

## TP noté 1 : Probabilités et statistiques descriptives

Durée : 1 heure.

Consignes :

- Les calculs éventuels sont à mener avec le logiciel R.
- Vous devez a minima rendre sur Moodle un fichier `.R` contenant vos calculs. Pensez à le nommer de façon explicite, et pour faciliter la correction, la première ligne du fichier doit être `#votre nom-prenom`.
- De préférence, ne rendez que le fichier `.R` en y intégrant les réponses littérales sous forme de commentaires (c'est-à-dire précédées de `#`). Si toutefois vous préférez autre chose (traitement de texte, photo d'un document manuscrit) n'oubliez pas de tout déposer sur Moodle.
- Le barème est uniquement donné à titre indicatif.

**Exercice 1 : Probabilités discrètes** Dans une population de 100 habitants, la probabilité pour qu'une personne souhaite être vaccinée contre la Covid est de 0.8.

On note  $X$  le nombre de personnes souhaitant être vaccinées.

1. Quelle est la loi de  $X$ ? On précisera les hypothèses faites. [2 points]
2. Déterminer la probabilité qu'il y ait entre 75 et 85 personnes demandant à être vaccinées. [1 point]  
(Le "entre" est à prendre au sens large)
3. Déterminer le nombre minimal de doses de vaccin à commander pour que la probabilité qu'on vienne à en manquer soit inférieure à 0.05. [1.5 point]

### Exercice 2 : Probabilités continues

La loi Gamma est une loi de probabilités à valeurs positives qui dépend de deux paramètres  $\alpha$  et  $\beta$ .

On la note  $\Gamma(\alpha, \beta)$ .

Le logiciel R la connaît sous l'appellation `gamma`.

1. Représenter graphiquement une densité de la loi  $\Gamma(1, 2)$  ainsi qu'une densité de la loi  $\Gamma(3, 2)$ .  
(Vous pouvez les faire figurer sur un ou sur deux dessins.) [1 point]
2. Déterminer, pour  $X$  de loi  $\Gamma(3, 2)$ , la probabilité  $P(1 \leq X \leq 4)$ . [1 point]
3. Simuler un échantillon de 1000 réalisations de la loi  $\Gamma(3, 2)$  et en déduire une estimation de l'espérance de cette loi. [2 points]

### Exercice 3 : Statistique descriptive

1. Importer le fichier `Yeux_cheveux.txt`. [0.5 point]
2. Combien y a-t-il d'individus? De variables? De quel type sont les variables? [1 point]
3. Faire un résumé numérique et une ou plusieurs représentations graphiques adéquates de la variable "cheveux". Commenter les observations en une ou deux phrases. [3 points]

### Exercice 4 : Statistique descriptive.

1. Importer le fichier `Nutrition.csv`, le séparateur étant ";". [0.5 point]

Sur la page suivante vous trouverez un descriptif de ce fichier de données.

2. Combien y a-t-il d'individus? [0.5 point]
3. Donner, parmi les données, un exemple de variable qualitative nominale, un exemple de variable qualitative ordinale, un exemple de variable quantitative discrète et un exemple de variable quantitative continue. [1 point]
4. Représenter sur un histogramme les tailles des individus. [1 point]
5. Effectuer un résumé statistique des variables poids et taille. On précisera notamment le coefficient de variation de chacune des deux variables. Commenter en une ou deux phrases brèves ces coefficients de variation. [3 points]
6. On rappelle que pour placer deux graphiques sur une seule fenêtre en R, on peut faire précéder les tracés de `layout(matrix(1:2,1,2))`.  
Tracer côte à côte les boîtes à moustaches correspondant au nombre de tasses de café prises quotidiennement par les femmes et celles prises par les hommes.  
Commenter succinctement. [2 points]

Le tableau ci-dessous liste les variables du fichier utilisé dans l'exercice 4.

### Présentation :

Un échantillon de personnes âgées résidant à Bordeaux (Gironde, France) a été interrogé en 2000 dans le cadre d'une enquête nutritionnelle. L'échantillon est constitué de 226 sujets.

### Variables et codage :

Description	Unité ou Codage	Variable
Sexe	2=Femme ; 1=Homme	sexe
Situation familiale	1=Vit seul 2=Vit en couple 3=Vit dans sa famille 4=Autre type de cohabitation	situation
Consommation journalière de thé	Nombre de tasses	the
Consommation journalière de café	Nombre de tasses	cafe
Taille	Cm	taille
Poids	Kg	poids
Âge le jour de l'entretien	Années	age
Consommation de viande	0=Jamais 1=Moins d'une fois par semaine 2=Une fois par semaine 3=2/3 fois par semaine 4=4/6 fois par semaine 5=Tous les jours	viande
Consommation de poisson	Idem	poisson
Consommation de fruits crus	Idem	fruit_crus
Consommation de fruits et légumes cuits	Idem	fruit_legume_cuits
Consommation de chocolat	Idem	chocol
Matière grasse préférentiellement utilisée pour la cuisson	1=Beurre 2=Margarine 3=Huile d'arachide 4=Huile de tournesol 5=Huile d'olive 6=Mélange d'huile (type Isio4) 7=Huile de colza 8=Graisse de canard ou d'oie	matgras