

الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي و البحث العلمي
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

جامعة ابن خلدون بتيـارت
UNIVERSITE IBN KHALDOUN DE TIARET

معهد علوم البيطرة
INSTITUT DES SCIENCES VETERINAIRES



Polycopie
Elément de statistiques descriptives
et inductives.

Rappel de cours & exercices corrigés.

A l'usage des étudiants
1ère année.

Présenté par : Mr AMIRAT Mokhtar.

LES STATISTIQUES DESCRIPTIVES

A. Statistiques a une variable

1. Vocabulaire de la statistique

la statistique est une **énumération** des données (obtenues, mesurées, collectées) permettant d'analyser et de disposer des méthodes afin d'arriver à un certain but. Donc, toute étude en bio-statistique peut être décomposée en deux phases au moins : le recueil ou la **collecte** des données statistiques, et leur **analyse** ou leur **interprétation** selon les phénomènes liés à la biologie.

Un ensemble d'objets ou de personnes d'une étude statistique est appelé **population**. Un élément de cette population est appelé **individu**.

L'étude statistique porte sur un caractère. Si le caractère est **quantitatif**, les mesures sont alors les valeurs d'une **variable statistique** (ex : un âge, une taille...). Si le caractère est **qualitatif**, on est obligé de le quantifier (ex : sexe...).

La variable est dite **discrète** si elle ne prend que des valeurs isolées (ex : entières). Elle est **continue** si elle peut prendre toutes la valeurs d'un intervalle (ex : **R**).

L'**effectif** d'une population est le nombre d'individus total de cette population. La **fréquence** d'un caractère est le nombre d'individus possédant ce caractère divisé par l'effectif total de la population.

2. Les variables discrètes

a) Représentation

On représente les variables aléatoires discrètes sous forme d'histogramme ou de camembert grâce aux différentes fréquences

b) Caractéristiques

(1) La moyenne Nommée X barre

Soit n valeurs distinctes ou non de la variable. Si cette variable prend p valeurs distinctes (p=n), x_1, \dots, x_p , d'effectifs respectifs n_1, \dots, n_p alors la moyenne est donnée par :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n n_i x_i$$

(2) La variance La variance sert à évaluer la dispersion d'une distribution autour d'une valeur centrale, la moyenne.

(3) L'écart-type est une mesure de dispersion par rapport à la moyenne, Elle présente de plus l'avantage d'avoir une signification probabiliste.

. C'est la racine carrée positive de la variance c'est un paramètre dont l'unité est celle de la variable

Elle est donnée par la formule : $V=S^2 = \frac{1}{n} \sum x_i^2 - \left(\frac{1}{n} \sum x_i\right)^2$

$$S^2 = \frac{1}{n} \sum x_i^2 - \left(\frac{1}{n} \sum x_i\right)^2$$

Dans chaque étude statistique il est très important de considérer la nature des données (observations, caractères, attributs) que l'on va tester. Il est donc primordial de préciser la nature de chaque caractère. Il existe deux types de caractères, ceux-ci peuvent être soit quantitatifs soit qualitatifs.

(4) -Le coefficient de variation: C'est le rapport de l'écart type sur la moyenne arithmétique.

Il est indépendant du choix des limites

$C.V = \text{l'écart type} / \text{la moyenne}$

Remarque :

$C.V \geq 0.33$ la dispersion est importante

$C.V < 0.33$ la dispersion est moins importante

1. Caractère quantitatif : C'est un caractère auquel on peut associer un nombre c'est-à-dire, pour simplifier, que l'on peut "mesurer" (grandeur mesurable).

On distingue deux variables statistiques pour ce caractère quantitatif :

- a. Variable discontinue ou discrète :** un tel caractère ne prend qu'un nombre fini de valeurs (valeur entière (N) dénombrable et sans aucune valeur intermédiaire). Par exemple le nombre d'enfants peut prendre : (0 ; 1 ; 2 ; ... ; 10 ; ...).

b. Variable continue : un tel caractère peut, théoriquement, prendre toutes les valeurs d'un intervalle de l'ensemble des nombres réels (R). par exemple la température peut prendre : (-100°C ; 0°C ; 37,8°C).

2. Caractère qualitatif : Dans ce type de caractère, les modalités ne sont pas quantifiables (pas mesurable) tel que les couleurs des yeux. Ce sont des qualités qui peuvent être réalisées dans deux échelles de mesure : **échelle ordinale** (rangement) et **l'échelle nominale**.

On appelle **effectif** de la modalité (x_i), le nombre (n_i). Et parfois on peut trouver dans la littérature le terme **fréquence absolue** pour l'effectif.

On appelle **Fréquence relative** de la modalité (x_i), le nombre (f_i) tel que $f_i = \frac{n_i}{n}$

Parfois on peut rencontrer le terme de **fréquence relative** pour les **fréquences**.

On appelle **Fréquence relative cumulée** le nombre ($f_{i cum}$) tel que $f_{i cum} = \sum_{n=1}^i f_n$

Source :A.Boukerma , AFEDDAL ,A.BENALLI Faculté de médecine d'Oran laboratoire de bio statistique

Exemple :

Parmi les variables suivantes, donner la qualité du caractère (qualitatif, quantitatif) et sa variable statistique (discrète, continue, ordinale ou nominale) :

- 1- Température ;
- 2- La race ovine ;
- 3- Taux de cholestérol ;
- 4- Nombre d'élèves par salle ;
- 5- Répartition par catégories-socio-professionnelles ;
- 6- Nombre de vaches a traité.

Solution :

- 1- La température est un caractère mesurable (quantifiable) et elle peut prendre une valeur de l'ensemble des nombres réels. Donc le caractère est quantitatif d'une variable statistique continue.
- 2- La race est une qualité qui identifie de quelle race appartient notre ovin ce qui implique le caractère qualitatif, et la variable statistique ne peut être que nominale
- 3- Taux de cholestérol d'abord est une mesure donnée par analyse sanguine, alors c'est un caractère quantitatif,
- 4- Nombre d'élèves par salle est un dénombrement ce qui implique directement qu'il s'agit d'une quantité (caractère quantitatif), et la variable statistique est discrète, le fait que le nombre d'élèves ne peut prendre que des valeurs fixes.
- 5- La répartition par catégorie Socio-Professionnelles est une qualité que l'on peut exprimer au rangement des postes (professions) dans la hiérarchie de la société. Donc il s'agit d'un caractère qualitatif avec une variable statistique ordinale.
- 6- Nombre de vaches a traité e c'est un caractère quantitatif avec une variable statistique discontinue.

Résumé :

	Types du caractère	Forme de variable statistique
Température	<i>Quantitatif</i>	<i>Continue</i>
La race ovine	<i>Qualitatif</i>	<i>Nominale</i>
Taux de cholestérol	<i>Quantitatif</i>	<i>Continue</i>
Nombre d'élèves par salle	<i>Quantitatif</i>	<i>Discontinue</i>
Répartition par catégories socio-professionnelles	<i>Qualitatif</i>	<i>Ordinale</i>
Nombre de vaches a traité	<i>Quantitatif</i>	<i>Discontinue</i>

EXERCICE N°1 :

Soit la quantité de créatine en Mgrs par 100 cm³ d'urine chez 80 vaches.

Quantité <i>xi</i>	Nombre de vache <i>ni</i>
2,5 – 3,5	2
3,5 – 4,5	11
4,5 – 5,5	20
5,5 – 6,5	30
6,5 – 7,5	14
7,5 – 8,5	03

Comment appelle t- on ce type de tableau ?

Donner la nature de ce caractère.

Calculer la moyenne, mode et l'écart type(MI).

Solution :

C'est un tableau statistique.

C'est un caractère quantitatif continu

Soit la quantité de créatine en Mgrs par 100 cm³ d'urine chez 80 vaches.

Quantité Classe	Centre de classe	Nombre de vache <i>ni</i>	<i>xi. ni</i>
2,5 – 3,5	3	2	6
3,5 – 4,5	4	11	44
4,5 – 5,5	5	20	100
5,5 – 6,5	6	30	180
6,5 – 7,5	7	14	98
7,5 – 8,5	8	03	24

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n n_i x_i$$

$$\bar{x} = 5,65$$

Le mode A est égale à 6.

$$S_2^2 = \frac{1}{n} \sum x_i^2 - \left(\frac{1}{n} \sum x_i \right)^2$$

La variance est égale donc à 1.22.

L'écart type est la racine carré de la variance est égale à 1,10.

Exercice N°02 :

L'étude du poids de 50 personnes a donné les résultats suivants (en Kg) présentés sous forme d'une série statistique ordonnée :

**37 43 47 50 52 54 55 55 58 58 61 62 63 63 64 65 66 66 67 68 68 69 69 70 71 72
72 72 73 73 74 74 75 76 76 77 79 79 80 82 82 84 86 87 88 90 92 93 98 98**

Dresser un tableau statistique contenant :

- 1- Les données de cette série regroupées en classes ;
- 2- Les centres de classes ;
- 3- Les fréquences absolues ;
- 4- Les fréquences relatives.

Solution :

Premièrement on doit identifier la nature du caractère étudié. Il s'agit de la mesure du poids de 50 personnes (ordonnées selon la série statistique) qui a donné des valeurs quantifiables ce qui implique directement que le caractère est quantitatif continu.

Dans le cas d'un caractère quantitatif continu, l'établissement du tableau de fréquences implique d'effectuer au préalable une répartition en classes des données. Cela nécessite de définir le nombre de classes attendu et donc l'amplitude associée à chaque classe.

En règle générale, on choisit des classes de même amplitude. Pour que la distribution en fréquence ait un sens, il faut que chaque classe comprenne un nombre suffisant de valeurs (n_i).

La formule permettant d'établir le nombre de classes pour un échantillon de taille n :

$$N = \sqrt{n} = \sqrt{50} = 7,07 \rightarrow 7 \text{ classes.}$$

L'amplitude de chaque classe est obtenue de la manière suivante :

$$\text{Amplitude de classe} = \frac{X_{max} - X_{min}}{N} = \frac{98 - 37}{7} = 8,71 \text{ que l'on arrondit à } 9 \text{ Kg par classe.}$$

Avec X_{max} et X_{min} respectivement la plus grande et la plus petite valeur de X dans la série statistique.

x_i : Classes	x_{icc} : centre de classe	n_i : fréquence absolue	f_{icum}
[37 - 46[41,5	2	0,04
[46 - 55[50,5	4	0,08
[55 - 64[59,5	8	0,16
[64 - 73[68,5	14	0,28
[73 - 82[77,5	11	0,22
[82 - 91[86,5	7	0,14
[91 - 100[95,5	4	0,08
		$n = \sum_{i=1}^n n_i = 50$	$\sum_{i=1}^n f_i = 1 = 100\%$

Exercice N°02 :

Soit la répartition de 150 grenouilles suivant le nombre de parasites qu'elles hébergent a obtenu le tableau suivants :

Nombre parasites par grenouille	0	1	2	3	4	5	6
Nombre de grenouilles correspondantes	11	22	45	40	19	11	02

a/Calculer la moyen

b/ calculer la fréquence relative

Solution :

a) Pour calculer le nombre moyen de trématodes par grenouille, on applique directement la formule de la moyenne arithmétique des effectifs.

x_i : Nb de trématode par grenouille	n_i : Nb de grenouille	$n_i x_i$	f_i
0	11	0	0,07
1	22	22	0,14
2	45	90	0,30
3	40	120	0,26
4	19	76	0,12
5	11	55	0,07
6	02	12	0,01
$n = \sum_{i=1}^n n_i = 150$		$\sum_{i=1}^n n_i x_i = 375$	1

$$\bar{x} = 2,5$$

Fréquence relative de la modalité (x_i), le nombre (f_i) tel que $f_i = \frac{n_i}{n}$

Exercice N°03 :

On a mesuré la longueur des tiges d'un lot fleurs et on a obtenu les résultats suivants exprimé par leur valeurs central.

