Variance et intervalles

Épreuve pratique d'algorithmique et de programmation
Concours commun des écoles normales supérieures

Durée de l'épreuve: 3 heures 30 minutes

Juin Juillet 2012

ATTENTION !

N'oubliez en aucun cas de recopier votre u_0
à l'emplacement prévu sur votre fiche réponse

Important.

Sur votre table est indiqué un numéro u_0 qui servira d'entrée à vos programmes. Les réponses attendues sont généralement courtes et doivent être données sur la fiche réponse fournie à la fin du sujet. À la fin du sujet, vous trouverez en fait deux fiches réponses. La première est un exemple des réponses attendues pour un $\widetilde{u_0}$ particulier (précisé sur cette même fiche et que nous notons avec un tilde pour éviter toute confusion!). Cette fiche est destinée à vous aider à vérifier le résultat de vos programmes en les testant avec $\widetilde{u_0}$ au lieu de u_0. Vous indiquerez vos réponses (correspondant à votre u_0) sur la seconde et vous la remettrez à l'examinateur à la fin de l'épreuve.

En ce qui concerne la partie orale de l'examen, lorsque la description d'un algorithme est demandée, vous devez présenter son fonctionnement de façon schématique, courte et précise. Vous ne devez en aucun cas recopier le code de vos procédures !
Quand on demande la complexité en temps ou en mémoire d'un algorithme en fonction d'un paramètre n, on demande l'ordre de grandeur en fonction du paramètre, par exemple: $O(n^2)$, $O(n \log n)$,...
Il est recommandé de commencer par lancer vos programmes sur de petites valeurs des paramètres et de ***tester vos programmes sur des petits exemples que vous aurez résolus préalablement à la main ou bien à l'aide de la fiche réponse type fournie en annexe***. Enfin, il est recommandé de lire l'intégralité du sujet avant de commencer afin d'effectuer les bons choix de structures de données dès le début.

1 Génération d'intervalles

Dans ce sujet, un intervalle désignera un ensemble fermé et borné de nombres réels, aussi appelé un segment. Un tel intervalle sera noté $\mathbf{x}$ tandis que x désignera un réel quelconque dans $\mathbf{x}$. La borne inférieure de l'intervalle $\mathbf{x}$ sera notée $\underline{x}$ tandis que sa borne supérieure sera notée $\overline{x}$. Autrement dit, $\mathbf{x} = \{x \in \mathbb{R} \mid \underline{x} \leq x \leq \overline{x}\}$.
Un n-uplet de valeurs $(x_1, \ldots, x_n)$ sera noté $\vec{x}$ tandis qu'un n-uplet d'intervalles $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ sera noté $\vec{\mathbf{x}}$. Par extension de la notation d'appartenance $x \in \mathbf{x}$, l'expression $\vec{x} \in \vec{\mathbf{x}}$ désignera la conjonction des propriétés $x_i \in \mathbf{x}_i$ pour $1 \leq i \leq n$.
Considérons la suite d'entiers $(u_k)_{k \in \mathbb{N}}$ définie par

$$u_{k+1} = 15091 \times u_k \mod 64007.$$

L'intervalle $\mathbf{x}_i$ est donné par $[\min(v_{2i}, v_{2i+1}), \max(v_{2i}, v_{2i+1})]$ avec $v_k = \lfloor u_k/320 \rfloor - 100$.

Question 1 *Combien valent les intervalles **a)** $\mathbf{x}_1$? **b)** $\mathbf{x}_{100}$? **c)** $\mathbf{x}_{10000}$?*

Le centre d'un intervalle $\mathbf{x}$ se note $\hat{x}$ et vaut $(\underline{x}+\overline{x})/2$. La méthode utilisée pour générer les intervalles de $\vec{\mathbf{x}}$ rend possible la présence de plusieurs intervalles ayant un même centre.

Question 2 *Quel est le centre le plus souvent présent dans $\vec{\mathbf{x}}$ et quel est son nombre d'occurrence pour **a)** $n = 100$? **b)** $n = 1000$? **c)** $n = 10000$? S'il y a plusieurs centres possibles, vous indiquerez le plus grand d'entre eux.*

2 Espérance et variance

On appelle espérance ou moyenne de $\vec{x}$ le nombre

$$E(\vec{x}) = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

On appelle variance de $\vec{x}$ le nombre

$$V(\vec{x}) = \frac{1}{n} \sum_{i=1}^{n} (x_i - E(\vec{x}))^2 = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right) - E(\vec{x})^2.$$

