Statistiques & Probabilités

A) Les Statistiques descriptives

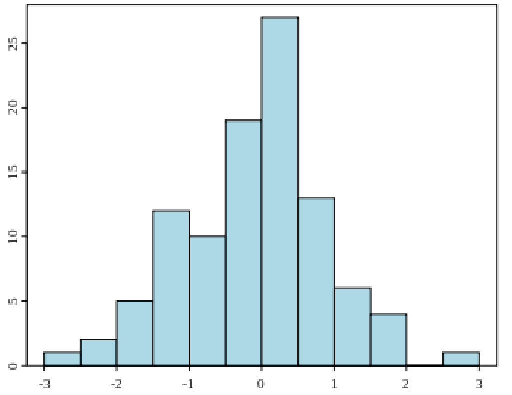
1) Représentations graphiques

Voici les séries de notes obtenues par 3 élèves :

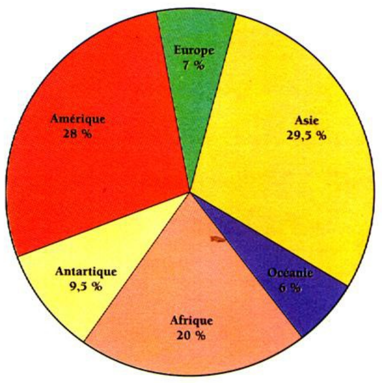
- Jérôme : 4 ; 6 ; 18 ; 7 ; 17 ; 12 ; 12 ; 18
- Bertrand : 13 ; 13 ; 12 ; 10 ; 12 ; 3 ; 14 ; 12 ; 14 ; 15
- Julie : 15 ; 9 ; 14 ; 13 ; 10 ; 12 ; 12 ; 11 ; 10

Construire un graphique permettant de comparer rapidement ces 3 élèves
Les différents <u>types de graphiques</u> sont :

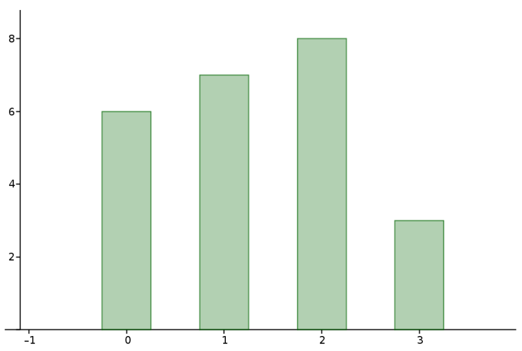
les ***histogrammes***



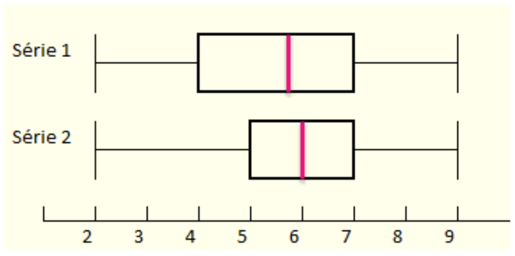
les ***diagrammes circulaires***



les ***diagrammes en barres***



les ***diagrammes en boîtes***



2) Les moyennes

Définition : Soit $x_1, x_2, x_3, \ldots, x_p$ valeurs pondérées par les coefficients $n_1, n_2, n_, \ldots, n_p$ alors la ***moyenne arithmétique pondérée*** est $\overline{x} = \frac{1}{N} \cdot \sum_{i=1}^{p} n_i \cdot x_i$

où $N = \sum_{i=1}^{p} n_i$ correspond à la somme des coefficients

<u>*exemple*</u> : Voici les notes obtenues par 3 élèves du type $(x_i ; n_i)$

- Jérôme : (4;1) ; (6;2) ; (18;1) ; (7;1) ; (17;2) ; (12;1) ; (12;2) ; (18;1)
- Bertrand : (13;1) ; (13;2) ; (12;1) ; (10;2) ; (12;1) ; (3;2) ; (14;1) ; (12;1) ; (14;1) ; (15;1)
- Julie : (15;1) ; (9;2) ; (14;1) ; (13;1) ; (10;2) ; (12;1) ; (12;1) ; (11;1) ; (10;2) ; (11;1)