X_i	110	120	130	140	150	160	170	180
ni cumulé	6	12	22	37	54	64	72	74

a / Donner la nature du caractère.

b / Mode, moyenne, la médiane et écart type(MI).

Solution :

a/ Puisque le caractère est défini par le centre de classe donc c'est un caractère quantitatif continu

Xi	110	120	130	140	150	160	170	180
ni cumulé	6	12	22	37	54	64	72	74
Classe	105-115	115-125	125-135	135-145	145-155	155-165	165-175	175-185
ni	6	6	10	15	17	10	08	02

b/ Le mode A = 150 et la classe modale est de 145 – 155.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n n_i x_i$$

$$\bar{x} = 143,91$$

écart type est de 17,91.

B. Représentation graphiques des series Statistiques

La **représentation graphique d'une série statistique** est une étape très importante dans l'analyse d'un problème **statistique** car elle donne une information sur la forme de la distribution observée. Cette forme est souvent caractéristique, comme l'allure en cloche observée pour les histogrammes de poids des nouveau-nés.

Le graphique est un mode d'expression qui permet "visuellement" de saisir et de mémoriser un certain nombre d'informations. C'est pourquoi, lors de la présentation des résultats statistiques et complémentairement aux tableaux, on utilise souvent une représentation graphique .

a/ caractères quantitatifs discontinus

Pour les **caractères quantitatifs discontinus**, la représentation graphique est le **diagramme en bâtons** où la hauteur des bâtons correspond à l’effectif (n_i) associé à chaque modalité du caractère (x_i).

b/ caractères quantitatifs continus

Pour les **caractères quantitatifs continus**, la représentation graphique est l’**histogramme** où la hauteur du rectangle est proportionnelle à l’effectif (n_i).

Autrement On portera en abscisses les valeurs des classes des caractères (variables) , et l’on portera en ordonnées les effectifs correspondants , on dit que l’on représente la « structure de la population étudiée ».

Principe de construction de l’histogramme.

Pour chaque classe, on élève un rectangle ayant une base proportionnelle à l’intervalle de classe et une hauteur proportionnelle à l’effectif simple. Dans ce cas, ce sont les « surfaces » , et non les hauteurs , qui « **sont proportionnelles à l’effectif** ».

Dans la pratique, cependant, deux cas peuvent se présenter. Le cas où les classes sont d’égales amplitudes et le cas ou les amplitudes sont inégales.

Cas 1 : Classes sont d’égales amplitudes.

Le **caractère qualitatif** peut se représenter soit par des **tuyaux d’orgue** soit par des **secteurs circulaires** (camembert).

Exercice N°01 :

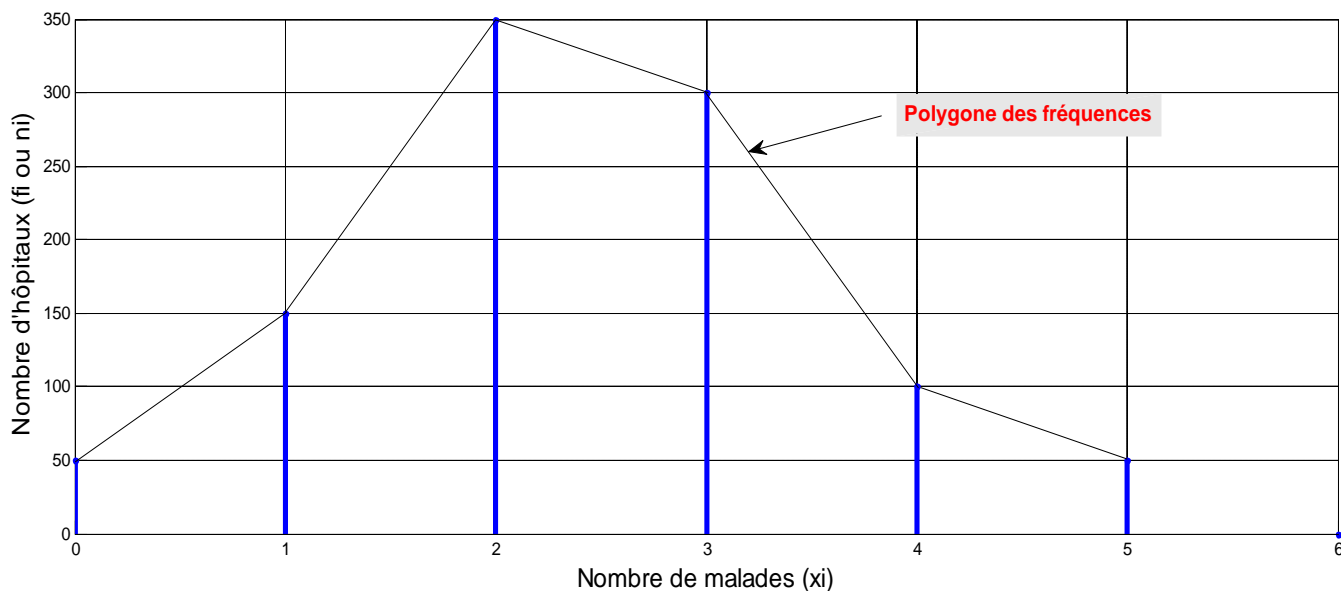
On recense dans 1000 écoles d’une région donnée le nombre d’eleves atteint d’une maladie M. On trouve les résultats :

Nombre d’élèves malades	0	1	2	3	4	5
Nombre d’écoles	50	150	350	300	100	50

Représenter graphiquement ces données et donner les conclusions nécessaires.

Solution :

La variable statistique est discontinue → **diagramme en bâtons**



Intitulé : Répartition des malades par école

Exercice N°02 :

Soit le tableau suivant dont la variable statistique est définie par le centre vde classe.

Centre de classe	45	75	105	135	165	195	225	255	285
ni cumulé	01	01	08	18	36	44	47	49	50

Questions :

Nature du caractère.

Déterminer l'amplitude a et les classes.

Calculer Fr

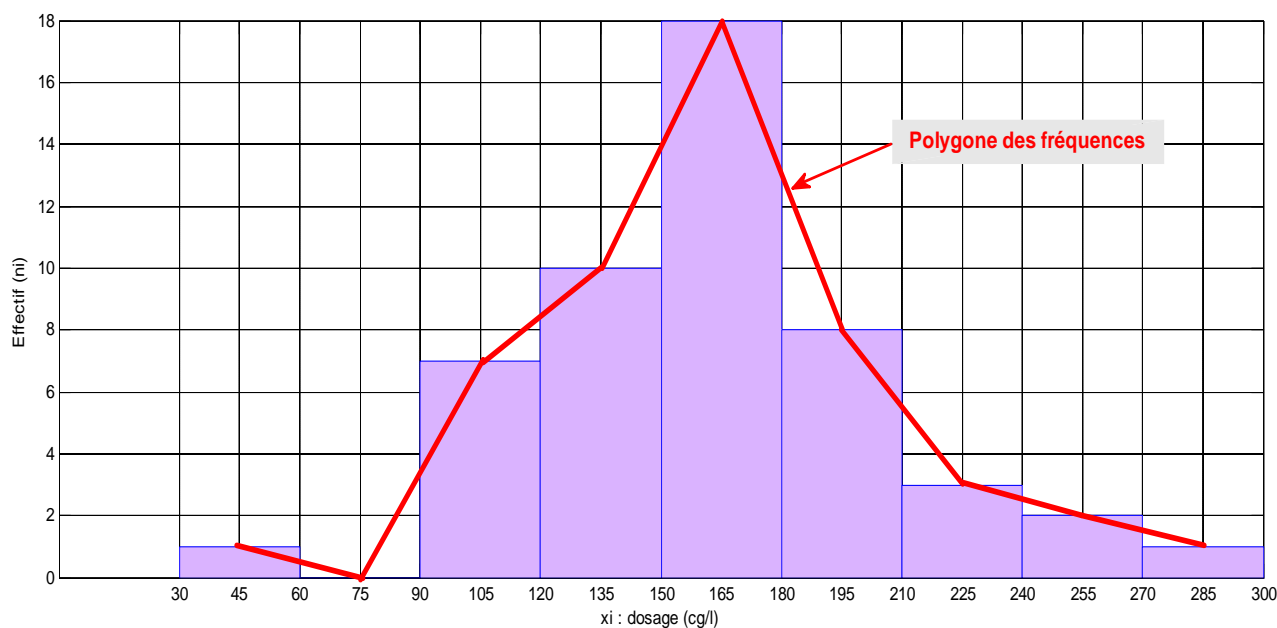
Représentations graphique

Solution :

On remarque à travers le tableau que l'amplitude $k = 30$ cg/l. on en déduit le tableau en classes :

X_i	Classes	n_i cumulée	n_i : effectif	Fr
45	[30 - 60[01	01	0.02
75	[60 - 90[01	00	0.0
105	[90 - 120[08	07	0.14
135	[120 - 150[18	10	0.2
165	[150 - 180[36	18	0.36
195	[180 - 210[44	08	0.16
225	[210 - 240[47	03	0.06
255	[240 - 270[49	02	0.04
285	[270 - 300[50	01	0.02

La variable statistique est continue → **Représentation par histogrammes**



Exercice N°04 :

Le tableau suivant montre les taux de naissance et de mortalité de 1000 personnes.

