G. Hunault
Maitrise Sciences de l'Environnement

Janvier 2005

Controle continu pour le module Statistiques

1. Modalités du Controle

Le controle continu se fait sous la forme d'une étude statistique d'un "dossier" de données à rendre qui comprendra une présentation des données, la mise en forme des résultats de l'étude statistique réalisée et le commentaire de ces résultats. La rédaction des documents se fera sous *Word* ou sous *Latex*. Les documents seront transmis sous la forme d'un seul fichier informatique transmis par *mail* à gilles.hunault@univ-angers.fr. Ces documents devront être remis au plus tard le jeudi 31 mars 2005 à 11 h. Les calculs pourront être faits avec *Dbase*, *Excel*, *StatBox*, *Sas* etc., mais dans tous les cas, on rendra au minimum un fichier .dsc et un fichier .xls. Pour éviter toute ambiguité, comme d'habitude, le nom des fichiers correspondra aux initiales de l'élève et le fichier transmis sera de type .zip ou .tar.

La note sera fonction des données traitées, de la qualité de la rédaction, de la justesse des commentaires. L'exposé des résultats devra être clair, facile et agréable à lire, sans faute d'orthographe détectable par un correcteur orthographique et grammatical informatisé.

Les fichiers transmis devront pouvoir être lus et utilisés tels quels à partir du répertoire courant, que ce soit sous *Dos/Windows* ou sous *Unix*.

On pourra être deux par dossier traité.

Il est impératif de faire valider les données par l'enseignant avant de commencer à les traiter. Cela évitera comme chaque année les erreurs de méthodes, les calculs sans aucun sens sur les densités, les moyennes de pourcentages etc.

2. Détail du travail à réaliser

Rien ne remplace la pratique pour les statistiques descriptives ! Aussi, chaque étudiant(e) devra fournir un fichier de données réelles (données de magazines, d'entreprise, etc.), rédiger son descriptif en termes d'unités, de codes, de variables qualitatives et quantitatives voire textuelles et traiter ensuite ce fichier au sens de l'analyse séparée et conjointe. La rédaction pourra se faire sous *Word*, *Excel* ou *Latex* ou un panaché des trois (par exemple les calculs sous *Excel* et la rédaction sous *Word*).

Le fichier de données devra comporter au moins 50 lignes et 10 colonnes statistiques dont au moins 5 variables qualitatives et 5 variables quantitatives. Il sera au minimum listé en annexe dans les documents transmis. Les résultats, triés selon l'ordre convenu en cours, seront commentés. On fournira quelques graphiques pour illustrer les variables les plus marquantes. En cas de découpage en classes pour des variables quantitatives, on justifiera le nombre de classes et les bornes de découpage.

Le même fichier de données peut être traité par plusieurs élèves (2 au maximum). Aucun des dossiers du répertoire k:\stat_ad ne pourra être utilisé comme dossier.

On trouvera sur la page

http://www.info.univ-angers.fr/pub/gh/wstat/redige.htm

quelques conseils pour bien rédiger. La page

http://www.info.univ-angers.fr/pub/gh/Datasets/datasets.htm

contient des références à des sites dans le monde entier qui proposent des jeux de données. Pour envoyer le projet, il faut utiliser la page

http://www.info.univ-angers.fr/pub/gh/wstat/chklprojmse.php

Enfin, ce document est accessible à partir de la page

http://www.info.univ-angers.fr/pub/gh/wstat/statgen.htm