

Statistiques élémentaires - Indicateurs de position (2^e partie)

- Préambule :**
0. (Rappel) Ouvrir une session en tapant **1L2** comme nom d'utilisateur – mot de passe « **jml** ».
 1. Ouvrir le classeur Excel à votre nom dans le dossier de **1L2_Ganesh / 1L2_2005_G1 ou G2**.
 2. Insérer une nouvelle feuille (Menu **Insertion / Feuille**) et la placer *après* les précédentes.
 3. Renommer cette feuille **TD7 Stat.2** comme pour les TD précédents.

I – Étude statistique d'un jeu de dés : on lance 200 fois une paire de dés à 6 faces et l'on note la somme des points.

Population / Échantillon : 200 lancers d'une paire de dés à 6 faces.

Effectif total de l'échantillon : $N = 200$

Caractère étudié : les faces.

Valeur du caractère : la somme des faces.

Classes : notées x_i , elles sont représentées par les nombres entiers de 2 à 12.

Série statistique donnée (ou engendrée aléatoirement par l'ordinateur) : 200 nombres entiers compris entre 2 et 12.

Effectif de chaque classe x_i noté n_i (nombre d'occurrences de la valeur x_i dans la série initiale)

Fréquence de chaque classe : $f_i = n_i / N$

1°) Construire le tableau suivant dans la feuille de calcul en utilisant les formules adéquates pour remplir les cellules automatiquement.

On observera en particulier que la **moyenne** s'obtient directement en faisant la somme des produits $x_i \cdot f_i$ dans la 4^e colonne.

On vérifiera que l'on obtient bien un total de 100% dans la colonne des fréquences.

La 5^e colonne contient les **fréquences cumulées** c'est à dire la somme des fréquences correspondant à toutes les classes inférieures ou égales à celles de la ligne correspondante.

On doit donc automatiquement obtenir la valeur 100% dans la ligne correspondant à la dernière valeur du caractère.

Ainsi cette dernière colonne permet-elle de déterminer approximativement la valeur de la **médiane** en recherchant quelle valeur du caractère correspond à environ 50% de l'effectif.

Classes	Effectifs	Fréquences	Produits	Fréqu. Cum.
x_i	n_i	f_i (%)	$x_i \cdot f_i$	Σf_i (%)
2	3			
3	8			
4	11			
5	25			
6	22			
7	33			
8	35			
9	27			
10	20			
11	10			
12	6			100 %
Totaux	N = 200	100%	Moyenne :	Médiane :

2°) Construire l'**histogramme des fréquences** dans la feuille de calcul, en n'omettant pas de mettre les valeurs des classes en abscisse et en réduisant les ordonnées au mieux. Indiquer la **moyenne** trouvée sur le graphique obtenu.

3°) Construire la **courbe des fréquences cumulées** (nuage de points reliés par des segments de droite) en n'omettant pas de mettre les valeurs des classes en abscisse, les ordonnées étant nécessairement de 0 à 100%.

Indiquer la **médiane** ainsi que le 1^{er} et le 3^e **quartile** sur le graphique obtenu et représenter la « **boîte à moustaches** ».

II – Construction automatique aléatoire d'un nouvel échantillon et calcul automatique des effectifs, des fréquences, de la moyenne et de la médiane à l'aide des formules statistiques du tableur.

On utilise la fonction **ALEA()** pour construire une plage de 200 valeurs comprises entre 2 et 12.

Pour cela on écrit la formule suivante dans une cellule : `= ENT(6*ALEA()) + 1 + ENT(6*ALEA()) + 1` puis on étend cette formule sur une plage de 200 cellules (par exemple en prenant 20 lignes et 10 colonnes).

On donne alors le nom « **série** » à cette plage [Menu : Insertion / NOM]

2°) Pour déterminer le nombre d'occurrences d'un nombre donné dans une **série brute**, on utilise la fonction suivante :

`=NB.SI(série ; valeur)` qui calcule le nombre d'occurrences de la **valeur** indiquée (réf cellule) dans la **série** indiquée (Références fixes de la plage de cellule contenant la série générée par la fonction **ALEA**)

Pour utiliser efficacement cette fonction on construit d'abord une colonne contenant les valeurs x_i des classes, dont les éléments serviront de référence pour les valeurs.

3°) Compléter alors le tableau obtenu comme dans le § I puis construire l'**histogramme** et la **courbe des fréquences cumulées** correspondants. On peut ainsi utiliser les fonctions `= MOYENNE(série)` et `= MEDIANE (Série)` , pour vérifier les résultats obtenus dans le tableau, ainsi que les fonctions

`=QUARTILE(Série ;1)` et `=QUARTILE (Série ;3)` pour obtenir le 1^{er} et le 3^e quartile.

4°) Si la feuille de calcul est configurée en mode « calcul sur ordre (voir Préférences /etc...) alors en appuyant sur la touche **F9** on voit se modifier automatiquement toutes les données et les paramètres de position. le tableau des valeurs et l'histogramme associé ainsi que la courbe des fréquences cumulées.