Cette fiche a été élaborée par des enseignantes et des enseignants des lycées et universités de l'académie de Créteil.

Thème	Intégration – Lois de probabilité à densité
Titre	Méthode de Monte-Carlo

Objectif

- Découvrir et utiliser deux méthodes/algorithmes classiques d'approximation d'intégrales : méthode des rectangles et méthode de Monte-Carlo.

Prérequis

- Intégrale d'une fonction positive
- (Loi uniforme)
- Facultatif : Intervalle de confiance

Mise en place

Séance de 1h d'accompagnement personnalisé dans une salle équipée d'ordinateurs ou à défaut utilisation des calculatrices graphiques des élèves.
Les questions étant relativement guidées, les élèves devraient pouvoir travailler en relative autonomie, le professeur intervenant pour vérifier l'avancée des élèves et donner des explications complémentaires si nécessaire.
Tous les élèves doivent pouvoir avoir fait les parties 1 et 2 à la fin de l'heure.
Les exercices page 4 pourront être donnés en travail à la maison :

- Exercice 1 : Application de la méthode de Monte-Carlo au calcul approché de π (accessible à tous);
- Exercice 2 : Requiert davantage d'initiatives (réservé aux élèves les plus à l'aise).

Sources/Ressources

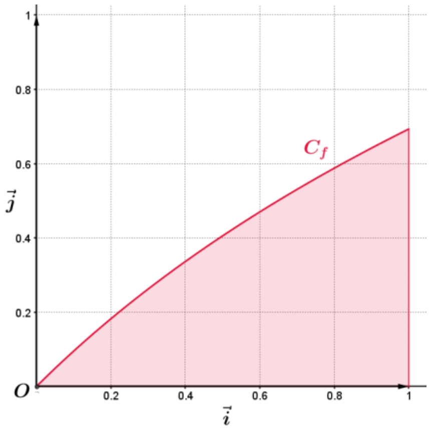
[1] De nombreuses animations concernant différentes méthodes d'intégration numérique (IREM d'Orléans-Tours).

[2] Recherche d'une valeur approchée de π par la méthode de Monte-Carlo au niveau 2^{nde} (Académie de Grenoble).

Prolongements possibles

- Calcul d'intégrale par intégration parties.
- Les différentes approximations de π et leur histoire.
- Les autres techniques d'intégration numérique (trapèzes, Simpson, …)

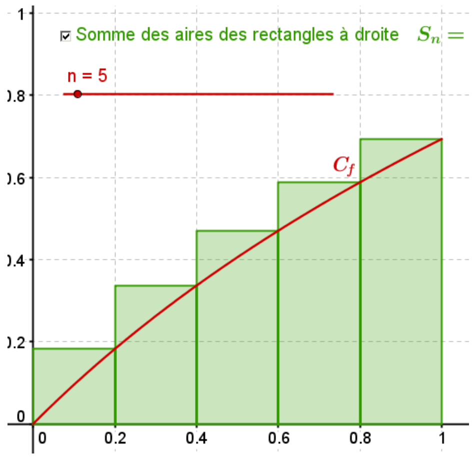
L'objectif de cette activité est d'utiliser deux méthodes différentes pour approcher l'aire d'une surface donnée lorsqu'on ne sait pas la déterminer par le calcul.



On considère l'intégrale $I=\int_0^1 f(x)dx$, où f est la fonction définie par $f(x)=\ln(x+1)$.
On note C_f la courbe de la fonction f.

Question préliminaire :
Justifier pourquoi on peut interpréter graphiquement l'intégrale I comme l'aire, en unités d'aire, du domaine colorié (délimité par C_f, l'axe des abscisses et les droites d'équation $x = 0$ et $x = 1$).

Partie 1 Une première méthode déterministe : la méthode des rectangles



<u>*Principe de la « Méthode des Rectangles » (à droite*[1]*) :*</u>

- On partage l'intervalle [0,1] en n intervalles de même amplitude (n entier naturel non nul) ;
- Sur chacun des intervalles on remplace la courbe représentative de f par un rectangle de hauteur l'image par f de la borne *supérieure* de l'intervalle en question ;
- La somme S_n des aires des n rectangles obtenus est une valeur approchée de I.

Sur la figure ci-contre $n = 5$.

1. Quelle est la largeur de chacun des 5 intervalles qui subdivisent l'intervalle [0,1] ?
2. Calculer S_5, puis une valeur approchée à 10^{-3} près.

