



---

# **Biostatistique en PASS**

**Faculté Lyon Est – UE 3**

**Pr Pascal ROY**

30/08/2024

1

**1. Présentation de l'équipe / programme**

**2. Introduction de l'UE 3**

**3. Statistiques descriptives**

**1. Présentation de l'équipe / programme**

**2. Introduction de l'UE 3**

**3. Statistiques descriptives**

# Equipe Pédagogique Hospitalo-Universitaire

Dr Claire	BARDEL	MCU-PH
Pr Delphine	MAUCORT-BOULCH	PU-PH
Dr Muriel	RABILLOUD	MCU-PH
Dr Nicolas	ROMAIN SCELLE	AHU
Pr Pascal	ROY	PU-PH

Service de Biostatistique des HCL  
Equipe Biostatistique Santé, UMR 5558 UCBL CNRS  
<http://mastersantepublique.univ-lyon1.fr>

# **Médecine Factuelle** **(Evidence-Based Medicine)**

## **Littérature scientifique médicale**

→ Diffusion des connaissances médicales

→ Décision médicale

# Accès à l'Information Médicale

## La lecture critique d'articles (LCA) (articles scientifiques médicaux)

- Extraire l'information appropriée
- Identifier l'information nouvelle
- Evaluer sa contribution à la connaissance

# Le rapport d'étude

= Base de la communication scientifique

Question d'étude

Synthèse de la littérature

Dessin d'étude

échantillon, recrutement,  
taille, données

Plan d'analyse statistique

Règles de publication

**Protocole de Recherche**



**Rapport**

## But de la Communication Scientifique

- Permettre un large accès aux découvertes
- Présenter les résultats de façon compréhensible
- Préserver l'intégrité de l'information

Salmi L-R. Lecture critique et communication médicale scientifique. Comment lire, rédiger et publier une étude clinique ou épidémiologique ? 3ème Edition. Elsevier Masson SAS; 2012.



## Quand (cursus médecine)

Année 1	→ Biostatistique* Pharmacologie
Année 2	→ Epidémiologie Pharmacologie Economie de la Santé
Année 3	→ Lecture Critique d'Articles LCA Pharmacologie Biomédecine Quantitative**
Années 4 à 6	→ LCA (modules de spécialité)

\*Beuscart R, Benichou J, Roy P, Quantin C Eds. Biostatistic. Omniscience

# Une question, un dessin d'étude, un niveau de preuve

Pour évaluer

- Un test diagnostique
- Une intervention de prévention
- Une intervention thérapeutique
- Un facteur du risque
- Un facteur pronostique

Pour décrire un phénomène

Pour faire une synthèse des connaissances

Pour générer des hypothèses

Salmi L-R. Lecture critique et communication médicale scientifique. Comment lire, rédiger et publier une étude clinique ou épidémiologique ? 3ème Edition. Elsevier Masson SAS; 2012.

## Quels types d'études en Recherche Clinique, en Recherche Epidémiologique

- Essais thérapeutiques (comparatifs O/N, randomisés O/N)
- Diagnostique
- Prognostique
- Descriptive (Qui ? Quand ? Où ?)
- Cas-témoins
- Cohorte

Roy P, Maucort-Boulch D: Eléments méthodologiques pour la lecture critique d'articles. [Methodological tools for critical reading of scientific articles in medicine]. *Rev Prat* 2009;**59**:1119-22.

## CALENDRIER DES RENTRÉES 2024

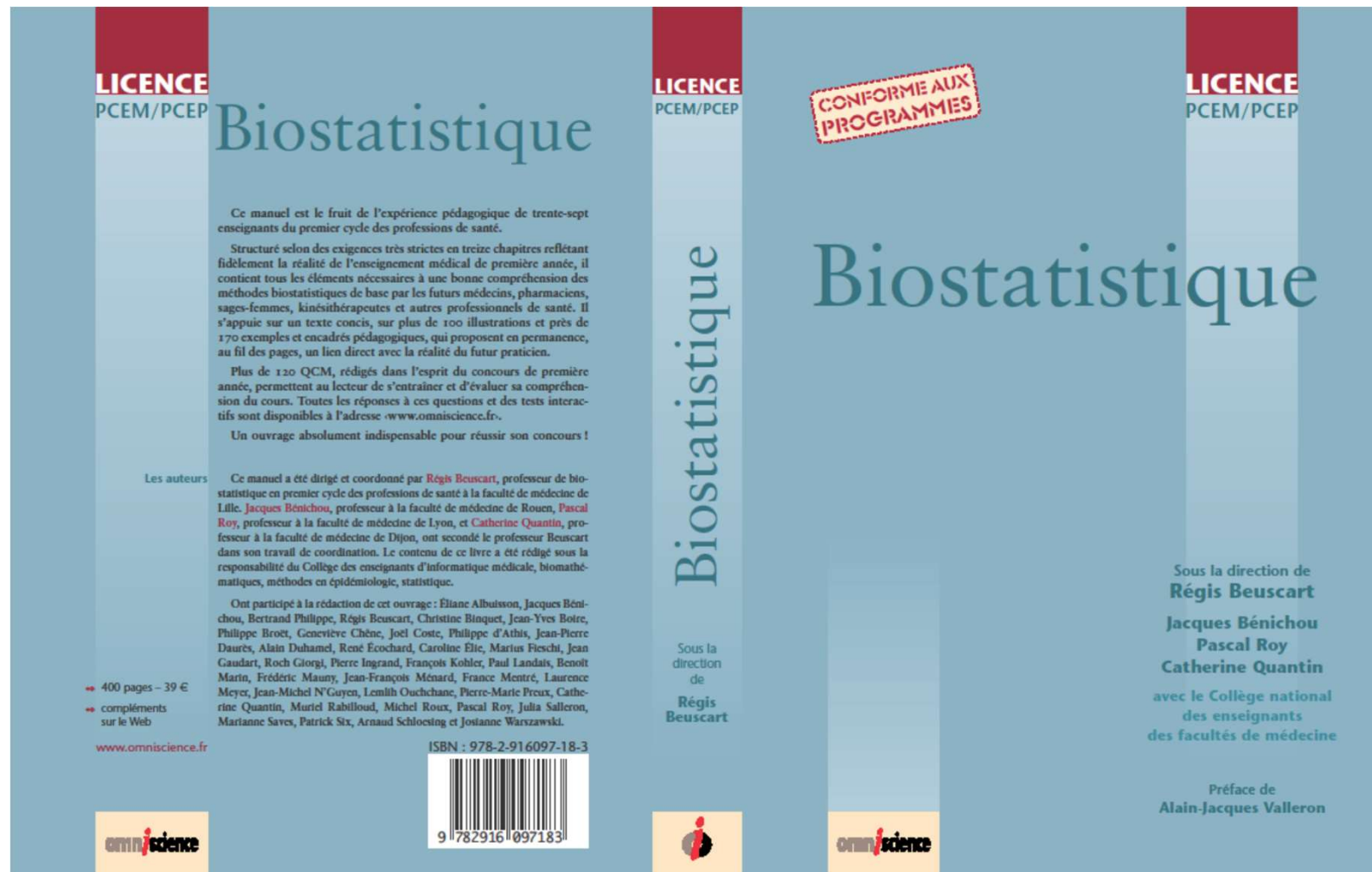
Retrouvez toutes les dates de la rentrée  
2024

[Lire la suite →](#)

1/1



# Un ouvrage conseillé



30/08/2024

13

# Interactions

## Avec vos enseignants

2 séances d'études de cas

## Vos tuteurs 2023-2024

### Questions

Recueil des questions (Tuteurs)

Réunions de synthèse Tuteurs / Enseignants

→ Réponses sur serveur, ++ structurées par chapitre

### Colles, et concours blanc

→ Sujets et corrections de colles validés

→ Sujet et correction du concours blanc validés

# Après la PASS → Formation complémentaire

du plaisir d'apprendre

→ **au double diplôme** s'inscrivant dans un projet de carrière académique ou industriel !

**Certificat RB7 Biostatistique et Modélisation**

Responsable Dr Maxime BONJOUR

**Certificat RB38 Intelligence Artificielle en Santé**

Responsables Pr Pascal ROY / Pr Delphine MAUCORT-BOULCH

**Master Santé Publique**

Coordonnateur Pr Pascal ROY

**Biostatistics, Bioinformatics, Biomathematics for Health (B3H)**

**Methods for Drug Development and Evaluation (MDDE)**

**1. Présentation de l'équipe / programme**

**2. Introduction de l'UE 3**

**3. Statistiques descriptives**



# Une Problématique (1)

Un patient n'est jamais identique à un autre !

sa maladie, sa réponse au traitement, son pronostic, tout varie .....

→ **Comment Analyser, Comprendre, Décider dans un monde où la variabilité est la règle ?**

# Une Problématique (2)

*Si les patients étaient identiques entre eux, il n'y aurait pas besoin de biostatistique – mais il n'y aurait pas de médecine non plus : un médecin disposant de la « notice de fonctionnement de l'homme » y suffirait.*

Valleron AJ *Préface*. Beuscart R, Benichou J, Roy P, Quantin C.  
*Biostatistique*, Omniscience ®