Comparer ces 3 élèves en calculant leur moyennes arithmétiques respectives

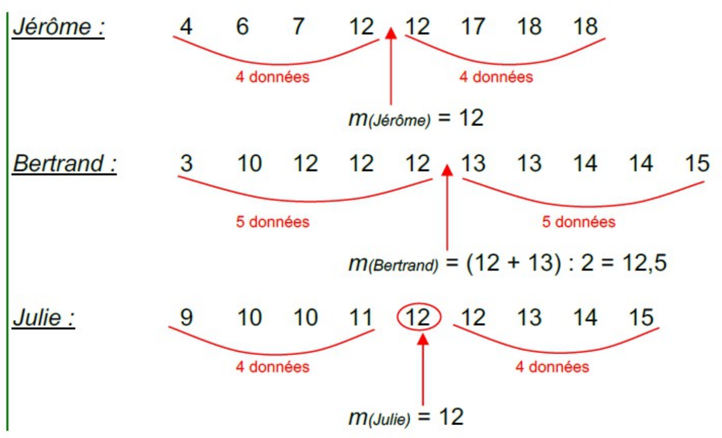
3) Les quartiles

Méthode : *Calculer une médiane*
Vidéo : https://youtu.be/kr90dXv0NFY

Voici les séries de notes obtenues par 3 élèves :

- Jérôme : 4 ; 6 ; 18 ; 7 ; 17 ; 12 ; 12 ; 18
- Bertrand : 13 ; 13 ; 12 ; 10 ; 12 ; 3 ; 14 ; 12 ; 14 ; 15
- Julie : 15 ; 9 ; 14 ; 13 ; 10 ; 12 ; 12 ; 11 ; 10

Calculer la médiane pour Jérôme, Bertrand et Julie.



Définition : Soit $x_1, x_2, x_3, \dots, x_p$ valeurs pondérées par les coefficients $n_1, n_2, n, \dots, n_p$ alors on définit les 3 quartiles de la série :

- le 1er quartile Q_1 correspond à la valeur de $x_{N/4}$
- le 2eme quartile Q_2 correspond à la valeur de $x_{N/2}$
- le 3eme quartile Q_3 correspond à la valeur de $x_{3N/4}$

<u>*exemples*</u> : déterminer les quartiles des 4 séries ci-dessous

Série **A**

x_i	7	8	9	10	12	14	15	17
n_i	2	3	1	4	2	1	3	1

Série **B**

x_i	7	8	9	10	12	14	15	17
n_i	1	3	4	5	1	1	2	1

Série **C**

x_i	7	8	9	10	12	14	15	17
n_i	2	4	1	3	1	4	1	3

Série **D**

x_i	7	8	9	10	12	14	15	17
n_i	3	2	2	3	4	3	2	1

<u>*Réponses*</u> : on obtient les résultats suivants

- Série **A** : $Q_1 = 8$; $Q_2 = 10$; $Q_3 = 14{,}5$
- Série **B** : $Q_1 = 9$; $Q_2 = 10$; $Q_3 = 12$
- Série **C** : $Q_1 = 8$; $Q_2 = 10$; $Q_3 = 14$
- Série **D** : $Q_1 = 8{,}5$; $Q_2 = 10{,}5$; $Q_3 = 14$

<u>*Rque*</u> : le ***2ème quartile*** s'appelle aussi la ***médiane*** puisque $\frac{N}{2} = \frac{2N}{4}$

4) Variance & écart-type

Définition : Soit $x_1, x_2, x_3, \dots, x_p$ valeurs pondérées par les coefficients $n_1, n_2, n, \dots, n_p$ alors la Variance est $V = \frac{1}{N} \cdot \sum_{i=1}^{p} n_i . (x_i - \overline{x})^2$

<u>*exemple*</u> : on donne la série $(x_i ; n_i)$ suivante

x_i	6	7	9	11	12	13	15	18
n_i	2	3	2	5	2	2	1	1

On vérifie que la moyenne est $\overline{x} = 10{,}5$
On calcule alors les "écarts" $e_i = x_i - \overline{x}$ entre les valeurs et la moyenne

e_i	$-4{,}5$	$-3{,}5$	$-1{,}5$	$0{,}5$	$2{,}5$	$3{,}5$	$5{,}5$	$8{,}5$
n_i	2	3	2	5	2	2	1	1