Plus généralement :

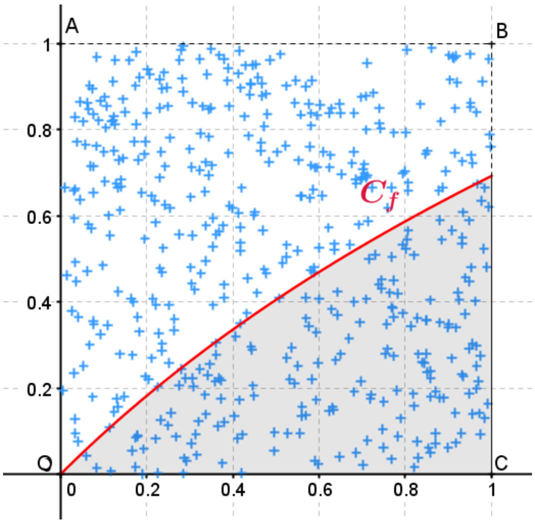
3. Exprimer la largeur d'un des n intervalles en fonction de n.
4. Exprimer S_n en fonction de n.
5. On considère le programme ci-contre écrit avec Scilab. Réécrire ce programme (lignes 5 à 10) en langage naturel puis expliquer son fonctionnement.
6. Programmer et exécuter ce programme afin d'obtenir :
   a) une valeur approchée de S_5
   b) une valeur approchée de S_{10}
   c) une valeur approchée de S_{100}
   d) une valeur approchée de S_{1000}.

   (Nommer ce programme **M_Rectangles** et l'enregistrer).

```
1  function y=f(x)
2      y=ln(x+1)
3  endfunction
4
5  S=0
6  n=input("nombre de rectangles n = ")
7  for i=1:n
8      S=S+1/n*f(i/n)
9  end
10 afficher(S)
```

[1] En utilisant sur chaque intervalle plutôt la borne *inférieure* (« *à gauche* ») de l'intervalle en question, on obtient la méthode équivalente des rectangles « à gauche ». Dans le cas d'une fonction monotone, on peut obtenir un *encadrement* de l'aire sous la courbe en utilisant à la fois les rectangles « à gauche » et « à droite ».

Partie 2 Une deuxième méthode (probabiliste) : la méthode de Monte-Carlo



Principe de la méthode :

- On prend au hasard n points dans le carré unité OABC[2] ;
- On compte le nombre de points k situés SOUS la courbe représentative de f ;
- La fréquence $\frac{k}{n}$ est une valeur approchée de $\frac{I}{aire(OABC)} = I$

On a traduit ci-dessous l'algorithme correspondant en langage Scilab.

```
1  function y=f(x)
2      y=ln(x+1)
3  endfunction
4  k=0
5  n=input("nombre de points n = ")
6  for i =1:n
7      x=rand()
8      y=rand()
9      if y<f(x) then k=k+1
10     end
11 end
12 afficher(k/n)
```

1. A quoi correspondent les variables n, k, x et y ?
2. Expliquer le test de la ligne 9 du programme.
3. Que renvoie le programme en fin d'exécution ?
4. Programmer l'algorithme correspondant dans le langage de votre choix *(Nommer le programme MonteCarlo et l'enregistrer).*
   Quels résultats obtenez-vous pour $n = 100$, $n = 1\,000$, $n = 10\,000$?
   Relancer une nouvelle fois l'algorithme pour chacune de ces valeurs de n. Que remarquez-vous ?

Partie 3 Comparaison de performances

La commande « integrate » de Scilab (voir ci-contre) permet d'obtenir une valeur approchée de I très précise (on admet ici que toutes les décimales affichées sont correctes).

```
-->integrate('ln(1+x)','x',0,1)
 ans  =

    0.3862943611199
```

1. Combien de décimales exactes de I avez-vous obtenu pour chacune des valeurs de n
   a) avec la méthode des rectangles ?
   b) avec la méthode de Monte-Carlo ? (relancer plusieurs fois les simulations si nécessaire)
2. *(Seulement si le chapitre « Echantillonnage et Estimation » a déjà été traité)*
   On s'intéresse maintenant au cas d'une fonction f quelconque, positive sur [0 ;1].
   a) On admet que la précision pour la méthode des rectangles obtenue avec n rectangles est de l'ordre de $1/n$.
   En déduire le nombre de rectangles n nécessaire pour obtenir une estimation de l'aire I sous la courbe de f sur l'intervalle [0 ;1] avec une précision égale à 10^{-2}.
   b) Donner à partir de la fréquence $f = k/n$ obtenue par la méthode de Monte-Carlo, l'expression d'un intervalle de confiance de I au niveau de confiance de 95%. En déduire le nombre de points n nécessaire pour obtenir, au niveau de confiance de 95%, une estimation de I avec une précision égale à 10^{-2}.
3. Quelle méthode est la plus performante ? Quel avantage a pourtant la méthode la moins performante ?