**Les questions posées sont  
d'ordre général**

**Mais...**

**On dispose du résultat d'études  
réalisées sur des échantillons**

# Exemple 1

Variabilité du résultat d'un test diagnostique

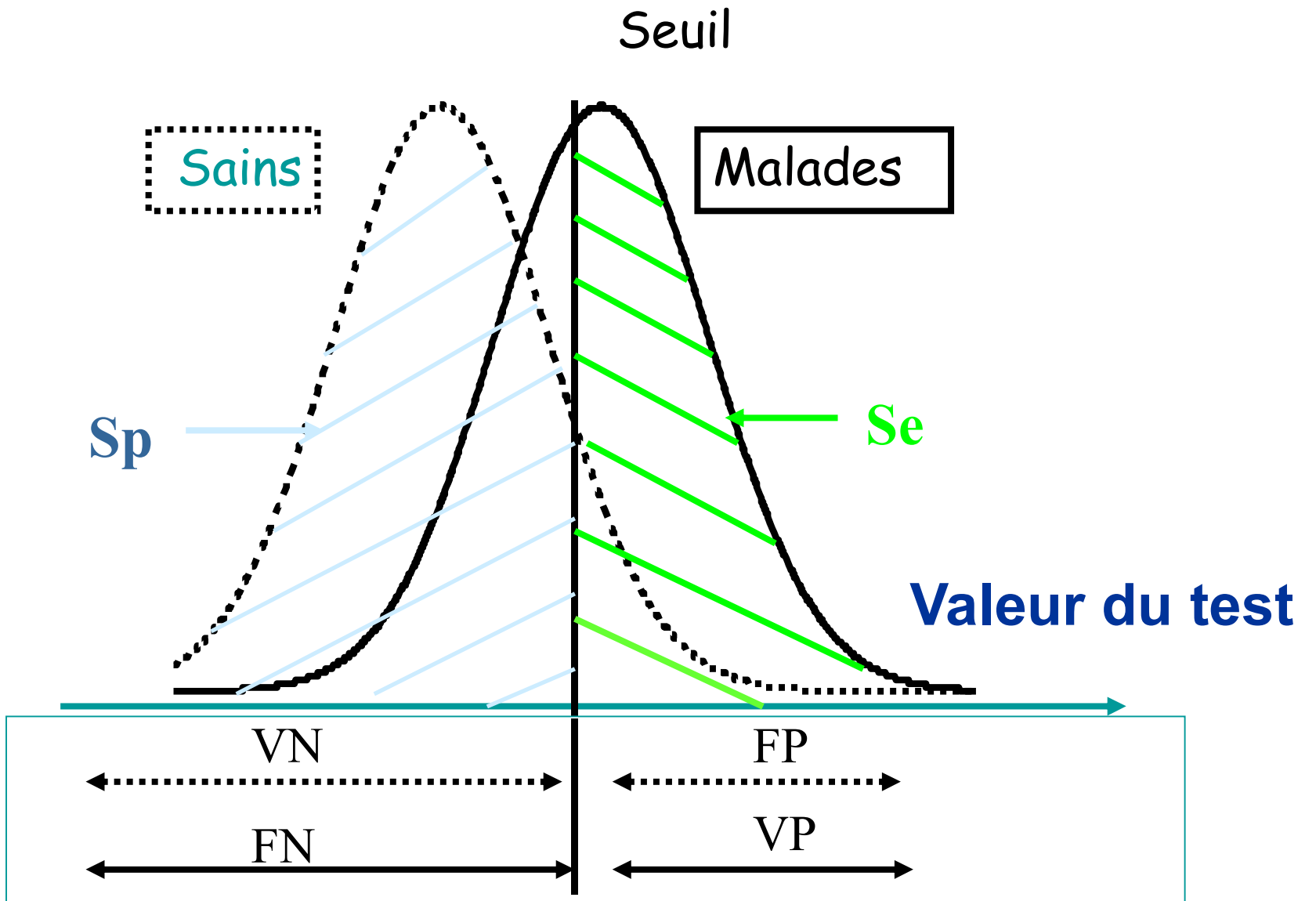
Résultats théoriques ...

d'un test diagnostique parfait sur un échantillon de 40 sujets malades et 80 sujets indemnes

	<b>Malade</b>	<b>Non Malade</b>
<b>Test positif</b>	<b>40</b>	<b>0</b>
<b>Test négatif</b>	<b>0</b>	<b>80</b>
	<b>40</b>	<b>80</b>

Résultat observé du test diagnostique sur un échantillon de 40 sujets malades et 80 sujets indemnes

	<b>Malade</b>	<b>Non Malade</b>
<b>Test positif</b>	<b>36</b>	<b>8</b>
<b>Test négatif</b>	<b>4</b>	<b>72</b>
	<b>40</b>	<b>80</b>

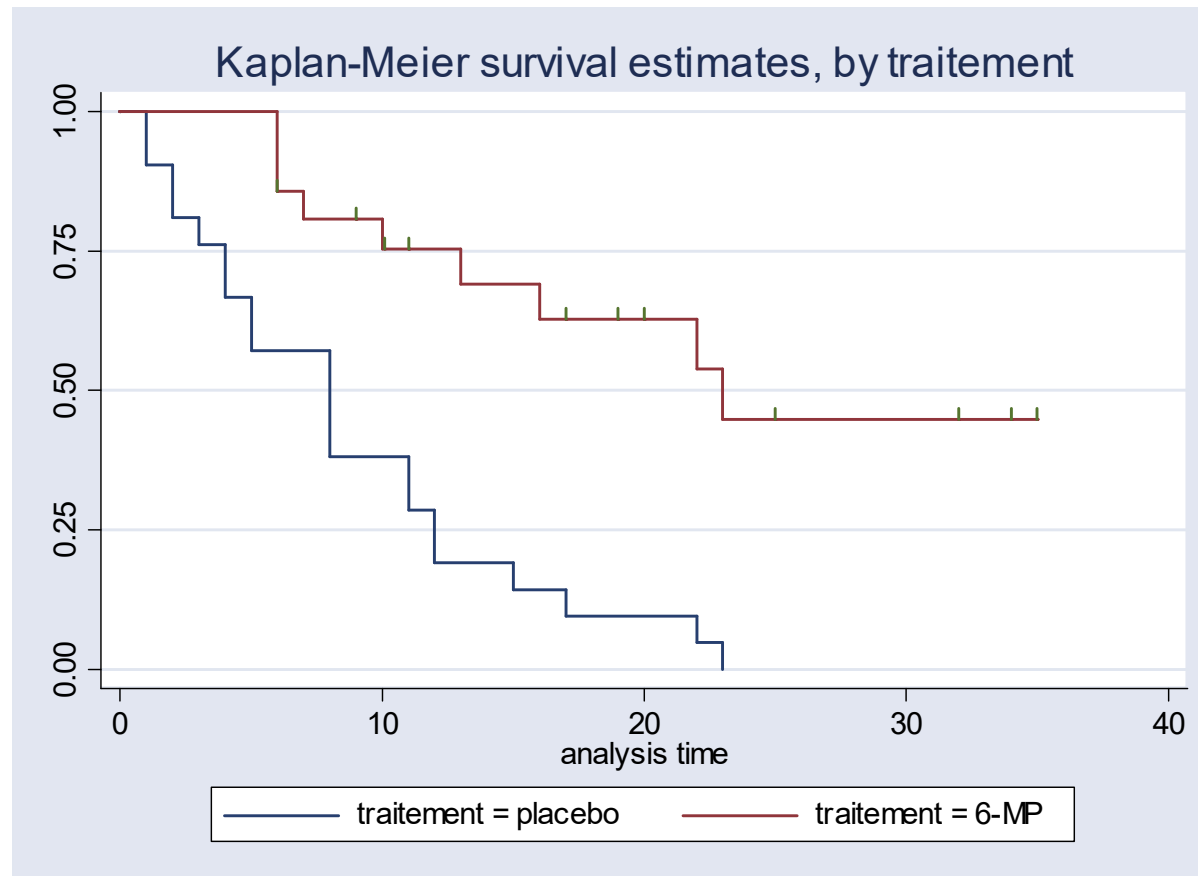


# Exemple 2

**Comparer l'efficacité de deux traitements sur la survenue d'une rechute (LAL, essai randomisé)**



# Rechute 6-MP vs Placebo (données d'Anderson)



**Les questions posées sont  
d'ordre général**

**Mais...**

**On dispose du résultat d'études  
réalisées sur des échantillons**

# INFERENCE STATISTIQUE

## → ESTIMATIONS

Epidémiologie Descriptive  
**Risque de cancer**

Recherche Clinique  
**Probabilités de guérison  $\neq$**   
**Modèle pronostique**

Intervalles de confiance à 95%

**Si l'étude était reproduite un très grand nombre de fois, 95% des intervalles de confiance obtenus contiendraient la vraie valeur du paramètre estimé.**

# Méthodologie de la recherche

(Clinique / Epidémiologique)

---

**Une SEULE question**

**dessin d'étude adapté**

**plan d'analyse adapté**

**Déroulement de l'étude**

**Analyse des résultats**

**Interprétation des résultats**

d'après Roy P., Maucort-Boulch D. *Revue du Praticien* 2009;**59**:1119-1122

# Comment analyser les études ?

---

→ **Collecte des données**

→ **Analyse proprement dite**

Analyse déductive ou descriptive

Analyse inductive ou inférence

**1. Présentation de l'équipe / programme**

**2. Introduction de l'UE 3**

**3. Statistiques descriptives**

# Population et Echantillon

**Population**

**Echantillon**

**Echantillon représentatif**

# Unité statistique

L'unité statistique est un élément de l'échantillon.

L'ensemble des unités statistique constitue l'échantillon.

On remplacera volontiers par la suite le terme « unité statistique » par le terme « individu ».

# Distribution statistique

Il s'agit de l'ensemble des couples  $(x_i, n_i)$ , où  $x_i$  est une modalité de la variable  $x$ , et  $n_i$  le nombre de fois où cette modalité est observée (effectif ou fréquence absolue).

Elle peut aussi être l'ensemble des couples  $(x_i, f_i)$ , où  $f_i$  est la proportion (ou réquence relative) de la modalité  $x_i$  :

$f_i = n_i / n$  ( $n$  étant l'effectif total des observations).



## **3.1. Différents types de variables et de tableaux**

### **3.1.1. Variables qualitatives nominales**

**(sans relation d'ordre)**

**sexe, groupe sanguin**

**Sémiologie initiale d'une maladie**

### **3.1.2. Variables qualitatives ordinales**

**intensité d'une relation inflammatoire**

**Stade d'une maladie**

# Quand la variable étudiée est qualitative nominale

## Localisation de l'infarctus

Modalité	Effectif	Proportion (%)	Fréquence cumulée (%)
antérieur	935	38.9	38.9
non antérieur	1177	48.9	87.8
sans onde Q	293	12.2	100.0
Total	2405	100.0	

# Quand la variable étudiée est qualitative binaire

## Sexe

Modalité      Effectif    Proportion (%)      Fréquence cumulée (%)

homme	1718	68.2	68.2
femme	801	31.8	100.0
Total	2519	100.0	

# Quand la variable étudiée est catégorielle ordinale

## Score de Killip

Modalité      Effectif    Proportion (%)      Fréquence cumulée (%)

stade 1	1670	66.6	66.6
stade 2	529	21.1	87.7
stade 3	253	10.1	97.8
stade 4	55	2.2	100.0
Total	2507	100.0	