On s'intéresse au cas où les valeurs de $\vec{x}$ ne sont pas connues exactement : seul $\vec{\mathbf{x}}$ est connu. On cherche à calculer les valeurs possibles de l'espérance et de la variance :

$$\begin{array}{rcl} E(\vec{\mathbf{x}}) & = & \{E(\vec{x}) \mid \vec{x} \in \vec{\mathbf{x}}\} \\ V(\vec{\mathbf{x}}) & = & \{V(\vec{x}) \mid \vec{x} \in \vec{\mathbf{x}}\} \end{array}$$

Remarque : les fonctions E et V sont continues et on s'intéresse à leurs valeurs sur un compact. Par conséquent, $E(\vec{\mathbf{x}})$ et $V(\vec{\mathbf{x}})$ sont en fait des intervalles que l'on notera $\mathbf{E}$ et $\mathbf{V}$.

Question 3 *Combien vaut $\mathbf{E}$ pour **a)** $n = 10$? **b)** $n = 100$? **c)** $n = 10000$?*

Question à développer pendant l'oral : Montrez qu'il existe un vecteur $\vec{x}$ tel que $V(\vec{x}) = \overline{V}$ et tel que x_i vaut soit $\underline{x}_i$ soit $\overline{x}_i$ pour tout $i \leq n$. Dans quels cas la connaissance de l'espérance E et du milieu $\hat{x}$ vous permet-elle de décider laquelle de ces deux bornes vaut x_i ?

La question précédente montre que l'on peut se contenter de calculer la valeur suivante pour obtenir la borne supérieure de $\mathbf{V}$:

$$\overline{V} = \max\{V(\vec{x}) \mid \forall i,\ x_i \in \{\underline{x}_i, \overline{x}_i\}\}.$$

Question 4 *Combien vaut $\overline{V}$ pour **a)** $n = 10$? **b)** $n = 20$? **c)** $n = 50$?*

Comme vu précédemment, lors de la recherche d'un vecteur $\vec{x}$ où le maximum $\overline{V}$ est atteint, la connaissance de E permet pour certaines valeurs de i de ne considérer qu'un seul des deux cas $x_i = \underline{x}_i$ et $x_i = \overline{x}_i$. Dans la question suivante, une restriction sur l'espérance sera fixée pour augmenter le nombre de valeurs de i pour lesquelles ce choix est possible.

Question 5 *Si l'on suppose lors de chaque choix $x_i \in \{\underline{x}_i, \overline{x}_i\}$ que l'espérance est comprise entre 3 et 4 exclus, combien vaut $\overline{V}$ pour **a)** $n = 100$? **b)** $n = 500$? **c)** $n = 2000$?*

Question 6 *Combien vaut $\overline{V}$ pour **a)** $n = 100$? **b)** $n = 500$? **c)** $n = 2000$?*

Question à développer pendant l'oral : Quelle est la complexité en temps de votre algorithme ?

Question à développer pendant l'oral : Soit $\vec{x} \in \vec{\mathbf{x}}$ un n-uplet tel que $V(\vec{x}) = \underline{V}$. Soit $J = \{j \mid E(\vec{x}) \notin \mathbf{x}_j\}$. Soit $E' = \frac{1}{\#J} \sum_{j \in J} x_j$ avec $\#J$ le cardinal de J. Montrez que $x_j = E(\vec{x}) = E'$ pour tout $j \notin J$. Que pouvez-vous dire de la valeur de x_j quand $j \in J$?

Question 7 *Combien vaut $\underline{V}$ pour **a)** $n = 100$? **b)** $n = 1000$? **c)** $n = 10000$?*

Question à développer pendant l'oral : Quelle est la complexité en temps de votre algorithme ?

3 Matrices d'intervalles

On étend les opérateurs arithmétiques usuels aux intervalles :

$$\begin{aligned}
\mathbf{a} + \mathbf{b} &= \{a + b \mid a \in \mathbf{a} \wedge b \in \mathbf{b}\} \\
\mathbf{a} \times \mathbf{b} &= \{a \times b \mid a \in \mathbf{a} \wedge b \in \mathbf{b}\} \\
\mathbf{a}^2 &= \{a^2 \mid a \in \mathbf{a}\}
\end{aligned}$$

Remarque : $\mathbf{a} + \mathbf{b}$, $\mathbf{a} \times \mathbf{b}$ et $\mathbf{a}^2$ sont des intervalles.