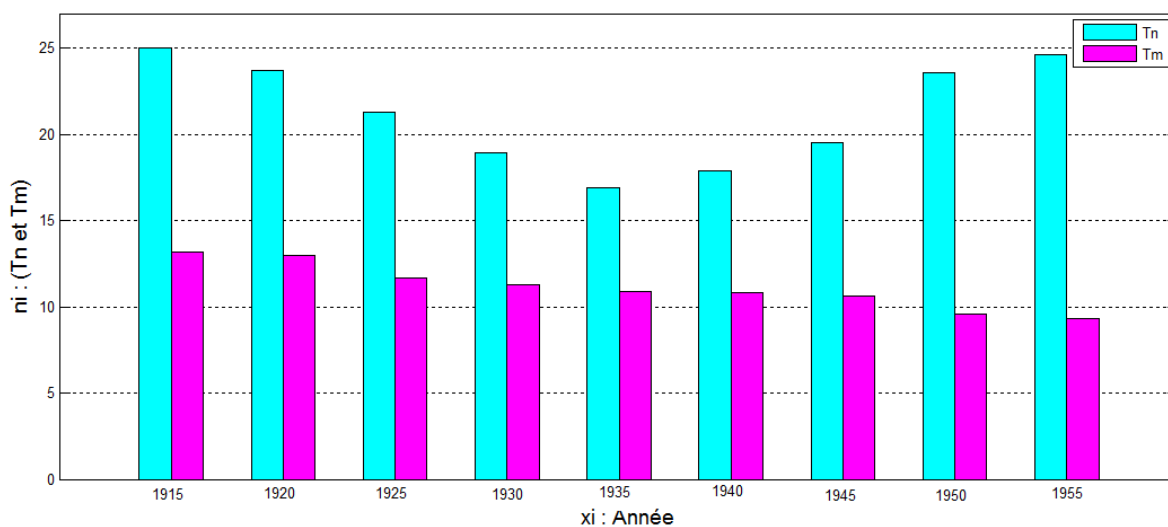
Année	1915	1920	1925	1930	1935	1940	1945	1950	1955
Tn (%)	25,0	23,7	21,3	18,9	16,9	17,9	19,5	23,6	24,6
Tm (%)	13,2	13,0	11,7	11,3	10,9	10,8	10,6	09,6	09,3

Représenter graphiquement ces résultats sur le même graphique. Faire le commentaire.

Source :A.Boukerma , AFEDDAL ,A.BENALLI Faculté de médecine d’Oran laboratoire de bio statistique

Solution :

Le caractère étudié est qualitatif → **Représentation par tuyaux d’orgue**



Intitulé : Répartition des taux de naissance et de mortalité par année.

La médiane Me

En théorie des probabilités et en statistiques, la médiane est la valeur qui sépare la moitié inférieure de la moitié supérieure d'un ensemble. Intuitivement, la médiane est ainsi le point milieu de l'ensemble. C'est un indicateur de tendance centrale de la série.

La médiane est le point milieu d'un jeu de données, de sorte que 50 % des unités ont une valeur inférieure ou égale à **la médiane** et 50 % des unités ont une valeur supérieure ou égale. Dans un jeu de données de petite taille, il suffit de compter le nombre de valeurs (n) et de les ordonner en ordre croissant

Pour **calculer la médiane** : On classe les valeurs de la série statistique dans l'ordre croissant : Si le nombre de valeurs est impair, la **médiane** est la valeur du milieu. S'il est pair, la **médiane** est la demi-somme des deux valeurs du milieu.

Autrement la médiane est la valeur de la variable statistique qui divise la série statistique en deux parties égales

(50% de l'effectif lui soit inférieur et 50% supérieur)

- Nombre d'observation impair :

7, 9, 13, 45, 70, 101, 115

$$\text{Me} = 45$$

- Nombre d'observations pair :

2, 5, 9, 10, 12, 14, 20, 22

$$\text{Me} = (10 + 12) / 2 = 11$$

La demi-somm

e de l'intervalle médian

Variable statistique discontinue

la médiane est la valeur de la variable statistique qui occupe le $(n / 2)$ ème rang (détermination direct)

Exemple : Soit le nombre d'enfant par famille comme l'indique le tableau suivant.

Nombre d'enfants	Nombre de familles ni
0	4
1	5
2	10
3	16
4	18
5	14
6	07
7	06
Total	80

$n / 2 = 80 / 2 = 40$, la valeur de la variable qui occupe le 40ème rang est égale à 4 donc $Me = 4$ enfants

Interprétation :

Il ya 50% (soit 40 familles) qui ont moins de 4 enfants et 40 plus de 4 enfants

Variable statistique continue

Il y a deux méthodes : direct, méthode d'interpolation linéaire

a) Méthode direct

Les pesées de 50 nouveau-nées

Poids (Kg) X_i	Poids (Kg) X_i
2 - 2.5	2
2.5 - 3	4
3 - 3.5	6
3.5 - 4 Classe médiane	30
4 - 4.5	8
Total	50

$3,0 - 3,5$ alors $Me = (3,0 + 3,5) / 2$

Donc $Me = 3,75$ Kg

Interprétation :

Il ya 50% (soit 25) nouveau-née qui ont un poids inferieur à 3,75 Kg et 50% (25) qui on un poids supérieur à 3,75 Kg

Source : Bio statistique Département pharmacie Oran Année 2008/2009 Faculté de médecine Oran

Exercice:

L'étude du périmètre encéphalique de deux échantillons de 36 enfants a fourni les résultats ci-contre :

PERIMETRE (cm)	n_i : (1 ^{er} échantillon)	n_i : (2 ^{ème} échantillon)
40 – 42	01	02
42 – 44	04	03
44 – 46	10	18
46 – 48	12	08
48 – 50	06	04
50 – 52	03	01

- a) Calculer les paramètres de tendance centrale.
- b) Comparer ces deux distributions en calculant la variance et l'écart-type de chacune.
- c) Pouvez-vous en tirer une conclusion ?

Source :A.Boukerma , AFEDDAL ,A.BENALLI Faculté de médecine d'Oran laboratoire de bio statistique

Solution :

1^{er} échantillon

x_i : classes	n_i : effectif	$x_i cc$	$n_i cum$	$n_i x_i$	$n_i x_i^2$
40 – 42	01	41	1	41	1681
42 – 44	04	43	5	172	7396
44 – 46	10	45	15	450	20250
46 – 48	12	47	27	564	26508
48 – 50	06	49	33	294	14406
50 – 52	03	51	36	153	7803
	36			1674	78044

a) Détermination des paramètres de tendances centrales (indicateurs de position):

- Calcul de la moyenne arithmétique :

$$\bar{x}_1 = \frac{1}{n} \sum_{i=1}^n n_i x_i cc = \frac{1674}{36} = 46,5 \Rightarrow \bar{x}_1 = 46,5 \text{ cm}$$

- Calcul de mode par interpolation linéaire:

$$M_{O1} = b_{min} + \frac{d_1}{d_1 + d_2} \times k = 46 + \frac{12 - 10}{(12 - 10) + (12 - 6)} \times 2 = 46,5 \Rightarrow M_{O1} = 46,5 \text{ cm}$$

- Calcul de la médiane par interpolation :

Il faut tout d’abord déterminer la classe médiane en procédant par la détermination de son rang.

$Rang = \frac{n}{2} = \frac{36}{2} = 18^{ème}$ rang, la 18^{ème} valeur se situe dans la classe [46 – 48[qui contient le nombre des enfants de 16 à 27 d’où avec $b_{min} = 46 \text{ cm}$, $n_{me} = 12$ enfants et $S = 15$.

$$M_{e1} = b_{min} + \frac{(\frac{n}{2} - S)}{n_{med}} K = 46 + \frac{18 - 15}{12} \times 2 = 46,5 \Rightarrow M_{e1} = 46,5 \text{ cm}$$

b) Détermination des paramètres de dispersion :

- Calcul de la variance :

$$S_1^2 = \frac{1}{n} \sum n_i x_i^2 - \bar{x}^2 = \frac{78044}{36} - (46,5)^2 = 5,638 \Rightarrow S_1^2 = 5,638 \text{ cm}^2$$

- Calcul de l’écart-type :

$$S_1 = \sqrt{S_1^2} = \sqrt{5,638} = 2,37 \Leftrightarrow S_1 = 2,37 \text{ cm}$$

2^{ème} échantillon

x_i : classes	n_i : effectif	$x_i cc$	$n_i cum$	$n_i x_i$	$n_i x_i^2$
40 – 42	02	41	2	82	3362
42 – 44	03	43	5	129	5547
44 – 46	18	45	23	810	36450
46 – 48	08	47	31	376	17672
48 – 50	04	49	35	196	9604
50 – 52	01	51	36	51	2601
	36			1644	75236

c) Détermination des paramètres de tendances centrales (indicateurs de position):

- Calcul de la moyenne arithmétique :

$$\bar{x}_2 = \frac{1}{n} \sum_{i=1}^n n_i x_i cc = \frac{1644}{36} = 46,5 \Rightarrow \bar{x}_2 = 45,66 \text{ cm}$$

- Calcul de mode par interpolation linéaire:

$$M_{O2} = b_{min} + \frac{d_1}{d_1 + d_2} \times k = 44 + \frac{18 - 3}{(18 - 3) + (18 - 8)} \times 2 = 45,2 \Rightarrow M_{O2} = 45,2 \text{ cm}$$

- Calcul de la médiane par interpolation :

Il faut tout d’abord déterminer la classe médiane en procédant par la détermination de son rang.

$Rang = \frac{n}{2} = \frac{36}{2} = 18^{\text{ème}}$ rang, la $18^{\text{ème}}$ valeur se situe dans la classe $[44 - 46[$ qui contient le nombre des enfants de 06 à 23 d’où avec $b_{min} = 44 \text{ cm}$, $n_{me} = 18$ enfants et $S = 5$.

$$M_{e2} = b_{min} + \frac{(\frac{n}{2} - S)}{n_{med}} K = 44 + \frac{18 - 5}{18} \times 2 = 45,44 \Rightarrow M_{e2} = 45,44 \text{ cm}$$

d) Détermination des paramètres de dispersion :

- Calcul de la variance :

$$S_2^2 = \frac{1}{n} \sum n_i x_i^2 - \bar{x}^2 = \frac{75236}{36} - (45,66)^2 = 5,053 \Rightarrow S_1^2 = 5,053 \text{ cm}^2$$

- Calcul de l’écart-type :

$$S_2 = \sqrt{S_2^2} = \sqrt{5,053} = 2,24 \Rightarrow S_2 = 2,24 \text{ cm}$$

Tableau récapitulatif :

	1 ^{er} échantillon	2 ^{ème} échantillon
Paramètres de tendance centrale	$\bar{x}_1 = 46,5$	$\bar{x}_2 = 45,66$
	$M_{O1} = 46,5$	$M_{O2} = 45,2$
	$M_{e1} = 46,5$	$M_{e2} = 45,44$
Paramètres de dispersion	$S_1^2 = 5,638$	$S_2^2 = 5,053$
	$S_1 = 2,37$	$S_2 = 2,24$

Analyse et interprétation

- $\bar{x}_1 > \bar{x}_2$: La moyenne du 1^{er} échantillon est supérieure à celle du 2^{ème} échantillon.
- Le périmètre le plus courant, répandu ou fréquent du 1^{er} échantillon est $46,5 \text{ cm}$, ainsi, le périmètre le plus courant, répandu ou fréquent du 2^{ème} échantillon est $46,5 \text{ cm}$.