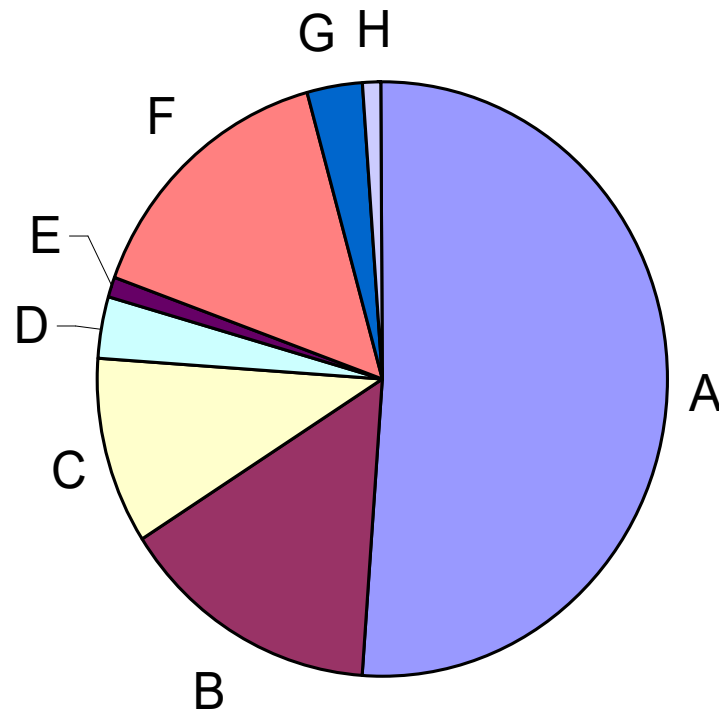
## Distribution de 1191 patients atteints de Sclérose en Plaques (SEP) selon la sémiologie initiale

Sémiologie initiale		effectifs	fréquence relative %
Atteinte FI isolée	(A)	609	51.1
No isolée	(B)	176	14.8
Atteinte Tc isolée	(C)	121	10.2
Atteinte FI et No	(D)	41	3.4
Atteintes Tc et No	(E)	12	1.0
Atteintes FI et Tc	(F)	184	15.4
Atteintes FI, Tc, No	(G)	35	2.9
Autres	(H)	13	1.1
<b>Totaux</b>		<b>1191</b>	<b>100.0</b>

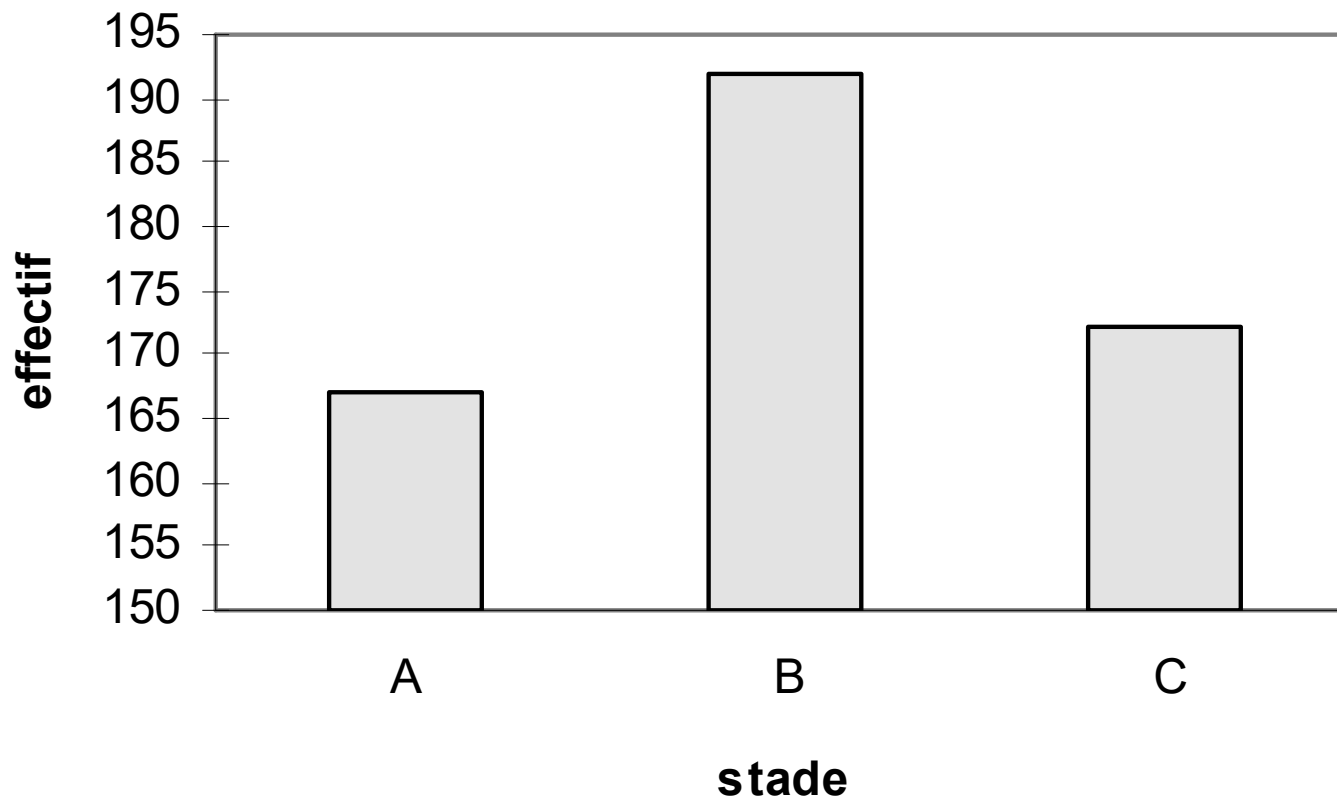
[FI : Fibres longues, No: Névrite optique, Tc: Tronc cérébral]

30/08/2024

## Distribution des 1191 patients atteints de SEP selon la semiologie initiale: diagramme circulaire



## Distribution de 531 patients atteints d'un cancer du rectum selon le stade de Dukes



## DISTRIBUTION DES GENOTYPES (taille de l'échantillon = n individus)

25374 individus sont homozygotes avec 2 allèles normaux

5482 individus sont hétérozygotes avec 1 allèle normal et un allèle muté

67 individus sont homozygotes avec 2 allèles mutés

---

GENOTYPES	AA	AS	SS	TOTAL
Notation	$n_1$	$n_2$	$n_3$	$n$
Effectifs	25 374	5 482	67	30 923

---

## FREQUENCE DES ALLELES NORMAUX ET PATHOLOGIQUES

Si le nombre total d'individus est  $n$ , alors le nombre total de gènes dans l'échantillon est  $2n$  ( $2 \times 30923$ ) puisque chaque individu de l'échantillon est diploïde.

fréquence de l'allèle normal A :  $p = (2 n_1 + n_2) / 2n = 0.90919$

fréquence de l'allèle muté S :  $q = (n_2 + 2 n_3) / 2n = 0.09081$  avec  $p + q = 1$



## 3.1. Différents types de variables et de tableaux

### 3.1.3. Variables quantitatives

#### discrètes

nombre d'enfants, titre d'anticorps sériques

#### continues

taille, poids, pression artérielle, concentration

plasmatique du cholestérol

discrétisation (taille en cm, âge en années)

## Série statistique (échantillon de taille n)

$(x_1, x_2, \dots, x_i, \dots, x_n)$ ,  $x_i$  étant la valeur de  $X$  pour le  $i^{\text{ème}}$  élément observé,  $i = 1, \dots, n$ .

## Série statistique triée

$(x_{(1)}, x_{(2)}, \dots, x_{(i)}, \dots, x_{(n)})$

avec  $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(i)} \leq \dots \leq x_{(n)}$ .

## Série double (deux variables), n couples de valeurs

## Exemple

### La série

$$(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) =$$
$$(15, 3, 623, 46, 126, 64, 1350, 279, 23)$$

### devient

$$(x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}, x_{(6)}, x_{(7)}, x_{(8)}, x_{(9)}) =$$
$$(3, 15, 23, 46, 64, 126, 279, 623, 1350)$$

### après cette opération

# Statistiques descriptives à une variable

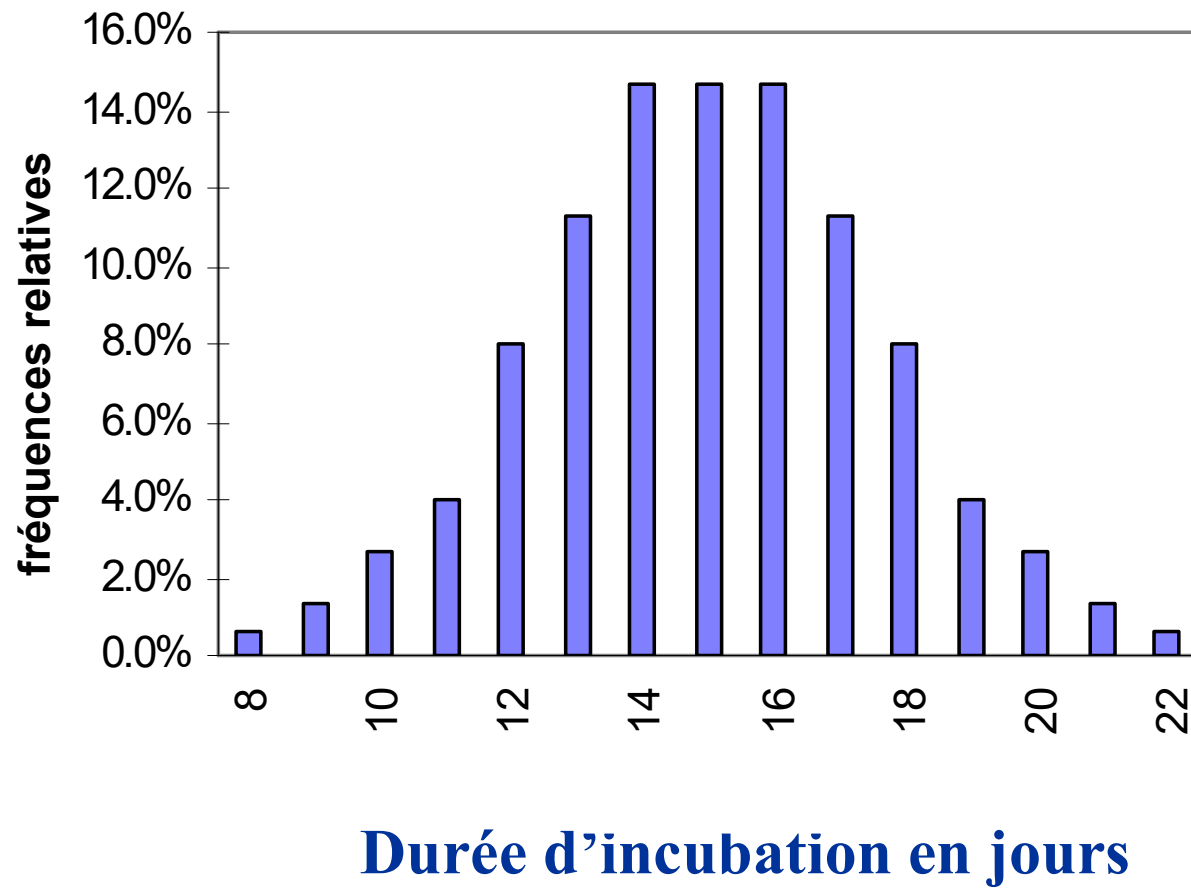
## Cas d'une variable quantitative

Délai d'incubation (en jours) d'une maladie infectieuse virale,  $n = 150$ .