---

[2] Prendre au hasard un point dans le carré *OABC* c'est prendre au hasard un point $M(x, y)$ où les coordonnées x et y sont des réalisations de variables aléatoires X et Y de **loi uniforme** sur [0 ;1].

A vous de jouer…

Exercice 1

On peut montrer (*aide* : « *cercle* ») que $J = \int_0^1 \sqrt{1-x^2}\,dx = \frac{\pi}{4}$.

a) En appliquant la méthode de Monte-Carlo pour déterminer une valeur approchée de J, on obtient donc également une valeur approchée de π (au facteur 4 près).

b) Ecrire un algorithme mettant en œuvre la méthode de Monte-Carlo, qui :

- demande le nombre n de points à utiliser,
- calcule les valeurs approchées de J et de π correspondantes,
- et qui affiche finalement la valeur approchée de π obtenue.

c) Programmer et faire fonctionner l'algorithme sur calculatrice ou ordinateur.
Donner les résultats obtenus pour n = 1 000, n = 10 000, n = 100 000.

Exercice 2 ***(Après le chapitre « Echantillonnage et Estimation »)***

a) En adaptant (*) le principe de la méthode de Monte-Carlo exposée dans la partie 2, donner une valeur approchée de $P = \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\,dx$

b) Donner, à partir de la fréquence $f = k/n$ obtenue par cet algorithme, l'expression d'un intervalle de confiance de P au niveau de confiance de 95% .

d) Quel est le plus petit entier n pour lequel on obtient, au niveau de confiance de 95%, une estimation de P à la précision (donnée par la longueur de l'intervalle de confiance) égale à 10^{-3} ? Donner la valeur approchée correspondante.

d) Un commentaire ?

(*) Cette fois-ci on ne prend plus les points au hasard dans un carré d'aire 1, mais dans un rectangle [– 1 ; 1] × [0, M] où M est un majorant de f sur [– 1 ;1].
*Pour simuler une **loi uniforme sur intervalle [a, b]** à l'aide de la fonction rand() de Scilab, on peut utiliser la commande : **a + (b-a)*rand().***

Au sujet de la Méthode de Monte-Carlo

Le terme **méthode de Monte-Carlo** désigne toute méthode visant à calculer une valeur numérique en utilisant des procédés aléatoires. Le nom de ces méthodes, qui fait allusion aux **jeux de hasard** pratiqués à Monte-Carlo, a été inventé en 1947 par Nicholas Metropolis. Le véritable développement des méthodes de Monte-Carlo s'est effectué sous l'impulsion de **John von Neumann** et Stanislas Ulam notamment, lors de la Seconde Guerre mondiale et des recherches sur la fabrication de la bombe atomique.
Les méthodes de Monte-Carlo sont souvent utilisées pour **calculer des intégrales** qu'on ne sait pas calculer autrement, en particulier pour des intégrales en dimensions plus grandes que 1 (**calculs de surfaces et de volumes** par exemple).
Elles sont également couramment utilisées en **physique des particules**, et pour la gestion de **risque dans une décision financière**.

d'après Wikipedia FR

Pour le professeur

Partie 1 – Méthode des rectangles

Fichiers Joints	**Scilab**	**Algobox**
Calcul d'une valeur approchée de l'aire sous une courbe sur l'intervalle [0;1] par la méthode des rectangles (à droite).	**M_Rectangles.sci**	**M_Rectangles.alg** *La fonction est définie dans l'onglet "Utiliser une fonction numérique".*
Calcul approché de l'aire sous une courbe sur l'intervalle [a;b] par la méthode des rectangles (à gauche et à droite). *On demande à la fonction d'être monotone pour obtenir un encadrement de l'aire recherchée.*	**M_Rectangles_plus.sci**	**M_Rectangles_plus.alg**
Geogebra		
Illustration de la méthode des rectangles n variable de 0 à 100 Rectangles à gauche et à droite possibles	**M_Rectangles.ggb**	

Résultats Numériques

nombre de rectangles	valeur approchée obtenue	nombre de décimales correctes
n = 5	**0.4**539462536246	0
n = 10	**0.4**205352957738	0
n = 100	**0.389**7559303803	2
n = 1000	**0.3866**408930435	3
		commandes
valeur exacte	**2ln(2)-1 ≈ 0.3862943611199**	XCAS en ligne integrate(ln(1+x),x,0,1)
valeur approchée à 10^{-13} près	**0.3862**943611199	Scilab integrate('ln(1+x)',x,0,1)