- Pour le 1er échantillon, il ya 18 enfants qui ont un périmètre inférieur à 46,5 cm et 18 enfants ont un périmètre supérieur à 46,5 cm. Par contre l'échantillon N°02, il ya 18 enfant qui ont un périmètre inférieur à 45,44 cm et 18 enfants ont un périmètre supérieur à 45,44 cm.
- $S_1 > S_2$: la dispersion des mesures autour de la moyenne pour le premier échantillon est plus importante que la dispersion pour le 2^{ème} échantillon.

Sujets d'examens

1/ Soit la répartition des poids et des tailles de 300 garçons de 10 ans définie par leurs centres de classe

POIDS Kg	TAILLE cm						
	120	125	130	135	140	145	150
20	4						
22.5	3	7	2				
25	2	17	20	7	1		
27.5		8	30	28	6		
30		1	17	36	19	4	
32.5			4	16	22	3	2
35			1	5	10	8	
37.5				1	4	5	2
40					2	1	2

a/ Donnez la nature du caractère b/ la moyenne et l'écart type (MI) c/ Fr d/ L'étendue de la série et la classe modale e/ Histogrammes.

2/ On a dénombré chez un individu 1000 leucocytes, et on s'est intéressé à leur catégorie .

Leucocytes	neutrophile	éosinophile	Basophile	Lymphocyte	Monocytes
Ni	600	20	10	110	260

a / La nature du caractère .

b / ni cumulé . c / le mode . d /Fr.

3/ Un coureur s'entraîne dans le but de participer à une course de 1 Km , pour chacune de ses 60 dernières courses d'entraînement, il a noté le temps en seconde pris définie par le centre de classe .

Xi	262,5	267,5	272,5	277,5	282,5	287,5	292,5
Ni	5	8	8	15	10	10	4

1. La nature du caractère.
2. La moyenne et l'écart type par la méthode indirecte.
3. La médiane.
4. la représentation graphique .

Analyse combinatoire

1. Définition :

L'**analyse combinatoire** est une branche des mathématiques qui étudie comment **compter** les objets. Elle fournit des méthodes de dénombrements particulièrement utiles en théorie des probabilités. Les **probabilités dites combinatoires** utilisent constamment les formules de l'analyse combinatoire développées dans ce chapitre. Un exemple des applications intéressantes de cette dernière est la démonstration du développement du **binôme de Newton** utilisé dans le calcul des probabilités d'une **loi binomiale**.

2. Arrangements

2.1. Définition

Etant donné un ensemble E de n objets, on appelle **arrangements** de p objets **toutes suites ordonnées** de p objets pris parmi les n objets.

Le nombre d'arrangements de p objets pris parmi n est noté : A_n^p

Remarque : On a nécessairement $1 \leq p \leq n$ et $n, p \in \mathbb{N}^*$

Si $n < p$, alors $A_n^p = 0$

Deux arrangements de p objets sont donc **distincts** s'ils diffèrent par la nature des objets qui les composent ou par leur ordre dans la suite.

Exemples :

(1) Une séquence d'**ADN** est constituée d'un enchaînement de 4 nucléotides [A (Adénine), C (Cytosine), G (Guanine) et T (Thymine)]. Il existe différents arrangements possibles de deux nucléotides ou di nucléotides avec $p=2$ et $n=4$.

(2) Le nombre de mots de 5 lettres (avec ou sans signification) formés avec les 26 lettres de l'alphabet correspond au nombre d'arrangements possibles avec $p=5$ et $n=26$.

(3) Le tiercé dans l'ordre lors d'une course de 20 chevaux constitue un des arrangements possibles avec $p=3$ et $n=20$.

Dans les exemples précédents, l'ordre des éléments dans la suite est essentiel. Ainsi pour le deuxième exemple, le mot NICHE est différent du mot CHIEN.

Mais dans les deux premiers exemples, une base ou une lettre de l'alphabet **peut se retrouver plusieurs fois** alors que dans le troisième exemple, les trois chevaux à l'arrivée sont forcément **différents**. Il faut donc distinguer le nombre d'**arrangements avec répétition** et le nombre d'**arrangements sans répétition** (arrangements au sens strict).

2.2. Arrangements avec répétitions

Lorsqu'un objet peut être observé **plusieurs fois** dans un arrangement, le nombre d'**arrangement avec répétition de p objets pris parmi n** , est alors :

$$A_n^p = n^p \text{ avec } 1 \leq p \leq n$$

Voici pourquoi :

Pour le premier objet tiré, il existe n manières de ranger l'objet parmi n .

Pour le second objet tiré, il existe également n possibilités d'arrangements car le premier objet fait de nouveau parti des n objets. On parle de **tirage avec remise**.

Ainsi pour les p objets tirés, il y aura $n \times n \times n \times \dots \times n$ (p fois) arrangements possibles, soit $A_n^p = n \times n \times n \times \dots \times n = n^p$

Exemples :

(1) Concernant l'exemple de **la séquence d'ADN**, le nombre de dinucléotides attendus si l'on fait l'hypothèse qu'une base peut être observée plusieurs fois dans la séquence (ce qui correspond effectivement à la réalité) est donc : $A_4^2 = 16$

dinucléotides possibles

Les 16 dinucléotides identifiables dans une séquence d'ADN sont :

AA AC AG AT CA CC CG CT

GA GC GG GT TA TC TG TT

2.3. Arrangements sans répétition

Lorsque chaque objet ne peut être observé qu'une seule fois dans un arrangement, le nombre d'**arrangements sans répétition** de p objets pris parmi n est alors :

$$A_n^p = \frac{n!}{(n-p)!} \quad \text{avec } 1 \leq p \leq n$$

3. Permutations :

La permutation de n objets constitue un cas particulier d'**arrangement sans répétition** de p objets pris parmi n lorsque $p = n$.

a) *Permutations sans répétitions :*

Etant donné un ensemble E de (n) objets, on appelle **permutations** de (n) objets distincts **toutes suites ordonnées** de (n) objets ou tout **arrangement n à n** de ces objets. Le nombre de permutations de (n) objets est noté :

$$P_n = A_n^n = \frac{n!}{(n-n)!} = n! = n(n-1)(n-2) \dots \dots 1$$

b) *Permutations avec répétitions :*

Dans le cas où il existerait plusieurs répétitions (k) d'un même objet parmi les (n) objets, le nombre de permutations possibles des (n) objets doit être rapporté aux nombres de permutations des (k) objets identiques.

Le nombre de permutations de n objets est alors : $P_n = \frac{n!}{k!}$

Exemple :

Considérons le mot « CELLULE ». Le nombre de mots possibles (avec ou sans signification) que l'on peut écrire en permutant ces 7 lettres est :

$$P_7 = \frac{7!}{2!3!}$$

= **420 mots possibles** en considérant deux groupes de lettres identiques : L (3 fois) et E (2 fois).

4. Combinaisons

1. Définition

Si l'on reprend l'**exemple de la séquence d'ADN**, à la différence des arrangements où les dinucléotides AC et CA formaient deux arrangements distincts, ces derniers ne formeront qu'une seule combinaison. Pour les combinaisons, on ne parle plus de suite ni de série puisque la **notion d'ordre des objets n'est plus prise en compte**. On parle alors de tirages avec ou sans remise.

On parle alors de tirages avec ou sans remise.

a) Combinaisons sans remise :

Etant donné un ensemble **E** de (**n**) objets, on appelle **combinaisons** de (**p**) objets tout ensemble de (**p**) objets pris parmi les (**n**) objets sans remise.

Le nombre de combinaisons de (**p**) objets pris parmi (**n**) est noté :

$$C_n^p = \frac{n!}{p!(n-p)!} \quad \text{avec } (1 \leq p \leq n)$$

Voici pourquoi :

Pour calculer le nombre de combinaison, on utilise le principe de la division.

- Il ya A_n^p manière de tirer **p** objet parmi **n** en les ordonnant soit $A_n^p = \frac{n!}{(n-p)!}$
- Une fois les **p** objets tirés, il y a **p!** manières de les ordonner.
- Il ya donc $\frac{A_n^p}{p!}$ manières de tirer **p** objets parmi **n** sans les ordonner.

$$C_n^p = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}$$

b) Combinaisons avec remise :

Le nombre de combinaisons de (**p**) objet parmi (**n**) avec remise est :

$$C_{n+p-1}^p = \frac{(n+p-1)!}{p!(n-1)!}$$

Exemples :

(1) Le tirage au hasard de 5 cartes dans un jeu de 32 (main de poker) est une combinaison avec $p=5$ et $n=32$.

(2) La formation d'une délégation de 5 personnes parmi un groupe de 50 constitue une combinaison avec $p=5$ et $n=50$.

Pour ces deux exemples, les objets tirés sont clairement distincts.

Exercices

Exercice N°01 :

Un médecin doit au cours d'une étude, visiter 6 villes :

- a) S'il se trouve 10 villes dans l'aire géographique à laquelle il s'intéresse, de combien de manière lui est-il possible de choisir un groupe de 6 villes ?
- b) Toujours dans la même hypothèse de 10 villes situées dans cette aire géographique, mais en admettant, en outre, que l'ordre dans lequel il visitera des 6 villes choisies présente de l'importance, combien d'itinéraires différents peut-on sélectionner ?
- c) En supposant que les 6 villes à visiter aient été choisies mais pas encore l'itinéraire qui leur est dévolu, de combien de manières peut se dérouler la visite des 6 villes ?

Solution :

- a) le médecin doit visiter 6 villes parmi 10 villes, le nombre de possibilités pour choisir un groupe de 6 villes sans tenir compte de l'ordre (combinaisons) est :

$$C_{10}^6 = \frac{10!}{6!(10-6)!} = \frac{3628800}{720 \times 24} = 210$$

Donc il existe 210 manières pour choisir un groupe de 6 villes parmi 10 villes.

- b) Toujours le médecin doit visiter 6 villes parmi 10 villes, le nombre de possibilités pour choisir un groupe de 6 villes en tenant compte de l'ordre (arrangements) est :

$$A_{10}^6 = \frac{10!}{(10-6)!} = \frac{3628800}{24} = 151200$$

Donc il existe 151200 itinéraires différents entre les 6 villes visitées parmi 10 villes.

- c) Maintenant, le choix des 6 villes a été réalisé. Le médecin doit visiter ces 6 villes. le nombre de possibilités pour faire sa visite entre les 6 villes sans tenir compte de l'ordre (permutations) est :

$$P_6 = 6! = 720$$

Donc il existe 720 manières de définir un itinéraire entre les 6 villes.