17	8	19	9	9	17	10	19	16	10
10	16	16	10	17	19	11	18	11	15
15	11	11	11	11	17	12	12	12	12
12	12	17	12	12	18	12	12	13	13
13	20	13	13	17	13	19	15	15	17
13	18	13	13	17	13	13	17	13	13
17	14	18	14	14	14	14	18	14	14
12	19	14	14	21	14	14	15	14	14
19	21	14	14	14	14	14	14	14	15
18	15	16	18	15	15	18	17	15	15
20	15	13	17	13	20	13	13	22	15
15	15	20	18	15	17	15	15	15	15
15	12	15	16	16	17	16	16	18	16
16	16	16	16	17	16	18	14	17	14
16	16	16	17	16	16	16	16	16	18

<b>Valeur</b>	<b>Effectif</b>	<b>Proportion (%)</b>	<b>fréquence cumulée</b>
	<b>fréquence absolue</b>	<b>fréquence relative</b>	<b>(%)</b>
<b>8.00</b>	<b>1</b>	<b>.7</b>	<b>.7</b>
<b>9.00</b>	<b>2</b>	<b>1.3</b>	<b>2.0</b>
<b>10.00</b>	<b>4</b>	<b>2.7</b>	<b>4.7</b>
<b>11.00</b>	<b>6</b>	<b>4.0</b>	<b>8.7</b>
<b>12.00</b>	<b>12</b>	<b>8.0</b>	<b>16.7</b>
<b>13.00</b>	<b>17</b>	<b>11.3</b>	<b>28.0</b>
<b>14.00</b>	<b>22</b>	<b>14.7</b>	<b>42.7</b>
<b>15.00</b>	<b>22</b>	<b>14.7</b>	<b>57.3</b>
<b>16.00</b>	<b>22</b>	<b>14.7</b>	<b>72.0</b>
<b>17.00</b>	<b>17</b>	<b>11.3</b>	<b>83.3</b>
<b>18.00</b>	<b>12</b>	<b>8.0</b>	<b>91.3</b>
<b>19.00</b>	<b>6</b>	<b>4.0</b>	<b>95.3</b>
<b>20.00</b>	<b>4</b>	<b>2.7</b>	<b>98.0</b>
<b>21.00</b>	<b>2</b>	<b>1.3</b>	<b>99.3</b>
<b>22.00</b>	<b>1</b>	<b>.7</b>	<b>100.0</b>
<b>Total</b>	<b>150</b>	<b>100.0</b>	<b>100.0</b>

## Distribution des durées d'incubation Diagramme en barres



30/08/2024

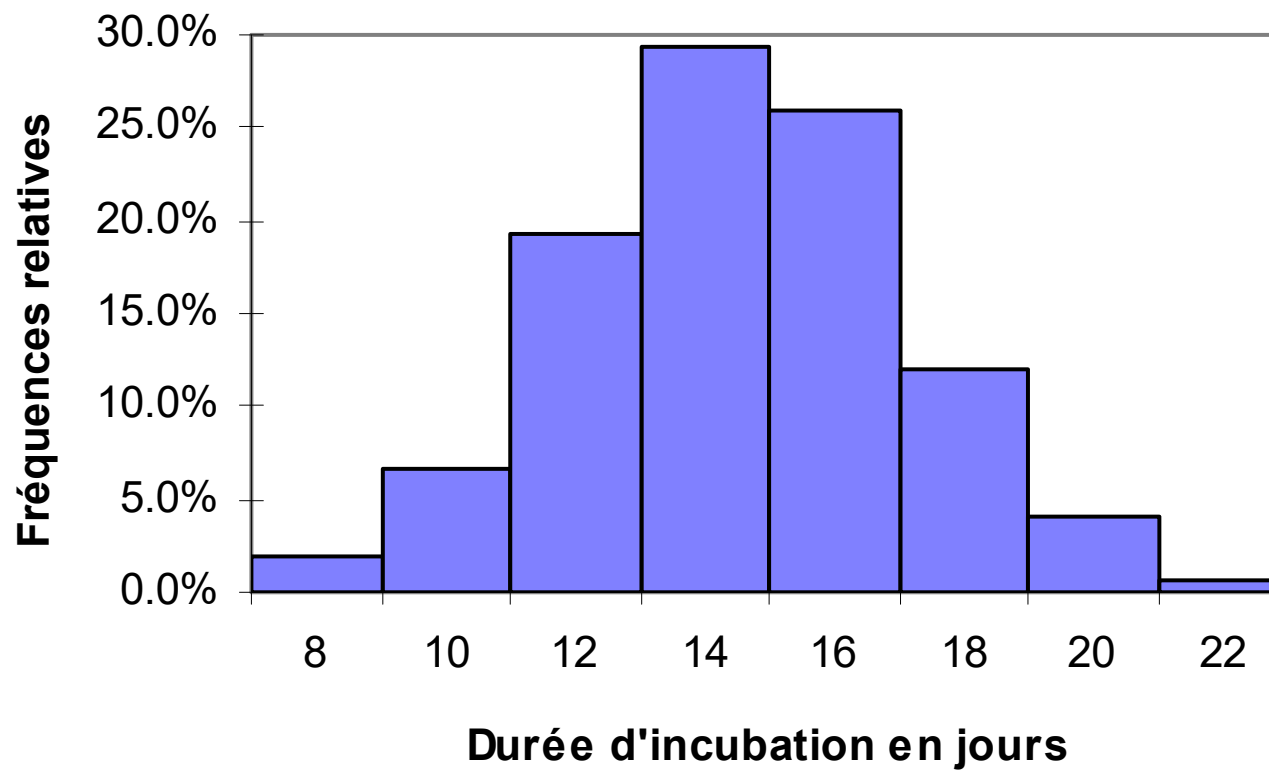
# Distribution des classes des durées d'incubation

## Tableau de fréquences absolues et relatives

Durée (en jours)	Durée(en jours)	Effectif	Fréquence relative (%)
Classe	Centres des classes	$n_i$	$(f_i)$
]7;9]	8.00	3	2.0
]9;11]	10.00	10	6.7
]11;13]	12.00	29	19.3
]13;15]	14.00	44	29.3
]15;17]	16.00	39	26.0
]17;19]	18.00	18	12.0
]19;21]	20.00	6	4.0
]21;23]	22.00	1	0.7
Total		----- 150	----- 100.0

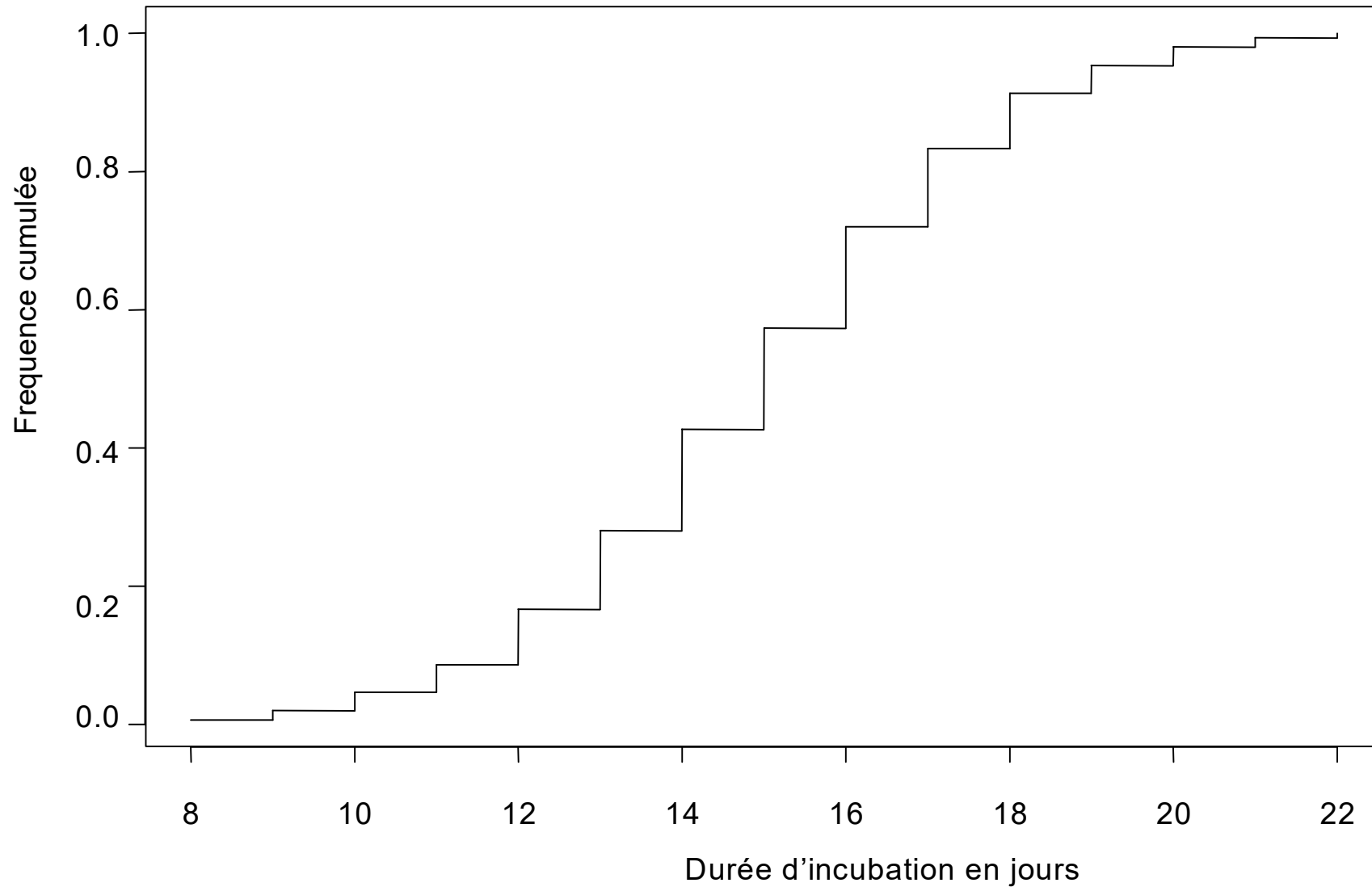


### Distribution des durées d'incubation (groupées en classes) de la maladie A chez les 150 enfants: histogramme



<b>Valeur</b>	<b>Effectif</b> fréquence absolue	<b>Proportion (%)</b> fréquence relative	<b>fréquence cumulée</b> (%)
8.00	1	.7	.7
9.00	2	1.3	2.0
10.00	4	2.7	4.7
11.00	6	4.0	8.7
12.00	12	8.0	16.7
13.00	17	11.3	28.0
14.00	22	14.7	42.7
15.00	22	14.7	57.3
16.00	22	14.7	72.0
17.00	17	11.3	83.3
18.00	12	8.0	91.3
19.00	6	4.0	95.3
20.00	4	2.7	98.0
21.00	2	1.3	99.3
22.00	1	.7	100.0
<b>Total</b>	<b>150</b>	<b>100.0</b>	<b>100.0</b>