Expliquez pour quelle raison un simple calcul de la moyenne des "écarts" ne correspond pas à un bon indicateur de dispersion

On calcule alors la moyenne des carrés des "écarts" qui donne la variance :

$$V = \frac{1}{18} \times (2 \times (-4{,}5)^2 + 3 \times (-3{,}5)^2 + \dots + 1 \times (8{,}5)^2) \simeq 9{,}81$$

Propriété : Soit $x_1, x_2, x_3, \dots, x_p$ valeurs pondérées par les coefficients $n_1, n_2, n, \dots, n_p$ alors la Variance est $V = \frac{1}{N} \cdot \sum_{i=1}^{p} n_i . x_i^2 - (\overline{x})^2$

<u>*exemple*</u> : avec la série précédente, on obtient : $V = \frac{2161}{18} - (10{,}5)^2 \simeq 9{,}81$
ainsi, cette méthode parait plus efficace pour calculer une variance

Définition : Soit $x_1, x_2, x_3, \dots, x_p$ valeurs pondérées par les coefficients $n_1, n_2, n, \dots, n_p$ alors l'écart-type est $\sigma = \sqrt{V}$

<u>*Rque*</u> : l'écart-type correspond alors à la moyenne quadratique des valeurs de la série $(x_i ; n_i)$ soit $\sigma = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{p} n_i . (x_i - \overline{x})^2}$; on trouve $\sigma = \sqrt{9{,}81} \simeq 3{,}13$

<u>*exercice*</u> : On reprend les 4 Séries **A, B, C, D** précédentes

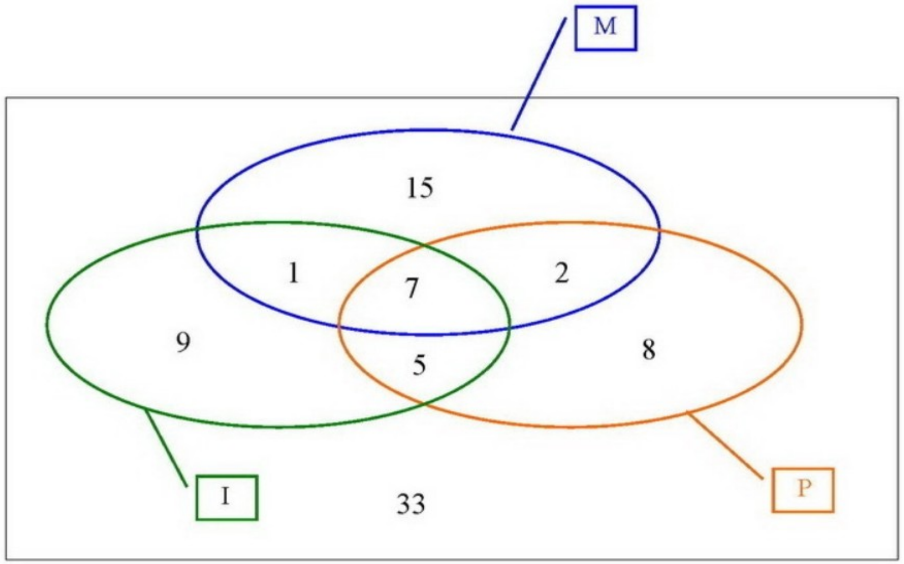
a) Calculer les moyennes des 4 séries
b) Calculer les variances des 4 séries
c) Calculer les écarts-type des 4 séries
d) Interpréter les résultats

B) Les Probabilités

1) Les diagrammes de Venn

exemple : Dans une classe de 80 élèves on sait que :