Exercice N°02 :

Soit une population composée de 12 éléments. Combien d'échantillons de 5 éléments peut-on extraire de cette population, si l'échantillonnage s'effectue sans remise et si l'on ne tient pas compte de l'ordre de sortie ?

Solution :

$$C_{12}^5 = \frac{12!}{5!(12-5)!} = \frac{12 \times 11 \times 10 \times 9 \times 8}{5 \times 4 \times 3 \times 2} = 792$$

On peut extraire 792 échantillons de 5 éléments d'une population composée de 12 éléments.

Exercice n° 3 :

A partir d'un groupe de 15 athlètes on veut former une équipe de 7 joueurs. Combien d'équipes peuvent on former ?

Exercice n° 4:

Une classe comprend 25 étudiants dont 15 garçons et 10 filles. Pour des travaux pratiques on choisit des sous-groupes de 5 étudiants.

a / Combien de sous groupes différents peut on former ?

b / Combien de sous groupes différents contenant 3 garçons et 2 filles seulement peut on former ?

Exercice N°05 :

On considère un lot de 35 boules identiques, numérotées de 1 à 35. Quelle est la probabilité de tirer :

- Une boule portant un numéro pair ?
- Une boule portant un numéro impair ?
- Une boule portant un numéro strictement supérieur à 5 ?

Solution :

L'épreuve (expérience) : « tirer une boule ».

a) L'évènement (a) : « avoir une boule portant un numéro pair ».

$$P(a) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_{17}^1}{C_{35}^1} = \frac{17}{35}$$

b) L'évènement (b) : « avoir une boule portant un numéro impair ».

$$P(b) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_{18}^1}{C_{35}^1} = \frac{18}{35}$$

c) L'évènement (c) : « avoir une boule portant un numéro strictement supérieur à 5 ».

$$P(c) = \frac{\text{nombre de cas favorable}}{\text{nombre de cas possible}} = \frac{\Omega}{N} = \frac{C_{30}^1}{C_{35}^1} = \frac{30}{35}$$

$$\text{Nb : } P(a) + P(b) = 1$$

Exercice N°06 :

Un vétérinaire doit au cours d'une étude, visiter 6 fermes :

- S'il trouve 10 fermes dans l'aire géographique à laquelle il s'intéresse, de combien de manière lui est-il possible de choisir un groupe de 6 fermes ?
- Toujours dans la même hypothèse de 10 fermes situées dans cette aire géographique, mais en admettant, que l'ordre dans lequel il visitera les 6 fermes choisies présente de l'importance, combien d'itinéraires différents peut-on sélectionner ?
- Déterminez le nombre de possibilités pour faire sa visite entre les 6 fermes sans tenir compte de l'ordre.

Exercice N°07 :

Trois garçons et trois filles s'assoient sur un banc.

Calculer les probabilités pour que :

- Les trois filles s'assoient l'une à côté de l'autre.
- L'on ait alternativement un garçon et une fille.

Exercice n°8

L'étude des groupes sanguins montre qu'il existe deux agglutinogènes A et B. ces derniers peuvent exister séparément (groupe A et groupe B) ou simultanément (groupe AB) ou peuvent manquer tous les deux (groupe 0). Pour les transfusions sanguines, on sait que :

- Les sujets du groupe A ne peuvent recevoir que du sang des groupes A et 0
- Les sujets du groupe B ne peuvent recevoir que du sang des groupes B et 0
- Les sujets du groupe AB peuvent recevoir le sang de n'importe quel groupe
- Les sujets du groupe 0 ne peuvent recevoir que du sang des groupes 0

Dans une population, un sujet pris au hasard, a la probabilité : 0,43 d'être du groupe A, 0,12 d'être du groupe B et 0,40 d'être du groupe 0

Exercice N°09 :

Deux postes de médecins sont offerts dans un hôpital, postes qui peuvent indifféremment être occupés par un homme ou par une femme. 4 femmes et 2 hommes posent leur candidature et les chances de chacun des 6 postulants sont à priori équivalentes.

- a) Quelle est la probabilité pour que les 2 postes soient occupés par 2 femmes ?
- b) Quelle est la probabilité pour que les 2 postes soient occupés par 2 hommes ?
- c) Quelle est la probabilité pour que les 2 postes soient occupés par 2 personnes de même sexe ?
- d) Quelle est la probabilité pour que les 2 postes soient occupés par 2 personnes de sexe différent ?

Solution :

Premièrement, il faut connaître l'épreuve (expérience) : « **extraire 2 postes** ».

a) L'évènement (A) : « **2 postes soient occupés par 2 femmes** ».

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_4^2 \times C_2^0}{C_6^2} = \frac{2}{5}$$

La probabilité pour que les 2 postes soient occupés par 2 femmes est : $P(A) = \frac{2}{5}$

b) L'évènement (B) : « **2 postes soient occupés par 2 hommes** ».

$$P(B) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_2^2 \times C_4^0}{C_6^2} = \frac{1}{15}$$

La probabilité pour que les 2 postes soient occupés par 2 hommes est : $P(B) = \frac{1}{15}$

c) L'évènement (C) : « **2 postes soient occupés par 2 personnes de même sexe** ».

$$P(C) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_2^2}{C_6^2} + \frac{C_4^2}{C_6^2} = \frac{7}{15}$$

La probabilité pour que les 2 postes soient occupés par 2 hommes ou 2 femmes est :

$$P(C) = \frac{7}{15}$$

d) L'évènement (D) : « **2 postes soient occupés par 2 personnes de sexe différent** ».

$$P(D) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_2^1 \times C_4^1}{C_6^2} = \frac{8}{15}$$

La probabilité pour que les 2 postes soient occupés par 1 homme et 1 femme est :

$$P(D) = \frac{8}{15}$$

On peut déterminer la probabilité de l'évènement (D) d'une autre manière, en utilisant la probabilité de l'évènement (C). Puisque $P(D) = P(\bar{C})$.

$$\text{Donc, } P(C) + P(\bar{C}) = 1 \Leftrightarrow P(C) + P(D) = 1 \Leftrightarrow P(D) = 1 - P(C) = 1 - \frac{7}{15}$$

$$\Rightarrow P(D) = \frac{8}{15}$$

Exercice N°10 :

Soit une population composée de 12 éléments. Combien d'échantillons de 5 éléments peut-on extraire de cette population, si l'échantillonnage s'effectue sans remise et si l'on ne tient pas compte de l'ordre de sortie ?

Solution :

$$C_{12}^5 = \frac{12!}{5!(12-5)!} = \frac{12 \times 11 \times 10 \times 9 \times 8}{5 \times 4 \times 3 \times 2} = 792$$

On peut extraire 792 échantillons de 5 éléments d'une population composée de 12 éléments.

Exercice N°11:

Dans une classe, on souhaite élire un comité. (Un comité est un petit groupe d'élèves auquel on confiera une mission particulière). On suppose que chaque élève de la classe peut-être élu.

1. Combien de comité de 3 personnes peut-on élire dans une classe de 31 élèves ?
2. Dans une classe de (n) élèves ; il ya 351 façons d'élire un comité de 2 personnes. Quel est le nombre n élève de cette classe ?

Solution :

1. Détermination du nombre de comités de 3 personnes dans une classe de 31 élèves :

$$C_{31}^3 = \frac{31!}{3!(31-3)!} = \frac{31 \times 30 \times 29}{3 \times 2} = 4495$$

On peut élire 4495 comités composés de 3 personnes dans une classe de 31 élèves.

2. Détermination du nombre n élève :

$$C_n^2 = 351 \Leftrightarrow \frac{n!}{2!(n-2)!} = 351$$

$$\Leftrightarrow \frac{n(n-1)(n-2)!}{2!(n-2)!} = 351$$

$$\Leftrightarrow n(n - 1) = 702 \quad \Leftrightarrow \quad n^2 - n - 702 = 0$$

$$\sqrt{\Delta} = 53 \Leftrightarrow \begin{cases} n_1 = -26 & \text{impossible d'avoir une classe avec un nombre négatif} \\ n_2 = 26 & \geq 0 \quad \text{donc } n = n_2 \end{cases}$$

Donc la classe est composée de 26 élèves, dont on peut élire 351 comités de 2 personnes.

Exercice N°12 :

Pour les besoins d'un test sur un vaccin V, nous disposons de 10 volontaires, 3 d'entre eux appartiennent à la même famille. Deux personnes sont tirées au hasard. Quelle est la probabilité P(F) que ces deux (02) personnes soient de la même famille ?

Solution :

Premièrement, il faut connaître l'épreuve (expérience) : « **tirer 2 personnes** ».

L'évènement (F) : « **2 personnes tirées soient de la même famille** ».

$$P(F) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_3^2}{C_{10}^2} = \frac{3}{45}$$

La probabilité de tirer 2 personnes de la même famille est : $P(F) = \frac{3}{45}$

CALCUL DE PROBABILITE

Les **probabilités** correspondent à la branche des mathématiques qui cherche à mesurer le caractère aléatoire de ce qui pourrait survenir. Calculer une **probabilité** revient donc à quantifier la possibilité qu'un évènement se produise lors d'une expérience qui ne découle que du hasard.

Le terme probabilité possède plusieurs sens : venu historiquement du latin *probabilitas*, il désigne l'opposé du concept de certitude ; il est également une évaluation du caractère probable d'un évènement.

La *probabilité* d'un évènement est un nombre réel compris entre 0 et 1.

La probabilité d'un évènement est donc

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Evènement : En théorie des probabilités, on appelle évènement élémentaire un ensemble de l'univers constitué d'un seul élément. Par exemple dans un jeu de carte classique de 52 cartes, tirer le roi de cœur est un évènement élémentaire car le paquet de carte ne contient qu'un seul roi de cœur.

Epreuve : une **épreuve** est le "tirage" d'un évènement élémentaire x de X , qui est le résultat de l'**épreuve**. Par exemple, lancer un dé, tirer une boule, mettre en service un appareil susceptible de tomber en panne sont autant d'**épreuves** aléatoires.

Exercice N°01

Une urne contient 8 jetons blanc, 12 noir, 18 jaunes et 23 rouge. On tire un jeton. Quelle est la probabilité qu'il soit : blanc ? Noir ? Jaune ? Rouge ? Blanc ou noir ? Jaune ou rouge ? Blanc, noir ou jaune ?

Exercice N°02 :

On extrait au hasard une boule d'une boîte contenant 6 boules rouges, 4 boules blanches et 5 boules bleues.

Déterminer la probabilité d'avoir une boule :

- a) noir b) jaune c) orange d) Pas rouge e) Rouge ou blanche.