# Histogramme cumulatif des fréquences



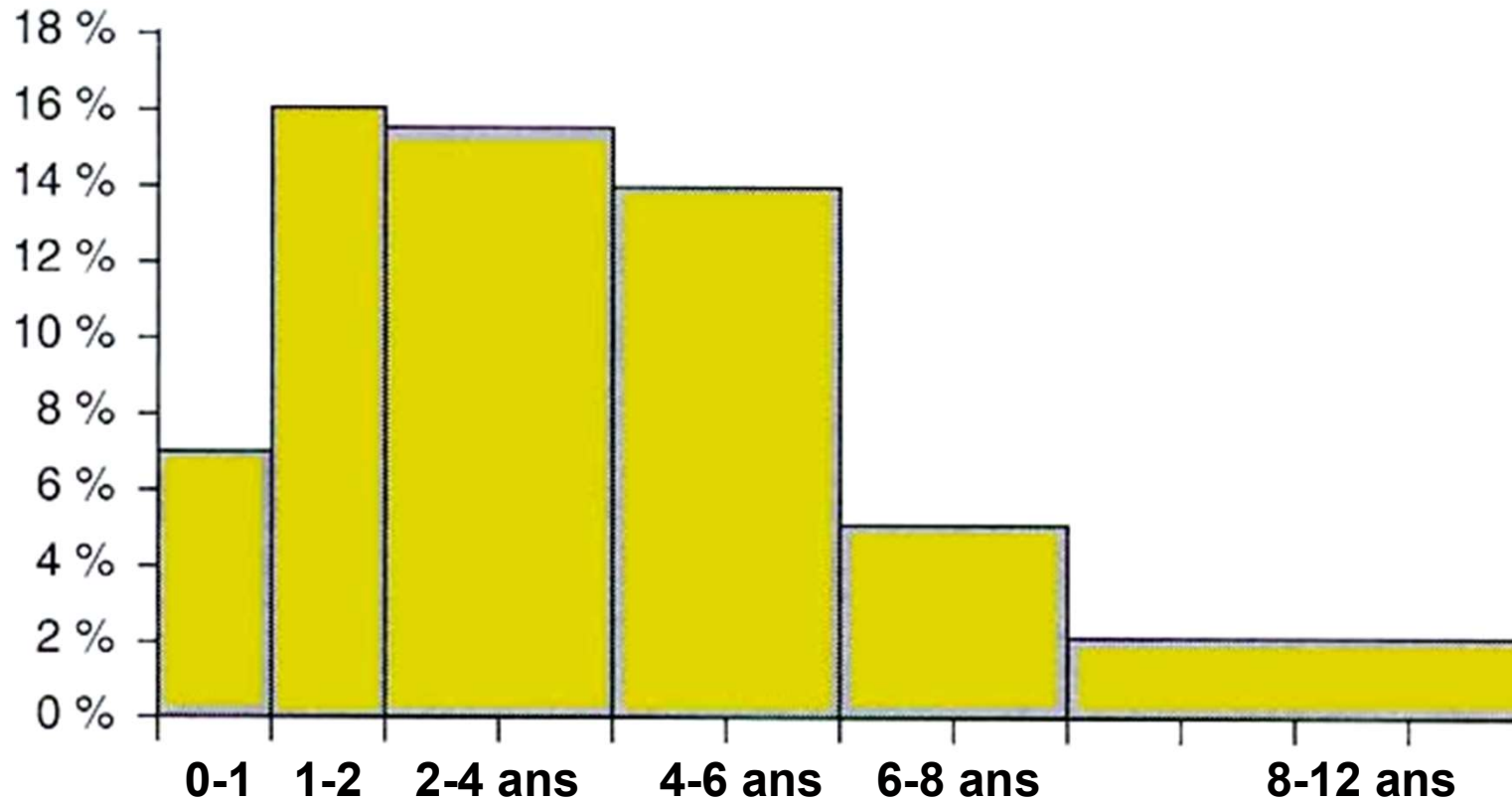
30/08/2024

# Histogramme : cas des classes inégales

Densité de fréquence

$$d_i = \frac{f_i}{\text{amplitude}_i}$$

## Age des enfants présentant une varicelle, 2001, Trouville



Age	0-1	1-2	2-4	4-6	6-8	8-12	Total
n <sub>i</sub>	13	30	58	52	19	15	187
%/an	7.0	16.0	15,5	13.9	5.1	2.0	

30/08/2024

53

# Représentation des valeurs d'une variable quantitative

## Paramètres de position

## La moyenne arithmétique

$$\bar{x} = \frac{\sum x_i}{n}$$

Maladie A. Incubation moyenne de 15 jours.

**La médiane partage la série des valeurs triées en 2 parties de même effectif. Il y a autant de valeurs inférieures que de valeurs supérieures à la médiane**

**Si n est impair : valeur de rang  $(n+1)/2$**

**Si n est pair : toute valeur de  $]X_{(n/2)} ; X_{((n/2)+1}[$   
(et en pratique, la moyenne de ces 2 valeurs)**

Maladie A. Incubation médiane de 15 jours.



## Le mode ou classe modale

**Le mode est la valeur observée avec la plus grande fréquence. Sa valeur s'obtient directement à partir du tableau statistique ou du diagramme en bâtons.**

**La classe modale est celle qui correspond au plus grand effectif si toutes les classes ont la même amplitude.**

# Représentation des valeurs d'une variable quantitative

## Paramètres de dispersion

**La variance est un paramètre de dispersion par rapport à la moyenne**

**La variance de l'échantillon (variance empirique) présentée ci-dessous fournit une estimation biaisée (car sous-estimée) de la variance de la variable en population**

$$\frac{\sum (x_i - \bar{x})^2}{n} = \frac{\sum x_i^2}{n} - \bar{x}^2$$

Maladie A. Variance de la durée d'incubation de 6.9 jours<sup>2</sup>.