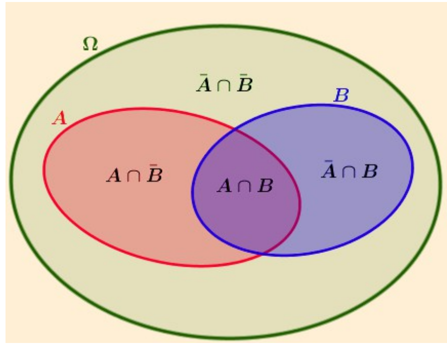
- 22 élèves pratiquent de l'informatique
- 25 élèves pratiquent de la musique
- 22 élèves pratiquent de la photo
- 7 élèves pratiquent les 3 activités à la fois
- 9 élèves pratiquent de la musique et photo
- 12 élèves pratiquent de l'informatique et de la photo
- 15 élèves pratiquent seulement de la musique
- 33 élèves ne pratiquent aucune activité



a) Calculer la probabilité qu'un élève
  - pratique de la musique
  - pratique de l'informatique
  - pratique de la photo

b) Calculer la probabilité qu'un élève
  - pratique de la musique et de l'informatique
  - pratique de l'informatique et de la photo
  - pratique de la photo et de la musique
  - pratique les 3 activités
  - ne pratique aucune activité

Définition : Dans le cas d'une équiprobabilité de toutes les éventualités la probabilité d'un événement A est égale à $P(A)=\frac{Card(A)}{Card(\Omega)}$ où Ω représente l'Univers des probabilités



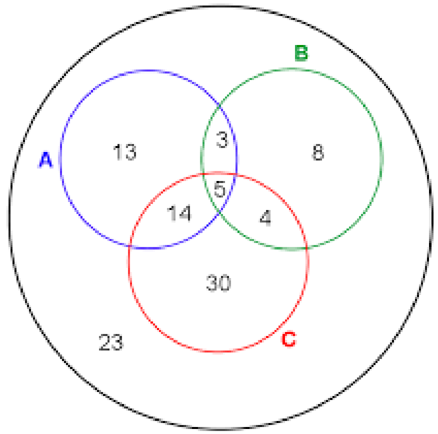
Propriétés : Pour tout événement $A, B\subset\Omega$,

- $0\leqslant P(A)\leqslant 1$
- $P(\Omega)=1$ et $P(\emptyset)=0$
- $P(\overline{A})=1-P(A)$
- $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

Notations :

- $A\cap B$ signifie « A **et** B »
- $A\cup B$ signifie « A **ou** B »
- $\overline{A}$ signifie « **contraire** de A »

exercice : On donne le diagramme de Venn



Calculer les probabilités suivantes :

a) $P(A), P(B), P(C)$

b) $P(A\cap B), P(A\cap C), P(B\cap C)$

c) $P(A\cup B), P(A\cup C), P(B\cup C)$

d) $P(A\cap B\cap C)$

e) $P(\overline{A}\cap B), P(A\cap\overline{C}), P(\overline{B}\cap\overline{C})$

solutions :

$Card(\Omega)=5+14+4+3+13+8+30+23=100$

$P(A)=\frac{13+3+5+14}{100}=0{,}35$; $P(B)=\frac{3+5+4+8}{100}=0{,}2$;

$P(C)=\frac{5+14+4+30}{100}=0{,}53$; $P(A\cap B)=\frac{5+3}{100}=0{,}08$;

$P(A\cap C)=\frac{14+5}{100}=0{,}19$; $P(B\cap C)=\frac{5+4}{100}=0{,}09$; $P(A\cap B\cap C)=0{,}05$;

$P(\overline{A}\cap B)=0{,}12$; $P(A\cap\overline{C})=0{,}16$; $P(\overline{B}\cap\overline{C})=0{,}36$;

$P(A\cup B)=0{,}35+0{,}20-0{,}08=0{,}47$; $P(A\cup C)=0{,}35+0{,}53-0{,}19=0{,}69$;

$P(B\cup C)=0{,}20+0{,}53-0{,}09=0{,}64$

ou $P(B\cup C)=1-P(\overline{B}\cap\overline{C})=1-0{,}36=0{,}64$

2) Les tableaux croisés

exemple : À l'occasion d'une cérémonie, un pâtissier confectionne un assortiment de 180 gâteaux composé d'éclairs au chocolat, d'éclairs au café, de religieuses au chocolat et de religieuses au café. On a donc $Card(\Omega)=180$

Les deux tiers de ces pâtisseries sont des éclairs. On sait également qu'il y a 100 gâteaux au chocolat parmi lesquels un quart sont des religieuses.

1) À partir des indications de l'énoncé, compléter le tableau suivant :

	Chocolat	Café	Total
Éclairs			
Religieuses			
Total			**180**

2) Antoine choisit au hasard un gâteau parmi toutes les pâtisseries.
Quelle est la probabilité qu'il s'agisse :
   a) d'une pâtisserie au chocolat ?
   b) d'une religieuse ?
   c) d'une pâtisserie au café ?

