G. Hunault
Licence Informatique

Angers, octobre 1999

Controle continu pour l'Option de Probabilités et Statistiques

Modalités du Controle

Le controle continu se fera sous la forme de deux projets à rendre, l'un concernant la partie Probabilités, l'autre la partie Statistiques. Chaque étudiant doit rendre les deux projets.

La rédaction des deux projets se fera au choix sous *Word* ou sous *Latex*. Les documents seront transmis "de la main à la main" sous forme de fichier informatique sur disquette ou par *mail*. Ces documents devront être remis au plus tard le mardi 14 décembre avant 12 h. On pourra rendre un seul fichier contenant les deux projets ou fournir un fichier .zip ou .tar sans sous-répertoire. Pour éviter toute ambiguité, tous les noms de fichiers correspondront aux initiales de l'élève. Il sera possible de vérifier auprès de l'enseignant le 14 décembre entre 14 et 18 h que les fichiers ont bien été reçus.

1. Projet de Probabilité

Les modèles probabilistes classiques utilisent des tirages, que ce soit avec répétition ou sans répétition. On voudrait simuler la loi binomiale et la loi "poissoniale" par un programme informatique. Une fois entré(s) le(s) paramètre(s) requis, à savoir respectivement n, p ou λ, on fournira au programme un paramètre N qui sera le nombre de valeurs à générer.

1. a) Cherchez à la bibliothèque une ou des méthodes pour simuler ces lois. Trouvez en particulier s'il y a des conditions à imposer sur N. Vous rédigerez soigneusement la méthode, vous citerez vos sources bibliographiques et vous fournirez si possible un exemple, quitte à en recopier d'un ouvrage que vous citerez.

1. b) Ecrire l'algorithme de la méthode retenue et traduisez-le dans un des langages suivants : Awk, Apl, Rexx, C, Perl, Tcl, Java, Pascal. Le programme détectera la loi à modéliser grace au nombre de paramètres. Par exemple si le programme s'appelle simuleLoi, simuleLoi 0.1 80 donnera 80 valeurs correspondant à la loi poissoniale pour le paramètre 0.1. Avant de vérifier la validité de votre simulation, vous fournirez la moyenne et l'écart-type de vos données, que vous comparerez aux valeurs théoriques attendues.

2. Projet de Statistiques

Soit x_i pour i de 1 à N avec N entier strictement positif des données entières correspondant aux indices des modalités d'une variable QL (qualitative). On nomme mode des x_i toute valeur x_j de plus fort effectif, c'est à dire toute valeur la plus fréquente. Par exemple pour `1 2 3 2` le mode est 2, pour `1 2 2 3 3`, il y a deux modes, 2 et 3.

On voudrait écrire un algorithme de calcul du ou des mode(s) d'une variable QL. On nommera $xData[i]$ les données, $xValue[j]$ les différentes valeurs rencontrées, $xEffect[j]$ les fréquences associées. On mettra le(s) modes dans le tableau $xMode[p]$. On notera T le nombre de valeurs distinctes $xData[i]$ et M le nombre de mode(s) trouvé(s).

2. a) Soient

   `1 4 1 2 1 2 1 2 3 1 4 4 2 3 3 3 3 4 4`

   les valeurs du tableau $xData[i]$. Donner les valeurs des différentes variables de l'algorithme une fois l'algorithme exécuté.

2. b) Ecrire l'algorithme qui calcule les différents $xMode[p]$ en respectant la syntaxe algorithmique vue en cours.

2. c) Implémenter l'algorithme dans l'un des langages suivants : Awk, Apl, Rexx, C, Perl, Tcl, Java, Pascal. Pour les programmes compilés, on fournira le code source et un exécutable. On indiquera dans le document d'accompagnement la version du langage et le(s) système(s) d'exploitation utilisé(s). Le programme devra accepter comme paramètre le nom du fichier des données.

Remarques concernant la rédaction

La note sera principalement fonction de la qualité de l'analyse du travail à effectuer, de l'exposé de la méthode et de la clarté de la rédaction. La valeur attribuée au programme ne représentera que 20 % de la note. Aucune interface n'est demandée, les programmes devront être exécutables en ligne de commande.