**L'écart-type est la racine de la variance.  
L'écart type empirique est donc la racine de  
la variance empirique**

$$\text{écart-type} = \sqrt{\text{variance}}$$

# Extrêmes et étendue

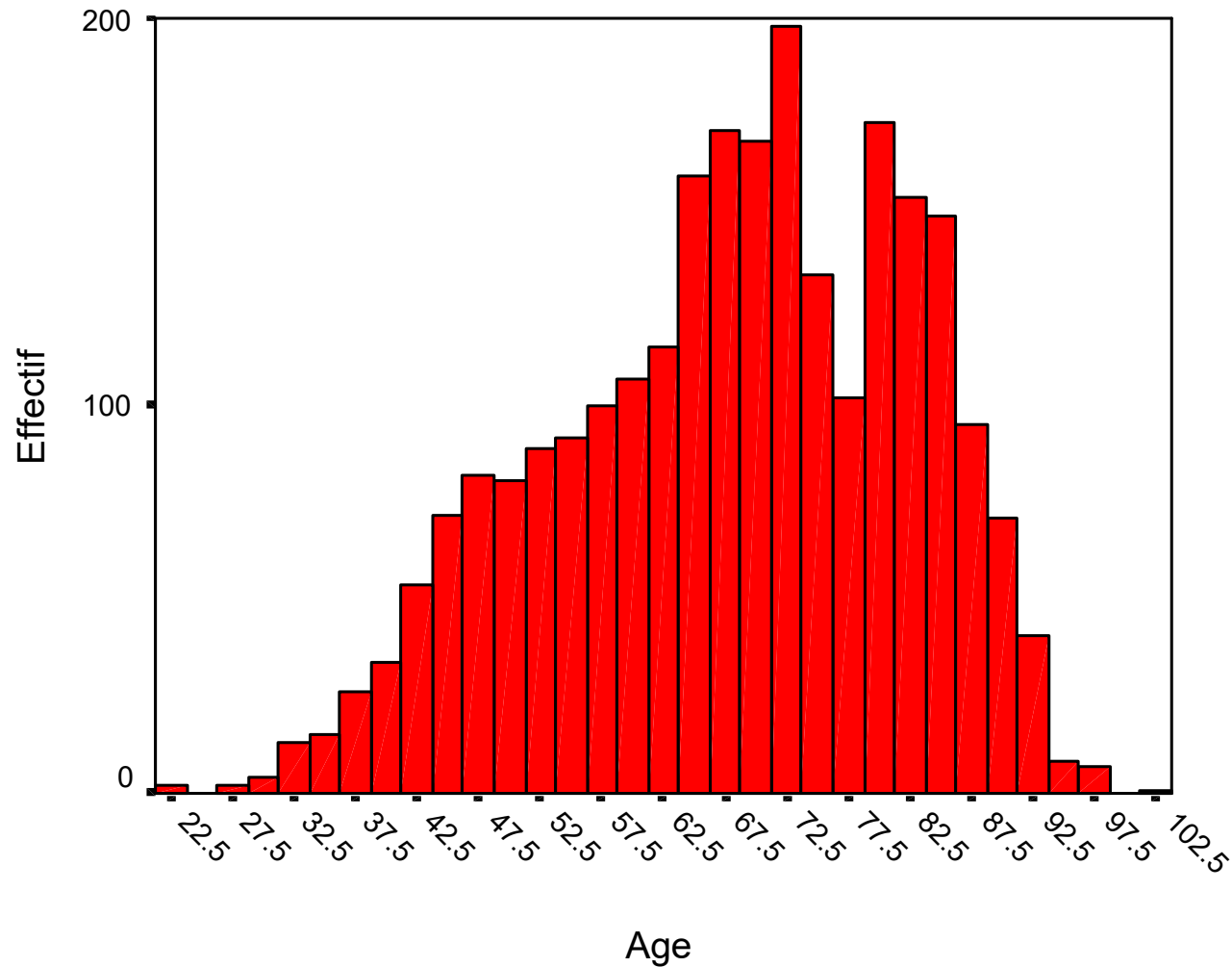
## Coefficient de variation

$$CV = \frac{\text{écart-type}}{\text{moyenne}}$$

Quantile de  $p$  : valeur de rang  $p(n+1)$

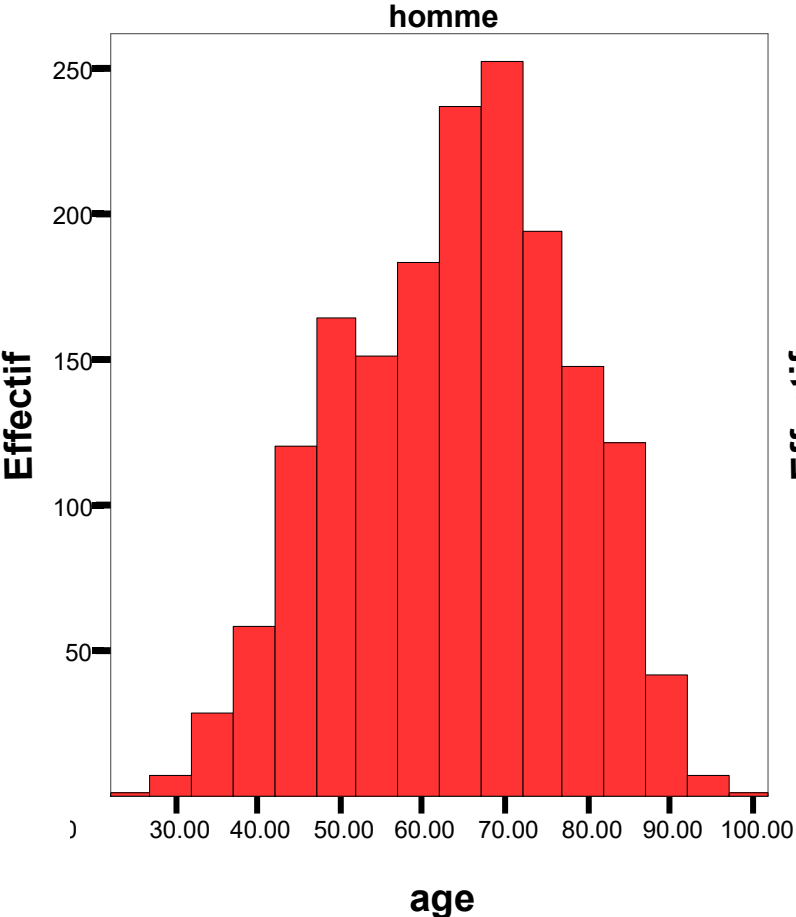
$Q_1$ ,  $Q_2$ ,  $Q_3$  sont les quartiles,  $Q_2$  la médiane.

# Age des patients hospitalisés entre septembre 1993 et janvier 1995 inclus pour infarctus du myocarde dans l'Isère, le Rhône et la Loire. Histogramme des effectifs

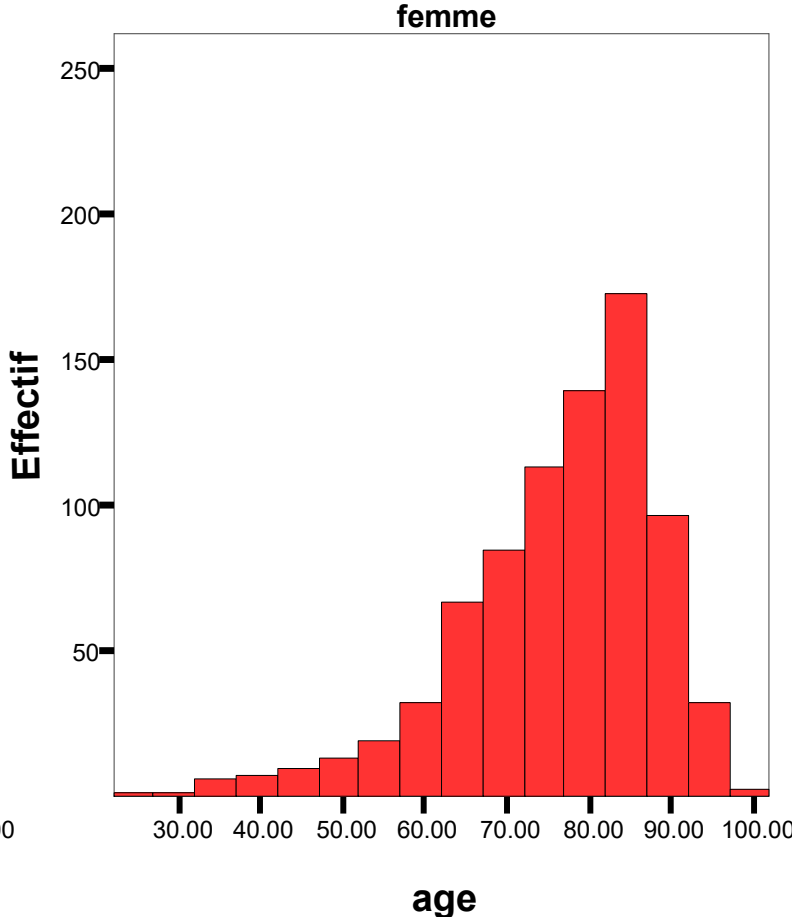


# Histogramme des effectifs

Moyenne = 64 ans

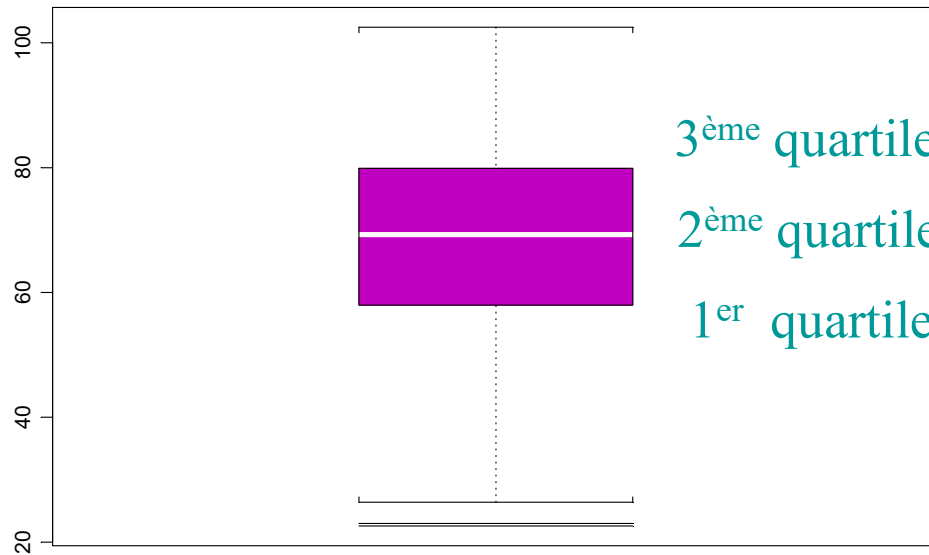


Moyenne = 76 ans



# Age des patients hospitalisés entre septembre 1993 et janvier 1995 inclus pour infarctus du myocarde dans l'Isère, le Rhône et la Loire \*

Age des 2517 patients de PRIMA



Distance interquartile

## LE BOX PLOT (Boîte à moustaches)

### Quantiles :

0%	25%	50%	75%	100%
22.6	58.0	69.3	79.9	102.5

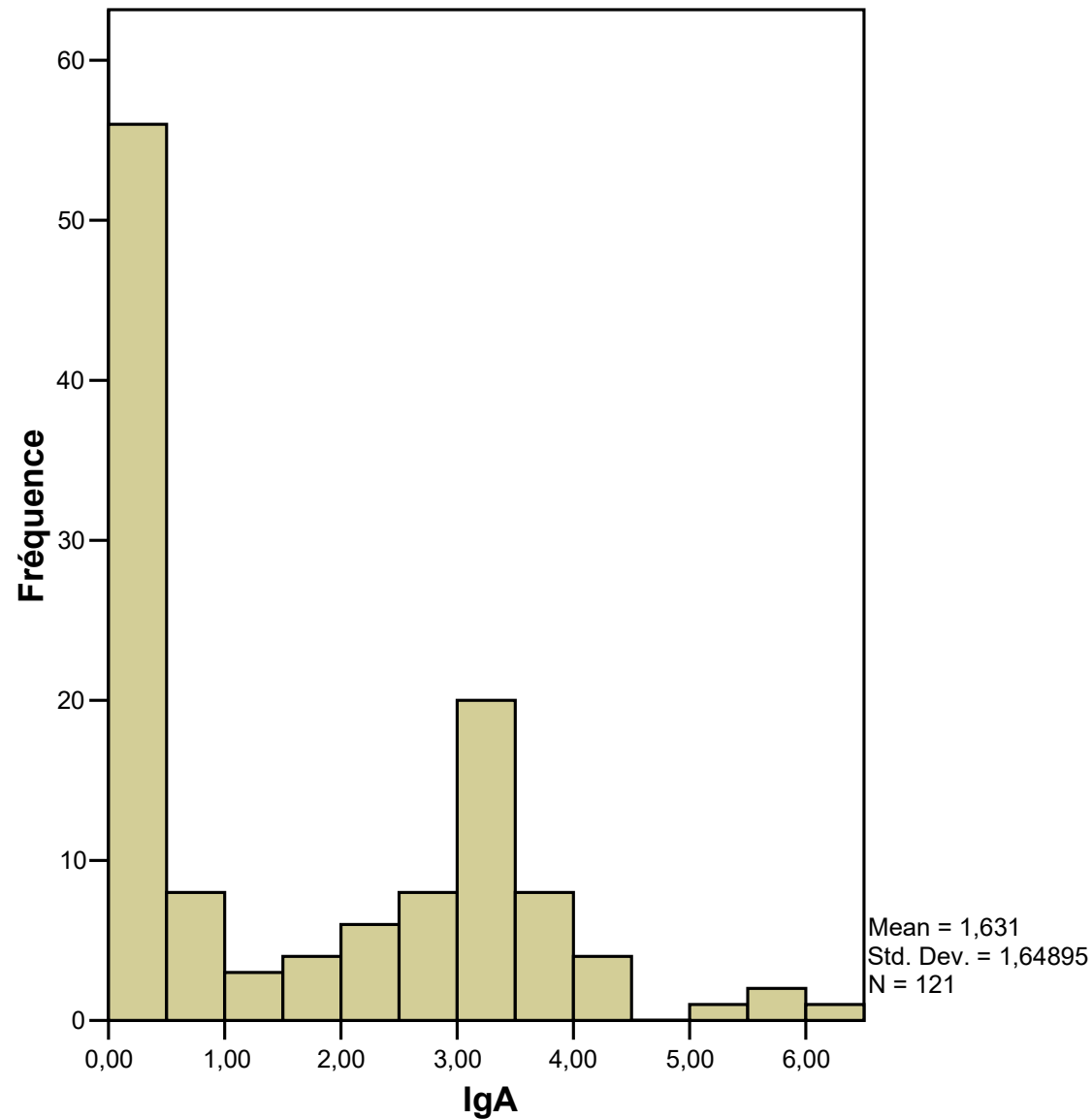
### Moyenne et déviation standard (ou : écart type)

Moyenne :	67.96 ans
Ecart type :	14.39 ans

\* quand, où, qui ?