3) déterminer la probabilité qu'il s'agisse :
   a) d'un éclair au chocolat ?
   b) d'un éclair ou d'une pâtisserie au chocolat ?
   c) D'une religieuse au café ?
   d) d'une religieuse ou d'une pâtisserie au café ?

Solution : On nomme les événements suivants

- H : "le gâteau est au chocolat"
- C : "le gâteau est au café"
- E : "le gâteau est un éclair"
- R : "le gâteau est une religieuse"

Le tableau complété est le suivant :

	Chocolat	Café	Total
Éclairs	75	45	120
Religieuses	25	35	60
Total	100	80	**180**

On obtient les probabilités suivantes :

$$P(H)=\frac{100}{180}=\frac{5}{9}\simeq 0{,}55 \quad ; \quad P(R)=\frac{30}{180}=\frac{1}{3}\simeq 0{,}33 \quad ; \quad P(C)=\frac{80}{180}=\frac{4}{9}\simeq 0{,}44$$

On obtient les probabilités suivantes :

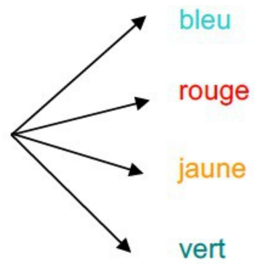
$$P(E\cap H)=\frac{75}{180}=\frac{5}{12}\simeq 0{,}42 \quad ; \quad P(E\cup H)=\frac{75+45+25}{180}=\frac{145}{180}=\frac{39}{36}\simeq 0{,}81$$

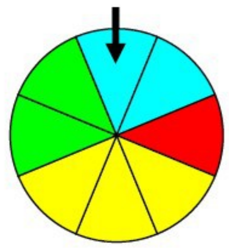
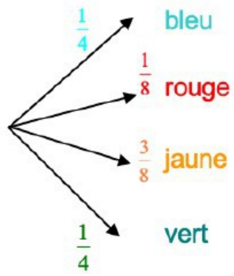
$$P(R\cap C)=\frac{35}{180}=\frac{7}{36}\simeq 0{,}19 \quad ; \quad P(R\cup C)=\frac{45+35+25}{180}=\frac{105}{180}=\frac{7}{12}\simeq 0{,}58$$

3) Les arbres pondérés

<u>les arbres à « 1 génération » :</u>

exemple : Lorsqu'on fait tourner la roue, quatre issues sont possibles.
On le schématise sur l'arbre des possibles :



ainsi $P(B)=0{,}25\,;P(R)=0{,}125\,;P(J)=0{,}375\,;P(V)=0{,}25$

Méthode : *Dénombrer pour calculer une probabilité*

- **Vidéo :** https://youtu.be/d6Co0q01QH0
- **Vidéo :** https://youtu.be/5ZNYG3e2g_k

exemple : On considère l'expérience aléatoire suivante :
On tire une carte dans un jeu de 32 cartes.
Soit ***E*** l'événement : « On tire un as ».
Quelle est la probabilité que l'événement ***E*** se réalise ?

Il a 32 issues possibles car il existe 32 façon différentes de tirer une carte.
L'événement ***E*** possède 4 issues possibles : As de cœur, as de carreau, as de trèfle et as de pique.

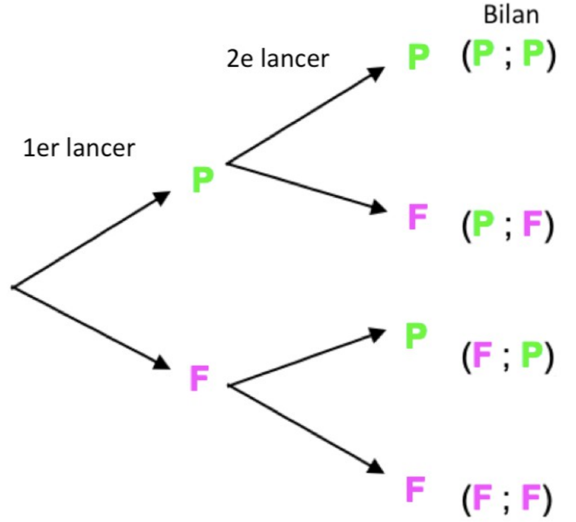
La probabilité que l'événement ***E*** se réalise est égale à : $P(E)=\frac{4}{32}=\frac{1}{8}=0{,}125$