Solution :

6 boules noir + 4 boules blanches + 5 boules bleues = 15 boules

L'épreuve (expérience) : « **extraire une boule** ».

a) L'évènement (a) : « **avoir une boule noir** ».

$$P(a) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_6^1}{C_{15}^1} = \frac{2}{5}$$

La probabilité d'avoir 1 boule noir est : $P(a) = \frac{2}{5}$

b) L'évènement (b) : « **avoir une boule jaune** ».

$$P(b) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_4^1}{C_{15}^1} = \frac{4}{15}$$

La probabilité d'avoir 1 boule jaune est : $P(b) = \frac{4}{15}$

c)

d) L'évènement (c) : « **avoir une boule orange** ».

$$P(c) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_5^1}{C_{15}^1} = \frac{1}{3}$$

La probabilité d'avoir 1 boule orange est : $P(c) = \frac{1}{3}$

e) L'évènement (d) : « **avoir une boule qui n'est pas noir** ».

$$P(d) = P(b \cup c) = P(b) + P(c) = \frac{4}{15} + \frac{1}{3} = \frac{3}{5}$$

La probabilité d'avoir une boule soit jaune soit orange (pas noir) est : $P(d) = \frac{3}{5}$

On peut déterminer la probabilité de l'évènement (d) d'une autre manière, en utilisant la probabilité de l'évènement (a). Puisque $P(d) = P(\bar{a})$.

$$\text{Donc, } P(\bar{a}) + P(a) = 1 \Leftrightarrow P(d) + P(a) = 1 \Leftrightarrow P(d) = 1 - P(a) = 1 - \frac{2}{5}$$

$$\Leftrightarrow P(d) = \frac{3}{5}$$

f) L'évènement (e) : « **avoir une boule noir ou jaune** ».

$$P(e) = P(a \cup b) = P(a) + P(b) = \frac{2}{5} + \frac{4}{15} = \frac{2}{3}$$

La probabilité d'avoir une boule soit noir soit jaune est : $P(e) = \frac{2}{3}$

On peut déterminer la probabilité de l'évènement (e) d'une autre manière, en utilisant la probabilité de l'évènement (c). Puisque $P(e) = P(\bar{c})$.

$$\text{Donc, } P(\bar{c}) + P(c) = 1 \Leftrightarrow P(e) + P(c) = 1 \Leftrightarrow P(e) = 1 - P(c) = 1 - \frac{1}{3}$$

$$\Leftrightarrow P(e) = \frac{2}{3}$$

Exercice N°03 :

On considère un lot de 35 boules identiques, numérotées de 1 à 35. Quelle est la probabilité de tirer :

- d) Une boule portant un numéro pair ?
- e) Une boule portant un numéro impair ?
- f) Une boule portant un numéro strictement supérieur à 5 ?

Solution :

L'épreuve (expérience) : « tirer une boule ».

d) L'évènement (a) : « avoir une boule portant un numéro pair ».

$$P(a) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_{17}^1}{C_{35}^1} = \frac{17}{35}$$

e) L'évènement (b) : « avoir une boule portant un numéro impair ».

$$P(b) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}} = \frac{\Omega}{N} = \frac{C_{18}^1}{C_{35}^1} = \frac{18}{35}$$

f) L'évènement (c) : « avoir une boule portant un numéro strictement supérieur à 5 ».

$$P(c) = \frac{\text{nombre de cas favorable}}{\text{nombre de cas possible}} = \frac{\Omega}{N} = \frac{C_{30}^1}{C_{35}^1} = \frac{30}{35}$$

Nb : P(a) + P(b) = 1

LOI DE PROBABILITE

En théorie des probabilités et en statistique, une *loi de probabilité* décrit le comportement aléatoire d'un phénomène dépendant du hasard.

L'étude des phénomènes aléatoires a commencé avec l'étude des jeux de hasard.

Les lois de probabilités sont des objets mathématiques qui permettent aux statisticiens de fabriquer des modèles pour décrire des phénomènes où le hasard intervient.

Une loi de probabilité est une distribution théorique de fréquences.

Variable aléatoire discrète

Une variable aléatoire discrète X est une application dont la valeur est la valeur du caractère étudié, c'est à dire le résultat d'une épreuve.

Si X prend n valeurs x_1, \dots, x_n , on définit :

- m la moyenne ou, $E(X)$ l'espérance de X par :

$$m = E(X) = \sum_{i=1}^n x_i P(X=x_i)$$
- μ_2 le moment d'ordre 2 par :

$$\mu_2 = E(X^2) = \sum_{i=1}^n x_i^2 P(X=x_i)$$
- $var(X)$ la variance de X par :

$$var(X) = \sum_{i=1}^n (x_i - m)^2 P(X=x_i) = E(X^2) - E(X)^2$$
- σ l'écart type par : $\sigma = \sqrt{var(X)}$

Loi uniforme : Une distribution de probabilité suit une **loi uniforme** lorsque toutes les valeurs prises par la variable aléatoire sont **équiprobables**. Si n est le nombre de valeurs différentes prises par la variable aléatoire,

La loi de Bernoulli

La variable aléatoire X suit une loi de Bernoulli de probabilité p , si X vaut 1 ou 0 avec les probabilités respectives p et $1-p$.

On a alors :

$$E(X)=p$$

$$E(X^2)=p$$

$$\sigma(X)=\sqrt{p(1-p)}.$$

Soit un univers Ω constitué de deux éventualités, **S** pour succès et **E** pour échec : $\Omega = \{E, S\}$.

La loi de probabilité associée à la variable de Bernoulli X telle que,

$$P(X = 0) = q$$

$$P(X = 1) = p$$

Est appelée loi de Bernoulli notée $\mathfrak{B}(1 ; p) \Rightarrow \begin{cases} E(x) = p \\ V(x) = pq \end{cases}$

La loi binomiale

Si la variable aléatoire X suit une loi binomiale $B(n,p)$, cela veut dire que X est égale au nombre de succès obtenus dans une série de n épreuves de Bernoulli de probabilité p . La variable aléatoire X peut donc prendre $n+1$ valeurs : $0,1,\dots,n$.

La loi binomiale $B(n,p)$ est la somme de n variables de Bernoulli indépendantes.

On a :

$$Pr(X=k) = C_n^k p^k (1-p)^{n-k}, \text{ pour } 0 \leq k \leq n,$$

$$E(X) = np,$$

$$V(x) = npq$$

La loi de probabilité suivie par la somme de n variables de Bernoulli où la probabilité associée au succès est p , est la loi binomiale de paramètres n et p notée : $B(n,p)$

$$\Rightarrow \begin{cases} E(x) = np \\ V(x) = npq \end{cases}$$

Loi de Poisson :

Lorsque n devient grand, le calcul des probabilités d'une loi binomiale devient très fastidieux.

Si $n \rightarrow \infty$ et $p \rightarrow 0$, alors : $X: \mathbf{B}(n; p) \rightarrow P(\lambda)$ avec $np \rightarrow \lambda$ (cette approximation est correcte si des conditions sont présentes).

Exercice N°01 :

On fabrique des pièces et on suppose que la probabilité pour qu'une pièce soit défectueuse est $p=0.05$ et donc il y a un contrôle de ces pièces.

Soit X la variable aléatoire égale à la valeur du nombre de contrôles effectués pour trouver une pièce défectueuse. Déterminer la loi de X ainsi que son espérance et son écart-type.

Ici X suit une loi géométrique de probabilité $p=0.05$.

$E(X)=1/p=20$

$\sigma(X)=\sqrt{1-p/p}=19.4935886896$

Exercice N°02 :

La variable aléatoire X est définie par :

x_i	-2	-1	0	1	2
$P(x_i)$	a	1/4	b	1/4	c

Déterminer a, b, c, si :

- La distribution de probabilité est symétrique.
- La variance est égale à l'unité.

Solution :

{ La distribution de probabilité est symétrique $\Rightarrow a = c$
 { La variance est égale à l'unité $\Rightarrow Var(x) = 1$

x_i	-2	-1	0	1	2
P_i	a	1/4	b	1/4	a
$x_i P_i$	-2a	$-\frac{1}{4}$	0	$\frac{1}{4}$	2a
$x_i^2 P_i$	4a	$\frac{1}{4}$	0	$\frac{1}{4}$	4a

$$E(x) = \sum_{i=1}^5 x_i P_i = 0$$

$$Var(x) = \sum_{i=1}^5 x_i^2 P_i - [E(x)]^2 = 8a + \frac{1}{2}$$

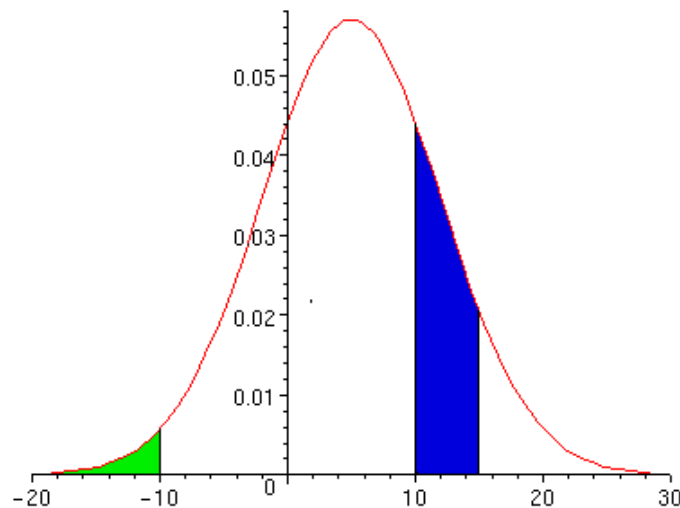
D'après les données (variance égale à l'unité) $\Rightarrow Var(x) = 1$

$$\Leftrightarrow 8a + \frac{1}{2} = 1 \quad \Leftrightarrow \begin{cases} a = \frac{1}{16} \\ c = \frac{1}{16} \end{cases}$$

On sait que :

$$\sum_{i=1}^5 P_i = 1 \quad \Rightarrow \quad 2a + \frac{1}{2} + b = 1 \quad \Leftrightarrow \quad b = \frac{3}{8}$$

$$\text{Donc : } \begin{cases} a = \frac{1}{16} \\ b = \frac{3}{8} \\ c = \frac{1}{16} \end{cases}$$



Source : A.Boukerma , AFEDDAL ,A.BENALLI Faculté de médecine d'Oran laboratoire de bio statistique

L'aire hachurée en vert correspond à la probabilité : $P(X < -10)$

L'aire hachurée en bleu correspond à la probabilité : $P(+10 < X < +15)$

Pour toutes les valeurs de x la loi de probabilité prend toute l'aire de la courbe qui est égale à 1.