# IgA à la naissance chez 124 enfants ayant une toxoplasmose congénitale



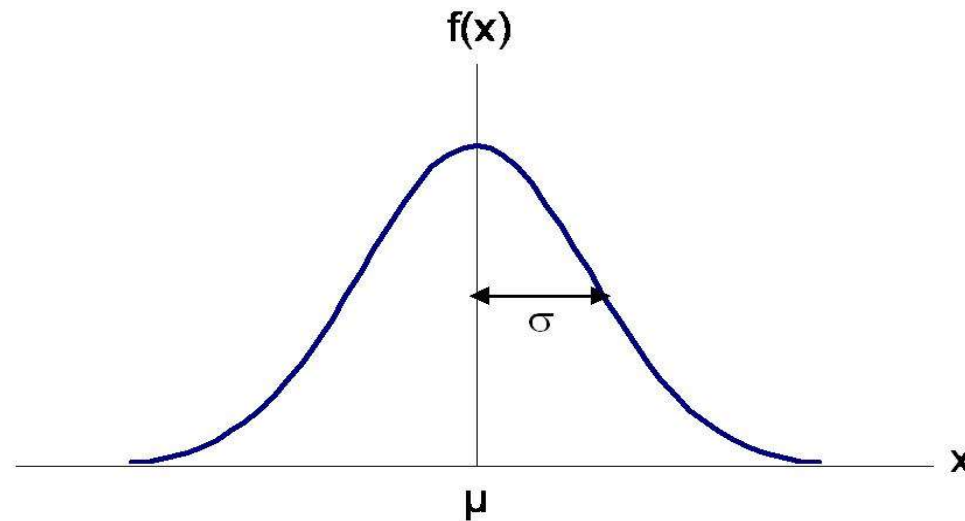
# La distribution de probabilité de la loi normale

- La distribution normale est entièrement décrite par sa moyenne  $\mu$  et sa variance  $\sigma^2$  ou son écart type  $\sigma$
- Sa densité de probabilité est définie par l'équation

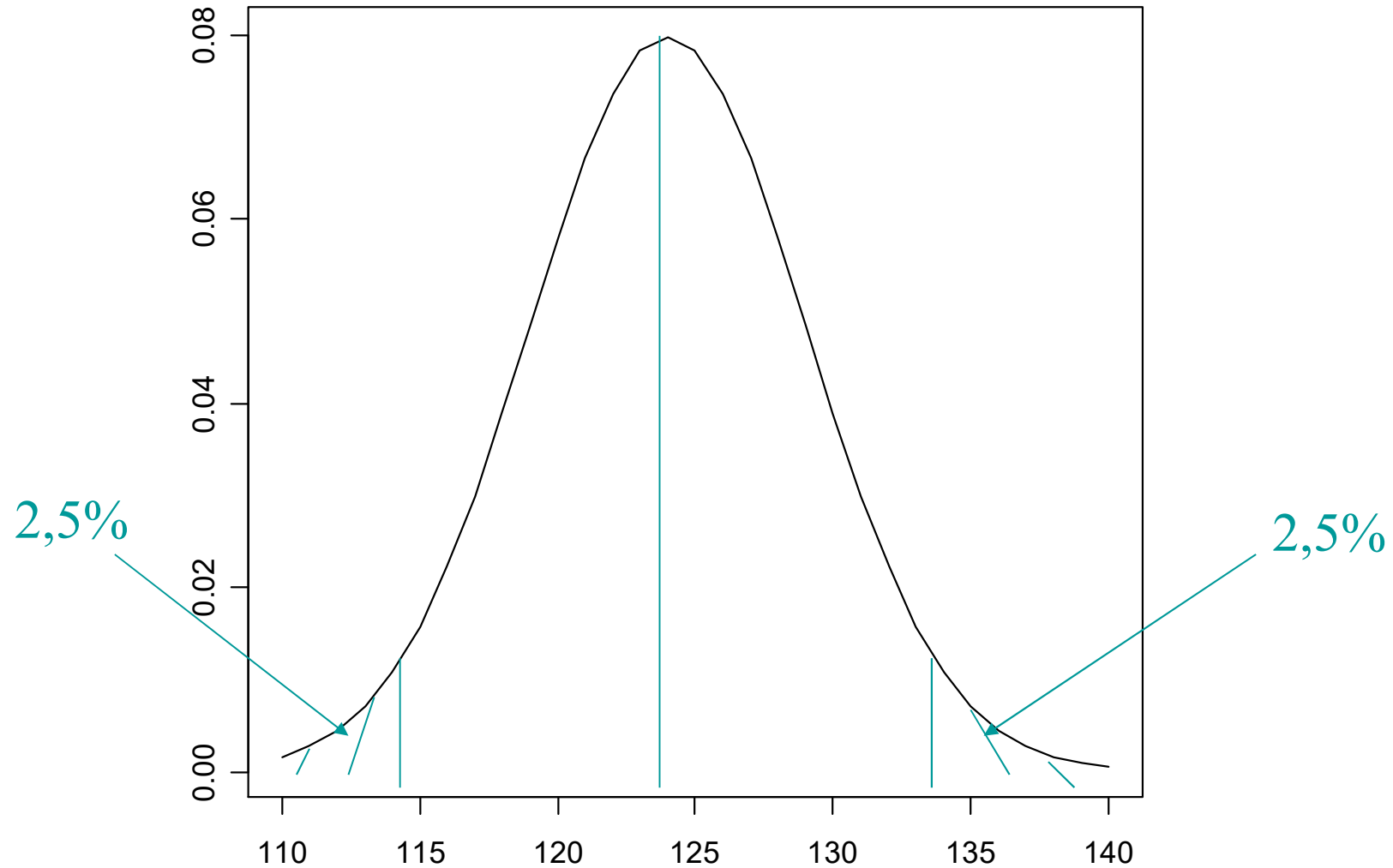
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

