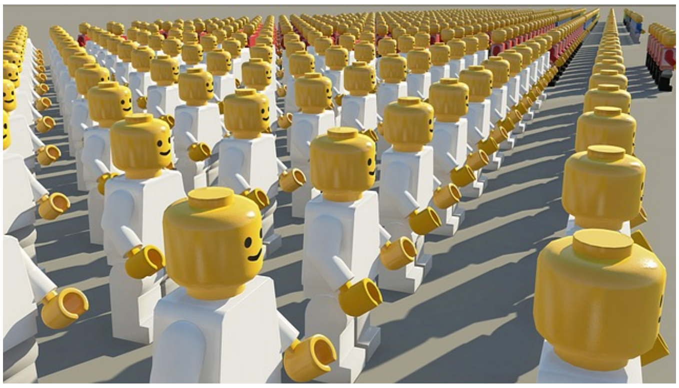
Chapitre 5 - Pourcentage - Statistique

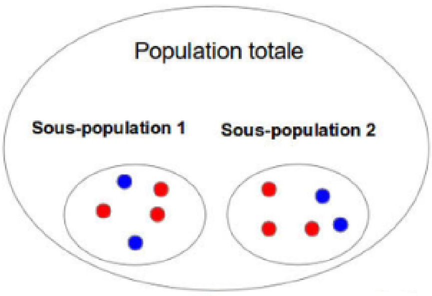


1 Proportion et Pourcentage

1.1 Population et sous-population

Définition

- **élément :** 1 élément (unitaire)
- **population (totale)** :
  tous les éléments
- **sous-population** :
  une sous-partie de la population



Définition et Propriété

n est le nombre d'individu de la sous-population et N la population totale (P_T)

- **proportion d'1 sous-population** : $\boxed{p = \dfrac{n}{N}}$
- **pourcentage d'1 sous-population** : $\boxed{\%p = \dfrac{n}{N} \times 100}$
- **pourcentage de pourcentage** :
  Si (P_1 est $\%p_1$ de la P_T) et (P2 est $\%p_2$ de P_1) Alors $\boxed{\%p = \%p_1 \times \%p_2 \div 100 \text{ de la } P_T}$

2 Variations d'1 quantité

Variation Absolue (V_A)ou Relative (t)

Définition

V une quantité qui varie (par exemple) au cours du temps
V_I la valeur initiale ; V_F la valeur finale

- $V_A = V_F - V_I$ qui possède la même unité que V_I ou V_F
- $t = \frac{\%}{100} = \frac{V_F - V_I}{V_I}$ (qui est sans unité) : proportion de variation "relativement" à V_I
- $\boxed{V_F = (1+t)V_I = (1 + \frac{\%}{100})V_I}$
- cette dernière formule est la plus importante :
  - j'achète un djinn à 100 qui est e remise cumulée de 30% puis de 20%, combien je paie ?
  - est-il plus intéressant d'acheter un article qui est monté de 10% puis à baisser de 10% ou l'inverse ?

3 Série statistique - Indicateur

3.1 Vocabulaire et Position du Problème

Définition

- on a une population
- cette population est constitué d'individus
- chaque individu possède des caractères (qualitatif ou quantitatif)
- un caractère quantitatif est discret (valeur) ou continue (positionné dans une classe)
- une série statistique est la liste associée à l'étude d'un caractère d'une population donnée

Exemple 1 : couleurs des voitures vendues

- population : voitures
- individu : 1 voiture
- caractère (qualitatif) étudié : couleur (vert, rouge, bleu, ...)
- série statistique : {bleu, vert, bleu, rouge, ...}

Exemple 2 : taille de chaussures vendues

- population : chaussure
- individu : 1 taille
- caractère (quantitatif discret) étudié : taille (36, 37, 38, ...)
- série statistique : {38, 43, 36, 39, 38 ...}

Exemple 3 : durée d'appels téléphonique

- population : appel téléphonique
- individu : 1 appel
- caractère (quantitatif continue) étudié : durée d'un appel ; de 0 à 5h
- série statistique : {10s, 3min20s, 1h20min ...}

Intérêt des Indicateurs et Calcul

- l'intérêt est de résumer l'information
(imaginez si votre série statistique possède 1 million de valeurs ...)
- tous les calculs sont en général faits par votre calculatrice ou par ordinateur
il faut donc apprendre à se servir de la calculatrice