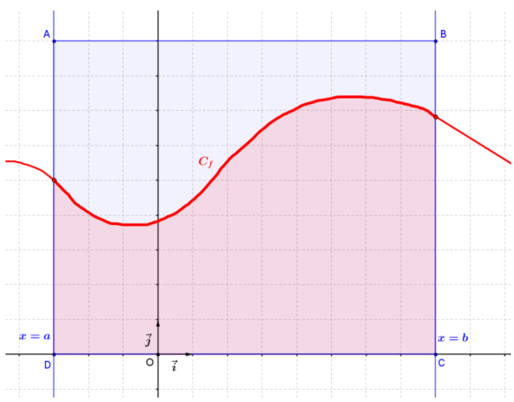
Partie 2 – Méthode de Monte-Carlo

Fichiers Joints	**Scilab**	**Algobox**
Calcul d'une valeur approchée de l'aire sous la courbe $y = \ln(x+1)$ sur l'intervalle [0;1] par la méthode de Monte-Carlo.	**MonteCarlo.sci**	**MonteCarlo.alg** *La fonction est définie dans l'onglet "Utiliser une fonction numérique".*
Calcul d'une valeur approchée de l'aire sous la courbe $y = \ln(x+1)$ sur l'intervalle [0;1] par la méthode de Monte-Carlo + *ILLUSTRATION graphique.*		**MonteCarlo_Graph.alg**

Compléments

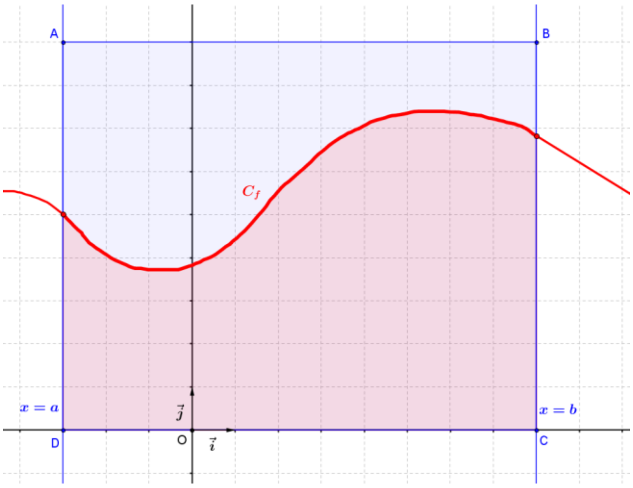
Soit f une fonction positive sur un intervalle $[a;\ b]$, et C_f sa représentation graphique dans un repère orthogonal.

On cherche à déterminer une valeur approchée de l'intégrale I de f sur $[a;\ b]$, c'est-à-dire de l'aire du domaine D compris entre l'axe des abscisses, la courbe C_f et les droites d'équations $x = a$ et $x = b$.



- On choisit un rectangle R (ABCD sur la figure) qui contient le domaine D.
- L'expérience aléatoire est la suivante : on prend un point au hasard dans le rectangle R.
- On admet que la probabilité de l'événement « le point appartient au domaine D » est égale au quotient : $p = \dfrac{aire(D)}{aire(R)}$

  Remarquons que dans le cas où aire(R) = 1, on a p = aire(D) = I.



Pour tout n entier compris entre 1 et n, notons X_i la variable aléatoire telle que :

- $X_i = 1$ si le point n°i est en-dessous la courbe de f ;
- $X_i = 0$ si le point n°i est au-dessus de la courbe de f.

On a :

- $P(X_i = 1) = p$
- $P(X_i = 0) = 1 - p$ où $p = \dfrac{aire(D)}{aire(R)}$

L'expérience consiste donc en la répétition de n expériences de Bernoulli, indépendantes, de même paramètre p : on a affaire à un shéma de Bernoulli.
Si on note S_n la variable aléatoire donnant le nombre de succès pour ce schéma de Bernoulli, c'est-à-dire le nombre de points situés sous la courbe, alors on sait que S_n suit une loi binomiale de paramètres n et p.
L'espérance et l'écart-type de S_n sont donnés par :

$$E(S_n) = np \qquad \text{et} \qquad \sigma(S_n) = \sqrt{np(1-p)}$$

Notons maintenant $F_n = \dfrac{1}{n} S_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$.

En utilisant les propriétés vues en 1ereS : $E(aX+b) = aE(X)+b$ et $Var(aX+b) = a^2 Var(X)$, on obtient :

$$E(F_n) = E\left(\frac{1}{n}S_n\right) = \frac{1}{n}E(S_n) = \frac{np}{n} = p \text{ et } Var(F_n) = Var\left(\frac{1}{n}S_n\right) = \frac{1}{n^2}Var(S_n) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

d'où $E(F_n) = p$ et $\sigma(F_n) = \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}$.