- On dit qu'une variable aléatoire continue a une densité de probabilités si :

$$\begin{cases} \forall x \in \mathbb{R} & f(x) \geq 0 \\ \int_{-\infty}^{+\infty} f(x) dx = 1 \end{cases}$$

Calcul de l'espérance mathématique de la variable aléatoire continue :

$$E(x) = \int_a^b x f(x) dx$$

Calcul de la variance de la variable aléatoire continue :

$$V(x) = \int_a^b f[(x - E(x))^2] dx = \int_a^b x^2 f(x) dx - [E(x)]^2$$

La Fonction de répartition :

$$F(x) = \int_{-\infty}^x f(t) dt \quad \text{ou} \quad F(x) = \int_a^x f(t) dt \quad (a : \text{la petite valeur de l'ensemble})$$

La densité de probabilité est une forme de représentation de la loi de probabilité sous forme d'intégrale.

Exercice N°03 :

On suppose que la probabilité élémentaire pour qu'un nouveau-né soit un garçon est de 0,5. Cette probabilité est indépendante des individus.

Quelle est la probabilité pour que sur cinq nouveau-nés il y ait :

- Deux garçons ?
- Au moins deux garçons ?
- Au plus deux garçons ?
- Calculer l'espérance mathématique et la variance de la distribution des garçons.

Solution :

P : Probabilité élémentaire pour qu'un nouveau né soit un garçon.

p = $\frac{1}{2}$: Cette probabilité suit une loi binomiale $\mathcal{B}(n ; p)$

$$q = 1 - \frac{1}{2} = \frac{1}{2} \text{ (Probabilité de l'évènement contraire).}$$

$$n = 5$$

$$\text{La loi Binomiale : } P(x) = C_n^x p^x q^{n-x}$$

1. La probabilité d'avoir 2 garçons sur 5 nouveaux nés :

$$x = 2$$

$$P(2) = C_5^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = 0,3125 \quad (\text{\`a partir du tableau}).$$

2. La probabilité d'avoir au moins deux garçons :

$$P(x \geq 2) = 1 - P(x < 2)$$

$$= 1 - [P(x = 0) + P(x = 1)]$$

$$= 1 - \left[C_5^0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 + C_5^1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 \right]$$

$$= 1 - (0,0312 + 0,1562) = 0,8126$$

3. La probabilité d'avoir au plus deux garçons :

$$P(x \leq 2) = P(x = 0) + P(x = 1) + P(x = 2) = 0,0312 + 0,1562 + 0,3125 = 0,5$$

4. Calcul de l'espérance mathématique et de la variance de la distribution des garçons :

$$E(x) = np = 5 \times \frac{1}{2} = 2,5$$

$$V(x) = npq = 5 \times \frac{1}{2} \times \frac{1}{2} = 1,25$$

Exercice N°04 :

Sachant que dans une population renfermant (regroupant) des familles de 8 enfants, la probabilité d’avoir un enfant aux cheveux bruns est de 4/10.

1. Quelle est la loi de probabilité de la variable Y ?
2. Quelle est la moyenne, la variance de ce nombre d’enfants aux cheveux bruns ?
3. Quelle est la probabilité pour que dans une famille de 8 enfants, il y ait 1 enfant aux cheveux bruns ?
4. Même question pour au moins 2 enfants aux cheveux bruns.

Solution :

1. Y suit une loi binomiale parceque :

- $P = \frac{4}{10}$ Valeur fini (Pas de fonction de densité) (Y : VAD).
- \exists 2 éventualités indépendantes (Soit 1 cheveux bruns, soit non).
- $n = 8 \geq 2$ est vrai.
- Tirage est avec remise.

$$Y \curvearrowright \mathcal{B}(8; 0.4)$$

2. Calcul de la moyenne et la variance :

$$E(Y) = np = 8 \times 0,4 = 3,2 \approx \mathbf{3 \text{ enfants}}$$

En moyenne on trouve 3 enfants aux cheveux bruns.

$$V(Y) = npq = 3,2 \times 0,6 = 1,92$$

$$S = 1,38$$

3. La probabilité d’avoir 1 enfant aux cheveux bruns dans une famille de 8 enfants :

$$P(Y = 1) = C_8^1 p^1 q^7 = C_8^1 (0,4)^1 (0,6)^7 = 0,0896$$

4. La probabilité d’avoir au moins 2 enfants aux cheveux bruns dans une famille de 8 enfants :

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y < 2) \\ &= 1 - [P(Y = 0) + P(Y = 1)] \end{aligned}$$

$$= 1 - [0,0168 + 0,0896]$$

$$P(Y \geq 2) = 0,9$$

Exercice N°05 :

La probabilité pour qu'un individu ait une mauvaise réaction à l'injection d'un certain sérum est de 0,001.

Déterminer la probabilité pour que sur 2000 individus :

1. 3 individus aient une réaction dangereuse.
2. Plus de 2 individus aient une réaction dangereuse.

A justifier la loi de probabilité utilisée.

Solution :

$$p = 0,001$$

$$q = 0,999$$

$$n = 2000$$

$$X \sim \mathcal{B}(2000 ; 0.001)$$

1. La probabilité pour que 3 individus aient une réaction dangereuse :

$$P(X = 3) = C_{2000}^3 p^3 q^{1997} = C_{2000}^3 (0,01)^3 (0,999)^{1997} = 0,18$$

On peut calculer la probabilité d'une autre méthode puisque les 3 conditions de la loi de probabilité présentent :

- a. **(n)** est très grand ($n \geq 30$) ;
- b. **(p)** est très faible ($p < 0,1$) ;
- c. $\lambda = np = 2 < 5$

On peut procéder par approximation à la loi de poisson et on pose : $P(X = x_i) = \frac{\lambda^x e^{-\lambda}}{x!}$

$$P(X = 3) = \frac{2^3 e^{-2}}{3!} = 0,18$$

2. La probabilité pour que plus de 2 individus aient une réaction dangereuse :

$$P(X > 2) = 1 - P(X \leq 2)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= \frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} + \frac{2^2 e^{-2}}{2!}$$

$$= 1 - [0,135 + 0,27 + 0,27] = 0,325$$

Remarque :

$X \sim \mathcal{B}(2000; 0.001)$ et pas de poisson.

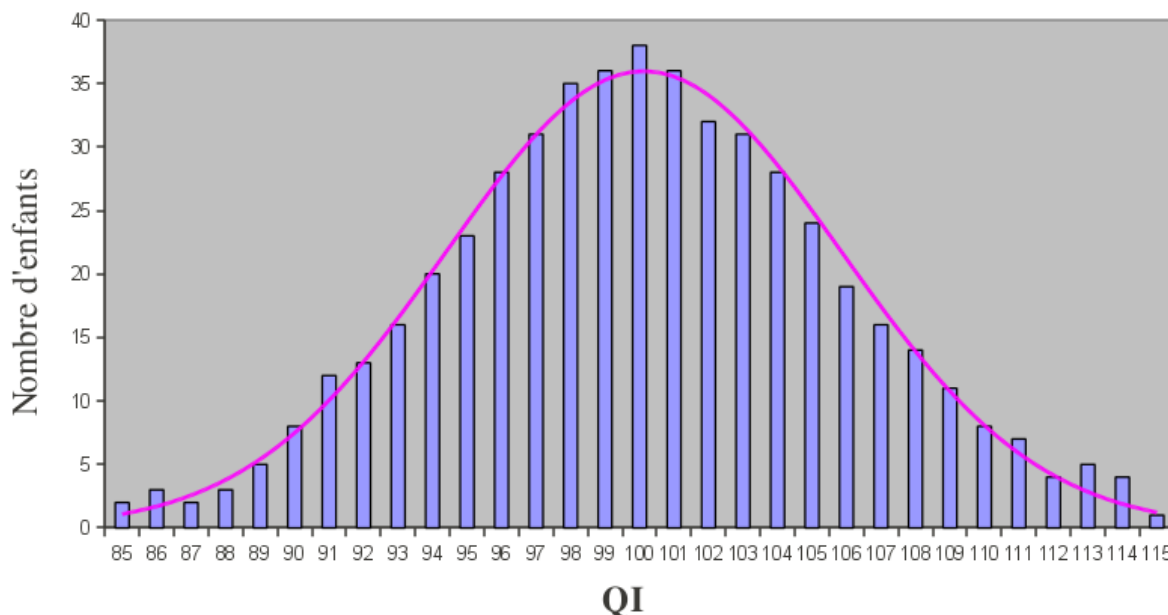
Justification :

1. X est VAD ($P=0.001$)
2. \exists 2 éventualités indépendantes
3. $n = 2000 \geq 2$
4. Le tirage est avec remise

Loi normale

En théorie des probabilités et en statistique, les lois normales sont parmi les lois de probabilité les plus utilisées pour modéliser des phénomènes naturels issus de plusieurs événements aléatoires

C'est une **loi** absolument continue, c'est-à-dire que la mesure est absolument continue par rapport à la mesure de Lebesgue. Autrement dit, il existe une densité de **probabilité**, souvent notée ϕ pour la **loi normale** centrée réduite, telle que : $N(dx) = \phi(x) dx$. Elle est généralisée par la **loi normale** multidimensionnelle.



Loi normale centrée réduite $N(0; 1)$

C'est une loi absolument continue, c'est-à-dire que la mesure est absolument continue par rapport à la mesure de Lebesgue. Autrement dit, il existe une densité de probabilité, souvent notée ϕ pour la loi normale centrée réduite, telle que : $N(dx) = \phi(x) dx$.

On suppose qu'une certaine variable $X \sim N(0; 1)$. Pour quelle proportion d'individus est-ce que $X \leq 1; 56$?

On cherche $P(X \leq 1; 56)$ (rappel : on écrit aussi $F(1, 56)$).

On cherche $P(X \leq 1; 56)$ (rappel : on écrit aussi $F(1, 56)$).

On cherche 1,56 dans la table :

Donc $P(X \leq 1, 56) = 0, 9406$.

Soit $P(X \geq 1; 49) = 1 - 0,9319 = 0,0681$.

Par définition, les variables aléatoires continues prennent des valeurs continues sur un intervalle donné.