les arbres à « 2 générations » :

exemple : On lance deux fois de suite une pièce de monnaie. Il s'agit d'une expérience aléatoire à deux épreuves.
Soit ***E*** l'événement : « On obtient au moins une fois la face PILE. »
Calculer P(***E***) en utilisant un arbre des possibles.

On construit un arbre des possibles présentant les résultats possibles aux deux épreuves de l'expérience.

- Le 1er niveau de l'arbre correspond les issues du 1er lancer (1ère épreuve).
- Le 2e niveau de l'arbre correspond les issues du 2e lancer (2e épreuve).



On compte **4 issues** en tout : (P ; P), (P ; F), (F ; P) et (F ; F).
L'événement ***E*** possède **3 issues** : (P ; P), (P ; F) et (F ; P).

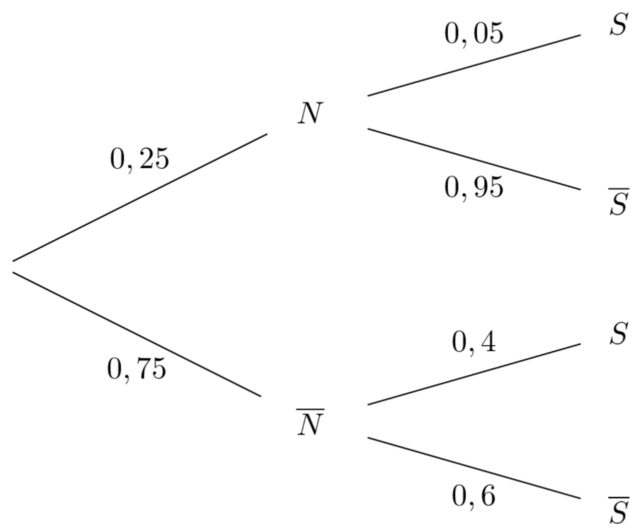
La probabilité que l'événement ***E*** se réalise est donc égale à $P(E)=\frac{3}{4}=0{,}75$

Il y a donc trois chances sur quatre d'obtenir au moins une fois « PILE » lorsqu'on lance deux fois de suite une pièce de monnaie.

Propriétés : On donne un arbre pondéré à n générations ($n\geqslant 2$)

- À partir d'un même nœud, la ***somme des probabilités*** est égale à **1**.
- Pour calculer la probabilité d'un chemin, on ***multiplie*** les probabilités des branches de ce chemin.
- La probabilité d'un événement associé à plusieurs chemins est égale à la ***somme des probabilités*** de chacun de ces ***chemins***.

Exemple : on donne l'arbre pondéré ci-dessous



calculer les probabilités suivantes :

a) de l'événement $N\cap S$
b) de l'événement $N\cap\overline{S}$
c) de l'événement $\overline{N}\cap S$
d) de l'événement $\overline{N}\cap\overline{S}$

solution : on obtient les résultats suivants :

$P(N\cap S)=0{,}25\times 0{,}05=0{,}0125$; $P(N\cap\overline{S})=0{,}25\times 0{,}95=0{,}2375$

$P(\overline{N}\cap S)=0{,}75\times 0{,}4=0{,}3$; $P(\overline{N}\cap\overline{S})=0{,}75\times 0{,}6=0{,}45$

4) Événements incompatibles

exemple : On considère l'expérience aléatoire suivante :
On tire une carte dans un jeu de 32 cartes à jouer.

On considère les événements suivants :
V : « On tire un valet » et R : « On tire un roi »
Les deux événements A et B sont incompatibles, en effet $V\cap R=\varnothing$
On en déduit que la probabilité de l'événement « Tirer un valet ou un roi » est égale à : $P(V\cup R)=\frac{4}{32}+\frac{4}{32}-0=\frac{8}{32}=0{,}25$

exercice : avec un tableau croisé : **Vidéo :** https://youtu.be/aVXgUHx6ICA