALEA :**
  - la valeur 1 aux valeurs comprises entre 0 et 1/6,
  - la valeur 2 aux valeurs entre 1/6 et 2/6,
  - etc.,
- Il suffit de multiplier le résultat de la fonction ALEA par 6, d'éliminer les décimales du résultat et d'ajouter une unité, ce que fait la formule

=ENT(ALEA()*6)+1

38

Loi de Bernouilli

- Pour simuler le résultat de lancer d'une pièce de monnaie, il suffit de dire qu'on a retourné le côté face si la résultat de la fonction **ALEA** est plus petit que 0,5:

  =SI(ALEA()<1/2;"Pile";"Face")
- De manière générale, on simule la réalisation d'un événement A de probabilité p par la formule

  =SI(ALEA()<p;"A";"Non A")
- et la variable indicatrice de l'événement, c'est-à-dire la variable qui prend la valeur 1 si l'événement est réalisé et 0 sinon par

  =SI(ALEA()<p;1;0)

39

Formules pour les lois usuelles

- En appliquant la méthode d'inversion, on simule avec Excel toutes les lois de probabilités usuelles

Loi	***Formule***
discrète uniforme sur les valeurs entières de 1 à n	**=ALEA.ENTRE.BORNES(1;n)**

40

Loi	*Formule*
Bernouilli de paramètre p	**=SI(ALEA()<p;1;0)**
binomiale de paramètres n et p	**=CRITERE.LOI.BINOMIALE(n;p;ALEA())**
Poisson de paramètre m	On peut utiliser la convergence de la loi binomiale vers la loi de Poisson et reprendre la formule du dessus en remplaçant **n par 10^4 et p par m/10^4**

41

Loi	*Formule*
hypergéométrique	utiliser **RECHERCHEV à partir des probabilités cumulées calculées avec la fonction LOI.HYPERGEOMETRIQUE**
géométrique de paramètre p	**=ARRONDI.SUP(LN(ALEA())/LN(1-p);0) le fractile d'ordre 1-α étant l'entier supérieur ou égal à ln(α)/ln(1-p)**
continue uniforme entre 0 et 1	**=ALEA()**

42

Loi	*Formule*
continue uniforme entre a et b	**=ALEA()*(b-a)+a**
gamma de paramètre r	**=LOI.GAMMA.INVERSE(ALEA();r;1)**
bêta de paramètres n et p entre les bornes a et b	**=BETA.INVERSE(ALEA();n;p;a;b)**

43

Loi	*Formule*
normale standard	**=LOI.NORMALE.STANDARD.INVERSE(ALEA())**
normale de paramètres m et sigma	**=LOI.NORMALE.INVERSE(ALEA();m;sigma)**
log-normale	**=LOI.LOGNORMALE.INVERSE(ALEA();m;sigma)**
Khi-deux à nu ddl	**=KHIDEUX.INVERSE(ALEA();nu)**

44

Loi	Formule
Student à nu ddl	=LOI.STUDENT.INVERSE(ALEA();nu)
Fisher à nu1 et nu2 ddl	=INVERSE.LOI.F(ALEA();nu1;nu2)
exponentielle de paramètre lambda	=-1/lambda*LN(ALEA())

45

Exemple

- Un examen consiste en une série de 40 questions indépendantes comptant chacune pour le même nombre de points. A chaque question, 4 réponses sont proposées dont une seule est exacte. A-t-on des chances d'avoir plus de la moyenne quand on répond au hasard ?

46

Exemple

- Le nombre de bonnes réponses suit une loi binomiale, qu'on peut simuler par la formule

  =CRITERE.LOI.BINOMIALE(n;p;ALEA())

  avec n = 40 (nombre de questions)

  et p = ¼ (4 choix possibles par question).
- Mais c'est avant tout la somme des indicatrices de succès à chaque question (on note 1 si la réponse à la question est bonne, 0 sinon).

47

Exemple

- C'est le « processus de Bernouilli », qu'on simule par la ligne de formule suivante :

Microsoft Excel - formules

Fichier Edition Affichage Insertion Format Outils Données Fenêtre ?

	A		AM	AN	AO
6	Q1	Q2	Q39	Q40	Nombre de bonnes réponses
7	=SI(ALEA()<1/4;1;0)	=SI(A	=SI(ALEA()<1/4;1;0)	=SI(ALEA()<1/4;1;0)	=SOMME(A7:AN7)

48

Exemple

- Recopiez les formules sur 1000 lignes. Représentez la distribution du nombre de bonnes réponses par un diagramme en bâtons.
- Comparez la à la distribution de probabilités (calculées avec la fonction **LOI.BINOMIALE).**
- L'étudiant a-t-il des chances d'avoir coché la bonne réponse à plus de la moitié des questions ?

49