Université Blaise Pascal

30 janvier 2006

ISIMA

Première année

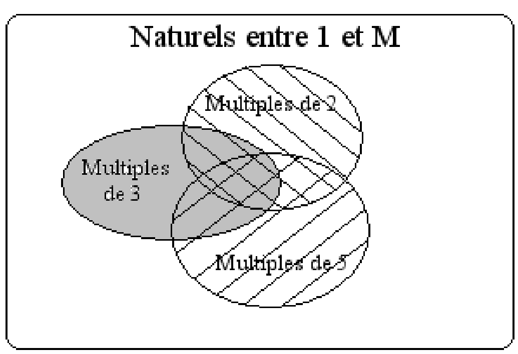
Probabilités et statistiques

Documents et calculatrices autorisées

Pour décider si un nombre entier naturel est composé (c'est à dire produit de deux nombres entiers autres que 1 et -1) ou au contraire premier, on dispose d'un algorithme stochastique $\mathcal{A}$ qui à un naturel n en entrée, associe soit la réponse "n est composé" (ce qui est alors vrai), soit la réponse "n est premier" mais, dans ce dernier cas, on sait qu'avec une certaine probabilité p, ce résultat est faux. On décide de tester cette méthode en lui fournissant en entrée des naturels multiples ni de 2, ni de 3 ni de 5 (ce qui est très rapide à vérifier), selon l'algorithme suivant :

programme test
| paramètre d'entrée : le naturel M
| variable locale : le naturel n
| **répéter**
| | tirer au hasard n entre 1 et M
| **jusqu'à ce que** n ne soit divisible ni par 2, ni par 3, ni par 5.
| **écrire** n et la réponse de $\mathcal{A}(n)$
fin programme

1. (a) Calculer la probabilité, pour M entier naturel très "grand", qu'un entier n tiré selon une loi uniforme entre 1 et M, ne soit multiple ni de 2, ni de 3, ni de 5.



Tournez s'il vous plaît

(b) Quelle est la loi du nombre I de boucles **répéter...jusqu'à...** effectuées dans l'algorithme ci-dessus ?

(c) Quelle est l'espérance mathématique du nombre moyen de boucles **répéter...jusqu'à...** effectuées dans l'algorithme ci-dessus ?

(d) On dispose de n réalisations $(x_1, x_2, ..., x_n)$ de variables aléatoires indépendantes réparties selon une loi géométrique de paramètre q, calculer l'estimateur du maximum de vraisemblance de q. On rappelle qu'une variable aléatoire X suivant une loi géométrique de paramètre q (avec $0 \leq q \leq 1$) ne prend que des valeurs entières k supérieures ou égales à 1, et avec la probabilité : $\mathbb{P}\{X = k\} = (1-q)^{k-1}.q$.

2. On a mesuré la durée d'exécution (en ms), sur une machine donnée, de l'algorithme précédent sur 200 essais. On a trouvé une durée moyenne $\overline{x}$ =13,4 avec un écart-type s =1,2. Donner, au seuil 5% un intervalle de confiance de la durée moyenne d'exécution et de l'écart-type de cette durée. **On précisera soigneusement les hypothèses qu'il est nécessaire de faire pour effectuer ces calculs.** On rappelle que, si une variable aléatoire X est distribuée selon une loi de χ^2 à n degrés de liberté on peut admettre, lorsque n est supérieur ou égal à 100, que $\sqrt{2.X} - \sqrt{2n-1}$ suit une loi normale centrée réduite.

3. On a mesuré également, sur la même machine, la durée d'exécution (en ms) de l'algorithme précédent sur 20 nombres entiers tirés au hasard entre 2^{63} et $2^{127} - 1$. On a trouvé une durée moyenne d'exécution $\overline{x}$ =120,2 avec un écart-type s =5,1. Tester, au seuil 5% l'hypothèse que la durée moyenne de fonctionnement est multipliée par 8, c'est à dire qu'elle vaut : μ=8×13,4=107,2. **On précisera soigneusement les hypothèses qu'il est nécessaire de faire pour effectuer ces calculs.**

4. On veut vérifier statistiquement l'hypothèse que le nombre de nombres premiers inférieurs ou égaux à J est d'environ $\frac{J}{\ln(J)}$. Pour cela, on lance 30000 fois le programme $\mathcal{A}$ sur des entrées entières tirées au hasard selon une loi uniforme entre 1 et 4294967295 et l'on a obtenu la répartition suivante :

de	à	nombre de nombres premiers
2	1048575	2
1048576	16777215	7
16777216	67108863	19
67108864	268435455	88
268435456	1073741823	317
1073741824	4294967295	1090
2	4294967295	nombre total : 1523

Sur les résultats précédents, tester, au seuil 5% le fait que le nombre de nombres premiers inférieurs à J soit $\frac{J}{\ln(J)}$. Pour cela, on pourra, par exemple, tester si la proportion des nombres premiers trouvés entre les entiers a et b ($a < b$) est proportionnelle à $\frac{b}{\ln(b)} - \frac{a}{\ln(a)}$.

5. On a lancé l'algorithme précédent sur des naturels entre 1 et $2^{63} - 1$, jusqu'à obtenir 1000 fois la réponse "n est premier", puis on a vérifié à l'aide d'un algorithme exact et l'on n'a jamais trouvé d'erreur. Donner, au seuil de confiance 1% un majorant de p.