

TP 14 informatique

BCPST 1 2019-2020

Statistique descriptive et applications

Points abordés

- Les bibliothèques mathématiques.
 - Des exemples de statistique.
 - les caractéristiques usuelles en statistique.
-

1 Exercices

- Exercice 1.**
1. Écrire une fonction `moyenne(L)` qui renvoie la moyenne d'une liste L de réels donnée.
 2. Écrire une fonction `variance(L)` qui renvoie la variance d'une liste L de réels donnée.
 3. Écrire une fonction `covariance(L,M)` qui renvoie la covariance de deux listes.
 4. Dédire des fonctions précédentes une fonction `regression(L,M)` qui renvoie un couple de réels (a, b) avec a le coefficient directeur et b l'ordonnée à l'origine de la droite de régression linéaire, où L correspond à la liste des abscisses et M à la liste des ordonnées.

Exercice 2. Grâce aux statistiques, il est possible de décoder des textes qui ont été chiffrés à l'aide des méthodes de César et de Vigenère. Décrivons ces méthodes de chiffrement. Concernant la méthode de César, on procède de la manière suivante : on choisit une clé qui correspond à une lettre. Par exemple, si la lettre clé est le c , alors le a est codé par un c , le b par un d , le c par un e etc. Pour le mot "biologie", on obtient alors "dkqnqikg". La méthode de Vigenère repose sur le même principe en utilisant non pas une lettre clé, mais un mot clé : on code la première lettre du mot à l'aide de la première lettre clé, la deuxième par la deuxième lettre clé..., et si le mot à coder est trop long, on revient au début. Par exemple, si le mot clé est "boum", on obtient pour le mot "biologie" le mot "cwixpucq". Pour le mot "biologie", le b a été codé à l'aide du b , le premier i à l'aide du o ,..., le deuxième o à l'aide du b etc.

À l'aide de l'analyse fréquentielle, il est possible de décrypter un texte ayant été crypté à l'aide de la méthode de César ou de Vigenère. En effet, la répartition des lettres dans la langue française n'a rien d'uniforme. En exploitant cette répartition, il est possible de retrouver la clé ayant permis de coder le texte et donc de le décrypter.

Le premier texte qui nous intéresse a été codé à l'aide de la méthode de César. L'objectif est de retrouver la clé.

1. Récupérer le fichier crypté "crypte1".
2. Écrire une fonction `codage_cesar(texte,cle)` prenant en arguments une chaîne de caractères `texte` et une lettre `cle` et qui renvoie `texte` chiffrée par `cle`. Seules les lettres suivantes de `texte` seront chiffrées :

abcdefghijklmnopqrstuvwxyz

les majuscules, lettres accentuées, symboles de ponctuation etc. seront laissés tels quels.

3. Écrire une fonction `frequence_lettre(texte,l)` qui prend en arguments une chaîne de caractères `texte` et une lettre `l` et qui renvoie le nombre de fois qu'apparaît `l` dans `texte`

4. Déterminer les fréquences d'apparitions de chacune des lettres chiffrées du texte "crypte1".
5. Un texte français comportant en moyenne plus de "e", en déduire une la lettre ayant permis codé ce texte. Á l'aide de `codage_cesar` déchiffrer le texte "crypte1".

Le deuxième texte ("crypte2"), lui a été codé à l'aide de la méthode de Vigenère avec un mot de 6 lettres. Dans ce contexte, une lettre du mot clé n'est pas nécessairement codé par une même lettre du mot clé. Par contre, la première lettre du texte et la septième le sont. Ainsi, il suffit de déterminer les fréquences dans le texte toutes les 6 lettres.

1. Écrire une fonction qui prend en argument un texte et une clé et qui renvoie le texte chiffré à l'aide de la méthode de Vigenère.
2. En adaptant la méthode utilisée précédemment, déterminer la clé.
3. En déduire le texte décrypté.