Question 8 *Combien valent **a)** $\mathbf{x}_{12} \times \mathbf{x}_{23}$? **b)** $\mathbf{x}_{33} \times \mathbf{x}_{33}$? **c)** $\mathbf{x}_{33}^2$? **d)** $\mathbf{x}_{44}^2$?*

Question à développer pendant l'oral : Soit $f(x_1, \ldots, x_n)$ une fonction calculant la valeur d'une expression e parenthésée qui n'utilise que les opérateurs arithmétiques $+$, $\times$ et $\square^2$. Qui plus est, chaque variable x_i n'apparaît qu'une seule fois dans l'expression. Par exemple, $x_1+x_4\times(x_3^2+x_2)$ est une telle expression. Substituons à chaque variable x_i l'intervalle $\mathbf{x}_i$ dans e puis évaluons l'expression en utilisant les opérateurs arithmétiques étendus aux intervalles. Montrez que l'intervalle résultant de cette évaluation vaut $f(\mathbf{x}_1, \ldots, \mathbf{x}_n)$.

On notera A une matrice de réels de taille $n\times n$ et ses éléments seront notés a_{ij}, le numéro de ligne i et celui de colonne j allant de 1 à n. Une matrice d'intervalles sera notée $\mathbf{A}$ et ses éléments seront notés $\mathbf{a}_{ij}$. La relation d'appartenance $A \in \mathbf{A}$ signifie que $a_{ij} \in \mathbf{a}_{ij}$ pour tout i et j.

On s'intéresse à l'ensemble $\mathbf{C}$ des valeurs que peut prendre le produit $C = A \times B$ de deux matrices A et B dont les éléments ne sont pas connus exactement :

$$\mathbf{C} = \{A \times B \mid A \in \mathbf{A}, B \in \mathbf{B}\}.$$

Remarque : $\mathbf{C}$ peut être vu aussi bien comme un ensemble de matrices que comme une matrice d'intervalles, l'intervalle $\mathbf{c}_{ij}$ représentant alors l'ensemble des valeurs que peut prendre l'élément c_{ij} de $C = A \times B$ pour $A \in \mathbf{A}$ et $B \in \mathbf{B}$.

Les éléments de la matrice $\mathbf{A}$ sont définis par $\mathbf{a}_{ij} = \mathbf{x}_{(10i+j)}$. Ceux de la matrice $\mathbf{B}$ sont définis par $\mathbf{b}_{ij} = \mathbf{x}_{(15i+2j)}$.

Question 9 *Pour $n = 20$, combien valent **a)** $\mathbf{c}_{23}$? **b)** $\mathbf{c}_{32}$? **c)** $\mathbf{c}_{33}$?*

On s'intéresse maintenant non plus au produit de deux matrices mais au carré D de la matrice A et on cherche à connaître l'ensemble des valeurs $\mathbf{D}$ qu'elle peut prendre :

$$\mathbf{D} = \{A^2 \mid A \in \mathbf{A}\}.$$

Question 10 *Pour $n = 20$, combien valent **a)** $\mathbf{d}_{23}$? **b)** $\mathbf{d}_{32}$? **c)** $\mathbf{d}_{33}$?*

◇
◇ ◇
◇

Fiche réponse type: Variance et intervalles

$\widetilde{u_0}$: 26

Question 1

a) $[11; 63]$

b) $[-80; 33]$

c) $[-22; 92]$

Question 2

a) $-9.5\ (\times 3)$

b) $-10\ (\times 12)$

c) $-10\ (\times 69)$

Question 3

a) $[-21.6; 55.7]$

b) $[-33.51; 33.48]$

c) $[-34.0979; 32.6532]$

Question 4

a) 5366.60

b) 5230.79

c) 4749.60

Question 5

a) 4916.81

b) 4764.98

c) 4892.47

Question 6

a) 4931.33

b) 4820.86

c) 4940.58

Question 7

a) 21.8655

b) 41.9421

c) 29.0627

Question 8

a) $[-3968; 3520]$

b) $[-2318; 3721]$

c) $[0; 3721]$

d) $[1764; 4225]$

Question 9

a) $[-72934; 71570]$

b) $[-47721; 55432]$

c) $[-62446; 76552]$

Question 10

a) $[-72934; 71570]$

b) $[-47721; 55432]$

c) $[-62446; 76552]$

Fiche réponse: Variance et intervalles

Nom, prénom, u_0: ....

Question 1

a)

b)

c)

Question 2

a)

b)

c)

Question 3

a)

b)

c)

Question 4

a)

b)

c)

Question 5

a)

b)

c)

Question 6

a)

b)

c)

Question 7

a)

b)

c)

Question 8

a)

b)

c)

d)

Question 9

a)

b)

c)

Question 10

a)

b)

c)