EPITECH
PROBABILITES ET STATISTIQUES
Année 20011-2012
Mini-projet 204param

1 Sujet / 1ère partie : intégration numérique

Le but est d'écrire un programme qui calcule les intégrales de la forme :

$$\int_a^b f(x)\,dx$$

Les données sont la fonction f(.) et les bornes d'intégration a et b. Pour ce calcul, on utilisera la méthode d'intégration numérique de Simpson décrite en annexe. Cette méthode est basée sur une subdivision de l'intervalle [a,b] en sous-intervalles réguliers. On choisira 20 sous-intervalles (ou plus) par intervalle unité (par exemple sur [0,1]). C'est à dire, que sur l'intervalle [a,b], on choisira comme nombre de subdivisions, l'entier le plus proche de (b-a)*20.

Pour tester votre programme, vous pourrez utiliser les résultats suivants sur l'intervalle [0,1] :

$$\int_0^1 1\,dx = 1$$

$$\int_0^1 x\,dx = 1/2$$

$$\int_0^1 x^2\,dx = 1/3$$

$$\vdots$$

$$\int_0^1 x^n\,dx = 1/(n+1) \qquad n \geq 0$$

2 Sujet / 2ème partie : espérance et variance

On demande maintenant au programme d'évaluer l'espérance mathématique et la variance de 5 variables aléatoires continues dont les densités de probabilités sont données ci-après. Naturellement, on utilise les formules suivantes :

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\,dx$$

$$var(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f(x)\,dx$$

où f(.) représente la densité de probabilité de la variable X.

Il n'est pas question numériquement d'intégrer sur l'intervalle $]-\infty,+\infty[$. On se restreindra donc, pour chacune des 5 lois, à un intervalle borné de type [a,b] en dehors duquel la densité est nulle ou quasi-nulle. Le choix de a et b vous est conseillé pour chacune des 5 lois.

Les lois dont on demande l'espérance et la variance sont les suivantes :

Loi	Densité de probabilité	Intervalle [a,b]
n°1	uniforme sur [2,5]	[2,5]
n°2	f(x) = 1-|x| sur [-1,1] (=0 sinon)	[-1,1]
n°3	f(x) = e^{-x} si x>0 (=0 sinon)	[0,100]
n°4	gauss $g_{0,1}$	[-10,10]
n°5	gauss $g_{3,4}$	[-10,20]

L'auteur a obtenu les résultats suivants :

f(x) = 1/(5-2) si x appartient à [2,5]
E(X) = 3,500
Var(X) = 0,750

f(x) = 1-|x| si x appartient à [-1,1]
E(X) = 0,000
Var(X) = 0,167

f(x) = exp(-x) si x>0
E(X) = 1,000
Var(X) = 1,000

f(x) = $gauss_{0,1}$(x) = $\frac{1}{\sqrt{2\pi}}.exp(-x^2/2)$
E(X) = 0,000
Var(X) = 1,000

f(x) = $gauss_{3,4}$(x) = $\frac{1}{2\sqrt{2\pi}}.exp\left(-\frac{(x-3)^2}{8}\right)$
E(X) = 3,000
Var(X) = 4,000

3 Le logiciel

Répertoire de rendu : $\tilde{/}$../rendu/math/204param/
Nom de l'exécutable : 204param
Exemple de lancer :

$>$ **204param**
En entrée : pas d'arguments.
En sortie : les cinq espérances et variances des cinq lois. Il est interdit de calculer ces quantités autrement que par intégration numérique (voir annexe).

4 Questions

1° Qu'est-ce qu'une variable aléatoire continue ?
2° Donner des exemples d'expériences aléatoires associées à une v.a. continue.
3° Quel sont les ensembles de départ et d'arrivée d'une v.a. continue réelle ?
4° Soit X une v.a. continue à valeurs dans $\mathbb{R}$. Que dire de P(X=x) où x est un réel quelconque ?
5° Quels sont les évènements caractéristiques associés à une v.a. continue ?

Annexe : méthode de Simpson

Dans ce paragraphe, on décrit la méthode de Simpson qui permet de calculer numériquement une intégrale :

$$\int_a^b f(x)\,dx$$

On choisit une subdivision à pas constant noté h de l'intervalle [a,b]. Le pas h est égal à :

$$h = \frac{b-a}{n}$$

Les (n+1) points de la subdivision sont notés x_i et vérifient :

$$\begin{cases} x_0 = a \\ x_n = b \\ x_{i+1} = x_i + h, \quad i = 0, ..., n-1 \end{cases}$$

On écrit alors l'intégrale sur [a,b] comme somme d'intégrales sur chacun des sous intervalles :

$$\int_a^b f(x)\,dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)\,dx$$

Si le pas h de la subdivision est suffisamment petit (n grand), alors on peut supposer que la fonction f(x) est quasiment constante sur chaque sous intervalle. On remplace alors la valeur de f(x) par une valeur constante. Plusieurs choix sont possibles. Soit $]x_i, x_{i+1}[$ un intervalle de la subdivision, on peut choisir :

$$\begin{cases} \text{soit } f(x) = f(x_i) & \text{valeur à gauche} \\ \text{soit } f(x) = \frac{f(x_i)+f(x_{i+1})}{2} & \text{valeur milieu} \\ \text{soit } f(x) = f(x_{i+1}) & \text{valeur à droite} \\ \ldots & \end{cases}$$

Mais ces différents choix ne procurent pas de très bons résultats. Dans la méthode de Simpson, sur chaque sous intervalle, on utilise une formule obtenue par intégration d'un polynôme de degré 2. Cette formule est la suivante :

$$\int_{x_i}^{x_{i+1}} f(x)\,dx = \frac{h}{6}\left(f(x_i) + f(x_{i+1}) + 4.f\left(\frac{x_i + x_{i+1}}{2}\right)\right)$$

On utilise alors une somme de cette formule pour l'intervalle [a,b]. On obtient l'approximation suivante :

Méthode de Simpson
$\int_a^b f(x)\,dx \simeq \frac{b-a}{6n}\left(f(a) + f(b) + 2.\sum_{i=1}^{n-1} f(a+i.h) + 4.\sum_{i=0}^{n-1} f(a+i.h+\frac{h}{2})\right)$

Cette formule est exacte pour les polynômes de degré 3. L'erreur d'approximation commise est en $\frac{1}{n^4}$. C'est une méthode qui donne de très bons résultats sans que l'on soit obligé de choisir une valeur élevée pour n.