Une variable aléatoire absolument continue X suit une loi normale de paramètres $(\mu ; \sigma)$ si sa densité de probabilité est donnée par :

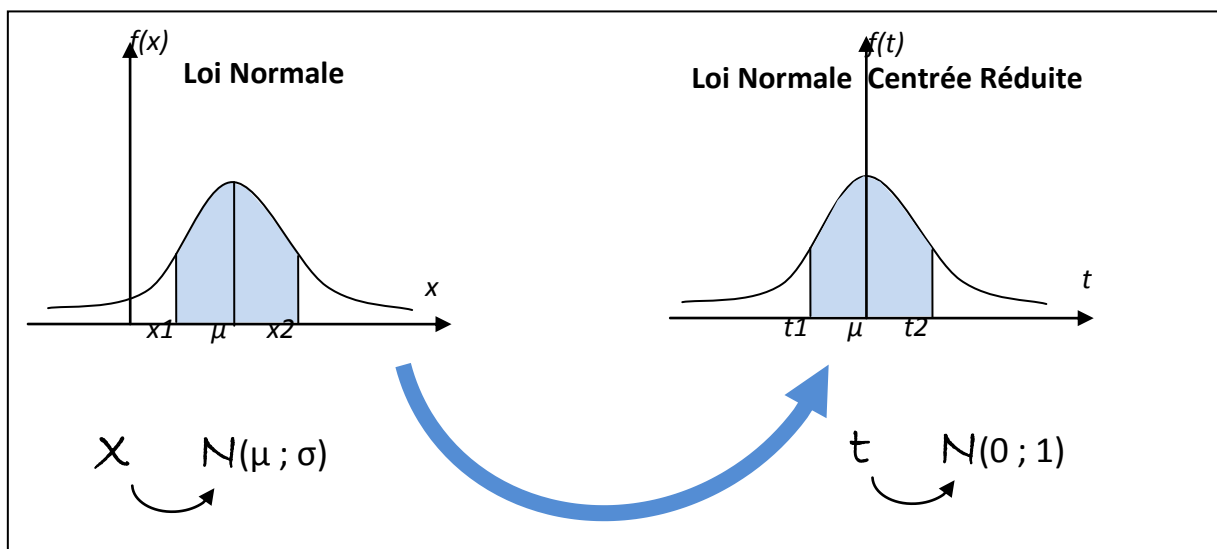
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$\text{Tel que : } \left. \begin{cases} \forall x \in \mathbb{R} : f(x) \geq 0 \\ \int_{-\infty}^{+\infty} f(x) dx = 1 \\ P(x_1 < x < x_2) = \int_{x_1}^{x_2} f(x) dx \end{cases} \right\}$$

Pour calculer la probabilité dans le cadre d'une loi Normale $X \sim N(\mu ; \sigma)$, il faut procéder par un changement de variable : $t = \frac{X-\mu}{\sigma}$ et pour cela la variable aléatoire (x) devient (t) qui va suivre une loi Normale Centré Réduite $t \sim N(0 ; 1)$

Donc la densité devient :

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \quad \text{et} \quad P(t_1 < t < t_2) = \int_{t_1}^{t_2} f(t) dt$$



Exercice N°01 :

Une variable aléatoire continue X suit une loi normale d'espérance mathématique égale à 400 et d'écart type 15,4.

Déterminer la probabilité pour que X soit supérieure à 393.

Solution :

$$\mu = 400$$

$$\sigma = 15,4$$

$$X \curvearrowright N(400, 15,4)$$

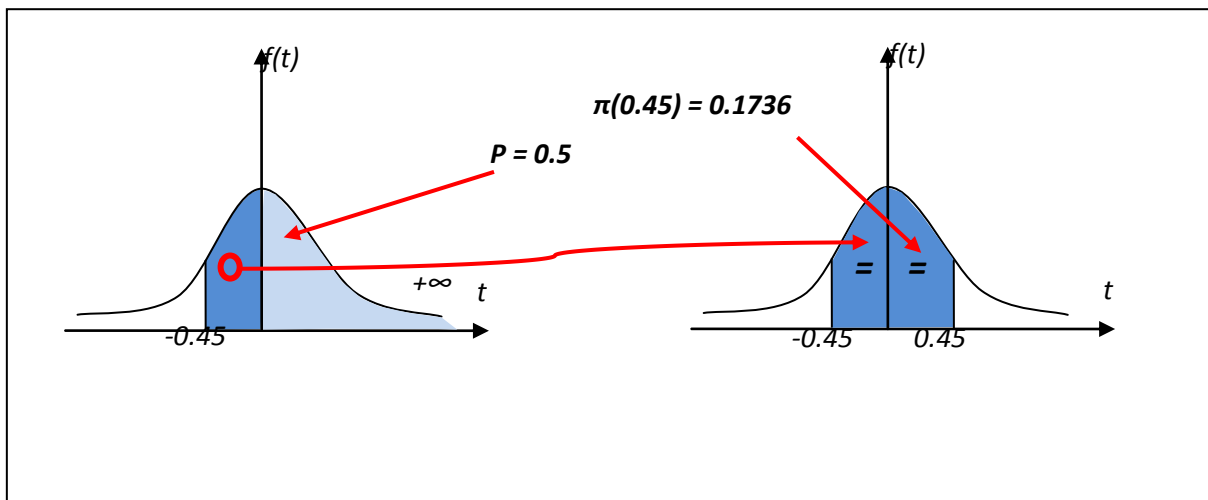
$$\text{On pose } t = \frac{x - \mu}{\sigma} \quad \rightarrow \quad dt = \frac{dx}{\sigma}$$

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \quad \text{Densité de la loi Normale Centrée réduite.}$$

$$t \curvearrowright N(0,1)$$

$$\begin{aligned} P(X > 393) &= P\left(\frac{x - \mu}{\sigma} > \frac{393 - \mu}{\sigma}\right) \\ &= P\left(t > \frac{-7}{15,4}\right) \\ &= P(t > -0,45) \\ &= \int_{-0,45}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \\ &= \int_{-0,45}^0 \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \end{aligned}$$

$$= \int_0^{0.45} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt + \frac{1}{2} = 0,1736 + \frac{1}{2} = 0,6736$$



Exercice N°02 :

La taille des élèves d’une école suit une distribution normale avec une moyenne de 150 cm et un écart type de 20 cm.

1. Quelle est la probabilité d’avoir une taille comprise entre 140 et 170 cm ?
2. Quel est le nombre d’élèves ayant une taille comprise entre 140 et 170 cm ? (L’effectif total est de 1000 élèves).

Solution :

$$\mu = 150 \text{ cm}$$

$$\sigma = 20 \text{ cm}$$

$$X \curvearrowright N(150, 20)$$

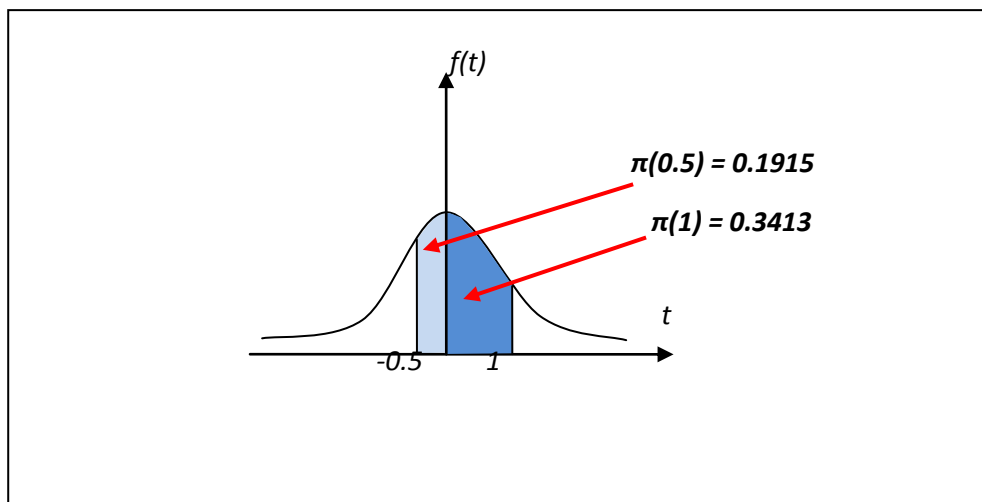
On pose $t = \frac{x - \mu}{\sigma} \rightarrow dt = \frac{dx}{\sigma}$

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \text{ Densité de la loi Normale Centrée réduite.}$$

$$t \curvearrowright N(0,1)$$

1. La probabilité d'avoir une taille comprise entre 140 et 170 :

$$\begin{aligned}
 P(140 \leq X \leq 170) &\xrightarrow{\text{Changement de variable}} P\left(\frac{x_1 - \bar{x}}{\sigma} \leq t \leq \frac{x_2 - \bar{x}}{\sigma}\right) \\
 &= P\left(\frac{140 - 150}{20} \leq t \leq \frac{170 - 150}{20}\right) \\
 &= P(-0.5 \leq t \leq 1) \\
 &= \pi(0.5) + \pi(1) \\
 &= 0.1915 + 0.3413 = 0.5328
 \end{aligned}$$



2. Le nombre d'élève ayant une taille comprise entre 140 et 170 :

$$n = N \times P(140 \leq X \leq 170) = 1000 \times 0.5328 = 532.8 \approx \mathbf{533 \text{ élèves}}$$

Sujets d'examen

EXERCICE N°1 :

Soit 1000 billets de loterie, un des billets gagne 500DA, 10billets 100 DA chacun, 50 billets 20DA chacun,100 billets 5 DA chacun, tout les autres ne gagne rien.

Trouvez la probabilité pour le propriétaire d'un seul billet de gagner au moins 20DA.

EXERCICE N°2

a/ Quelle est la probabilité d'obtenir K faces avec une pièce lancée n fois successivement, ou avec une pièces lancées simultanément.

b/ Calculer $P(X=K)$ avec $K= 0,1, 2, 3, 4,5$ et $n= 5$.

c/ Calculer l'espérance mathématique et l'écart type

EXERCICE N°3

Dans un amphi présentant 120 étudiants dont 46 sont de sexe féminin, X_i est une variable aléatoire suivant une loi normale de moyenne 89 et d'écart type 7,5 .

Calculez :

a/ la probabilité pour que la demande soit inférieure à 92.

b/ la probabilité pour que la demande soit supérieure à 95

c/ la probabilité pour que la demande soit comprise entre 97 et 100

EXERCICE N° 4 :

Sachant que 0.53 des nouveau-nés sont des garçons .La variable aléatoire X est le nombre de garçons parmi 1000 nouveau-nés .

1° - a / Donner la loi de probabilité .

-b / Calculer l'espérance mathématique, la variance et l'écart type .

2° Calculer la probabilité d'avoir au plus 350 garçons.

3° Calculer la probabilité d'avoir de 350 à 550 garçons .

4° Calculer la probabilité d'avoir au moins 450 garçons .

2° Calculer la probabilité d'avoir au plus 350 garçons.

3° Calculer la probabilité d'avoir de 350 à 550 garçons .

4° Calculer la probabilité d'avoir au moins 450 garçons .

Références bibliographiques

- Statistique cours et exercices pour les étudiants du tronc commun biomedical

INESSM ALGER / ADMANE.O HOANG-KY / OUAkli N 1990

- Polycopie des cours de statistique A.Boukerma , AFEDDAL ,A.BENALLI Faculté de médecine d'Oran laboratoire de bio statistique 2010
- Cours de statistiques d'IUT Jean-Michel BERNABOTTO
- Notion de probabilité HEC MONTREAL centre d'aide en mathématique.
- Cours et exercices de statistique appliquée Abdelaziz BERRAH 1975.
- Polycopie des cours Bio statistique Département pharmacie Oran Année 2008/2009 Faculté de médecine Oran.
- Cours et exercices de statistique E.AZOULAY / D.COHEN 1968