3.2 Indicateur de Position

Définition

- **moyenne** (3 formules à connaître)
  - normale : $M = \dfrac{x_1 + x_2 + ...}{n} = \dfrac{1}{n} \times \displaystyle\sum_{i=1}^{n} x_i$
  - pondérée : $M = \dfrac{n_1 \times x_1 + n_2 \times x_2 + ...}{n} = \dfrac{1}{n} \times \displaystyle\sum_{i=1}^{k} n_i \times x_i$ où $\displaystyle\sum_{i=1}^{k} n_i = n$
  - fréquence : $M = f_1 \times x_1 + f_2 \times x_2 + ... = \displaystyle\sum_{i=1}^{k} f_i \times x_i$ où $\displaystyle\sum_{i=1}^{k} f_i = 1$
  - lien entre pondérée et fréquence : $\dfrac{n_i}{n} = f_i$
- **médiane** : valeur qui coupe la série statistique en 2 morceaux de même taille **une fois triée par ordre croissant** (pour un nombre pair de valeur, prendre la moyenne des 2 valeurs)
- **mode** : valeur la plus fréquente
- **étendue** = max - min

remarque HP : utilisation méconnue de la moyenne pondérée

- tout nombre peut être construit à partir de sa structure de nombres premiers
- par exemple, $18\,900 = 2^2 \times 3^3 \times 5^2 \times 7$
- grâce à cette écriture unique, on définit la notion de radical et de qualité d'un nombre :
  - $rad(18\,900) = 2 \times 3 \times 5 \times 7 = 210$; c'est la base du nombre
  - $qualite(18\,900) = \dfrac{\ln 18\,900}{rad(18\,900)} = \dfrac{2\ln 2 + 3\ln 3 + 2\ln 5 + \ln 7}{\ln 2 + \ln 3 + \ln 5 + \ln 7}$; c'est la hauteur moyenne du nombre
  - ceci débouche sur un déplus important problème mathématique non prouvé actuellement : le problème A B C - écrit ou vidéo

3.3 Indicateur de Dispersion

Définition

- Q_1 : **première** valeur de la série qui permet d'avoir à sa gauche (Q_1 inclus) au moins 25% des valeurs
- Q_3 : **première** valeur de la série qui permet d'avoir à sa gauche (Q_3 inclus) au moins 75% des valeurs
- **min** , **max**
- **variance** = moyenne des carrés des écarts = $\dfrac{1}{n} \times \displaystyle\sum_{i=1}^{n} (x_i - M)^2$
  où M est la moyenne et n l'effectif total de la série
- **écart type** = racine de la variance = $\sqrt{variance}$
  traduit comment cela "bouge" autour de la moyenne

3.4 Un exemple

Analyser la série suivante

- sujet - corrigé

4 Une activité rigolote

- interprétation de la notion de pourcentage
- ajouter des pourcentages

5 Un peu de python

5.1 Algorithme 1 à chercher : pour des données réelles ou issues dune simulation, lire et comprendre une fonction écrite en Python renvoyant la moyenne m, l'écart type s, et la proportion déléments appartenant à [m-2s,m+2s]

```python
from math import sqrt
def stats(tableau):
        n=len(tableau) #On mesure la taille du tableau
        m=sum(tableau)/n #calcul de la moyenne
        tableau2=[(x-m)**2 for x in tableau] #On crée le tableau des carrés des écarts à la moyenne
        variance=sum(tableau2)/n #Calcul de la variance
        s=sqrt(variance) #Calcul de l'ecartype
        a=m-2*s
        b=m+2*s
        compteur=0
        for i in range(n):
                if tableau[i]<=b and tableau[i]>=a:
                        compteur+=1
        proportion=compteur/n
        return [m,s,proportion] #On renvoie un tableau qui contient la moyenne, l'écart-type et la proportion déléments appartenant à [m-2s,m+2s].
```

Après avoir compilé le code, on tape dans la console :

- tableau=[i**2 for i in range(100)]
- stats(tableau)

5.2 Algorithme 2 à chercher : lire et comprendre une fonction Python renvoyant le nombre ou la fréquence de succès dans un échantillon de taille n pour une expérience aléatoire à deux issues

```python
from random import*
def nb_succes(n,p):
    #n est le nombre de simulations
    #p est la probabilité testée
    c=0
    for k in range(1,n+1):
        t=random()
        if t<p:
            c=c+1
    return c
```