# Loi normale

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

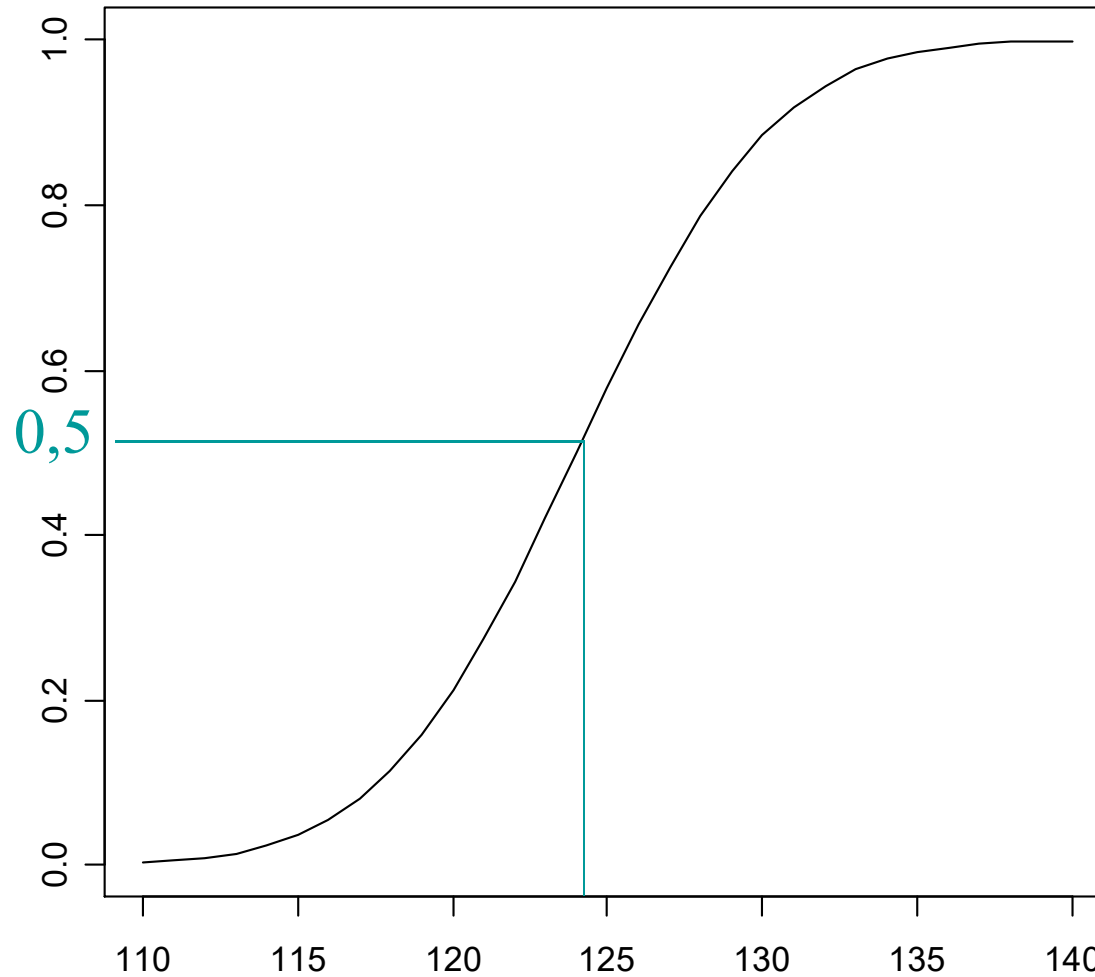


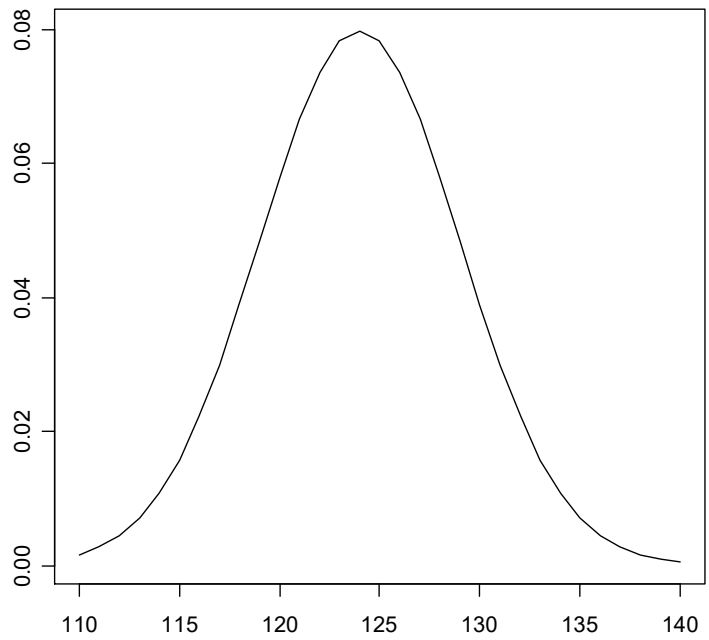
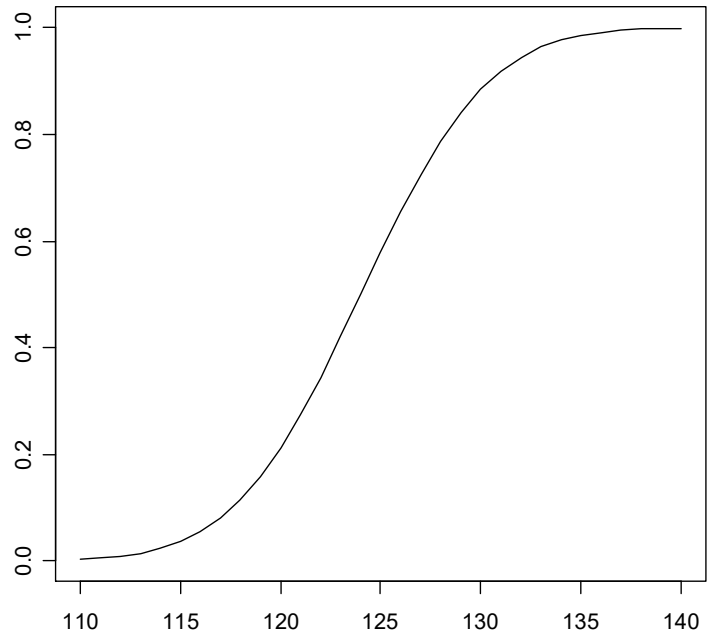
# Distribution normale de moyenne = 124 cm et d'écart type = 5 cm



30/08/2024

# Fonction de répartition de la distribution normale de moyenne = 124 cm et d'écart type = 5 cm





30/08/2024

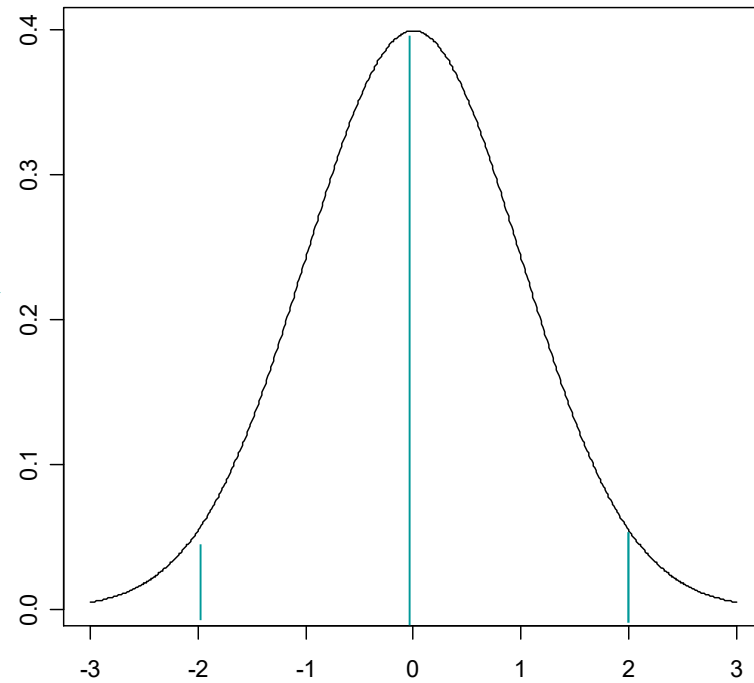
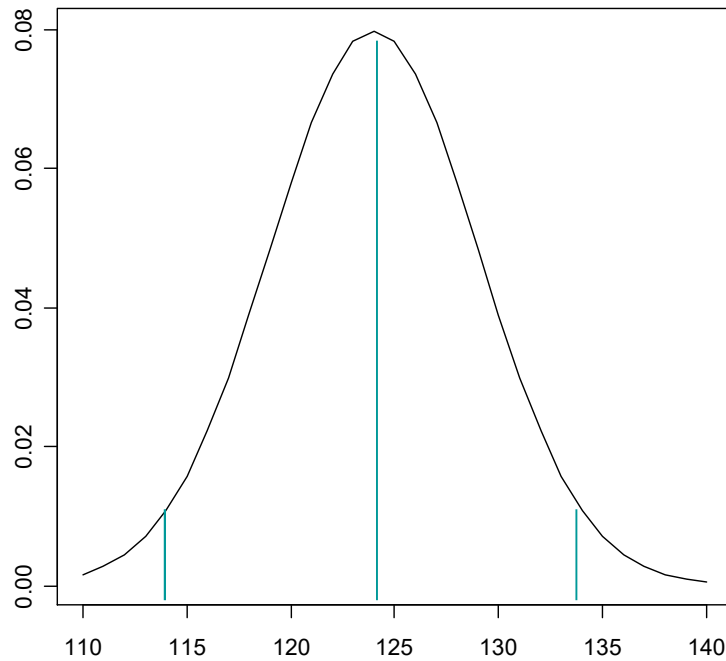
# Distribution normale centrée réduite

- Caractéristique étudiée centrée sur sa moyenne et divisée par son écart type = écart réduit

$$\varepsilon = \frac{x_i - \mu}{\sigma}$$

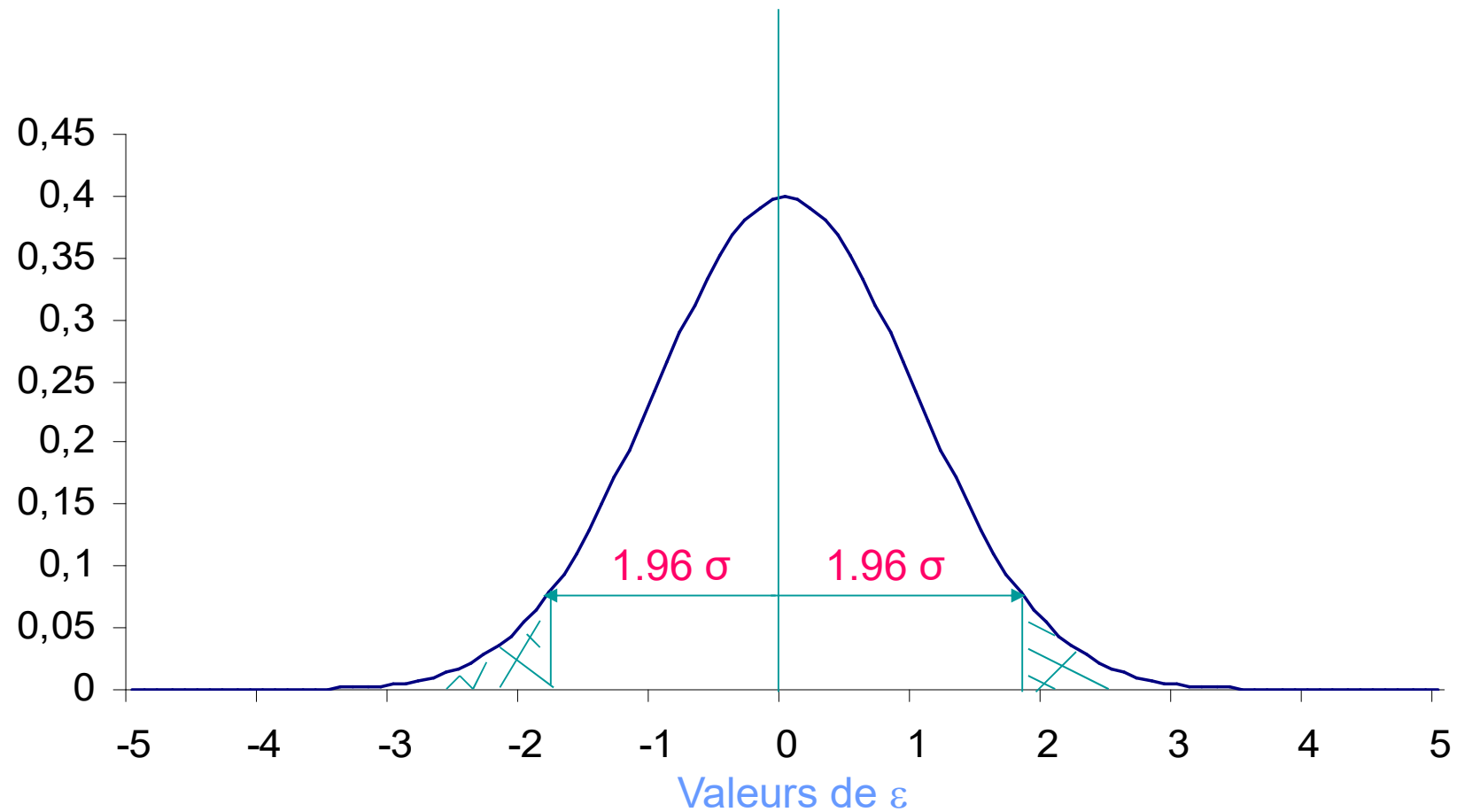
- Si  $X$  suit une distribution normale de moyenne  $\mu$  et d'écart-type  $\sigma$ ,
- alors l'écart réduit suit une distribution normale standard de moyenne 0 et d'écart type 1

# Distribution normale centrée réduite





# Loi normale centrée réduite



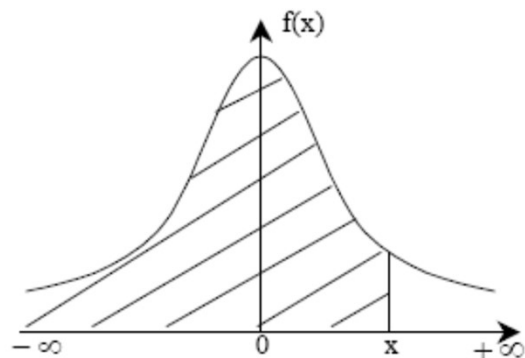
## La distribution normale

$$[\mu - \sigma; \mu + \sigma]$$

**68% des valeurs**

$$[\mu - 2\sigma; \mu + 2\sigma]$$

**95% des valeurs**



$$F(x) = P(X < x)$$

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

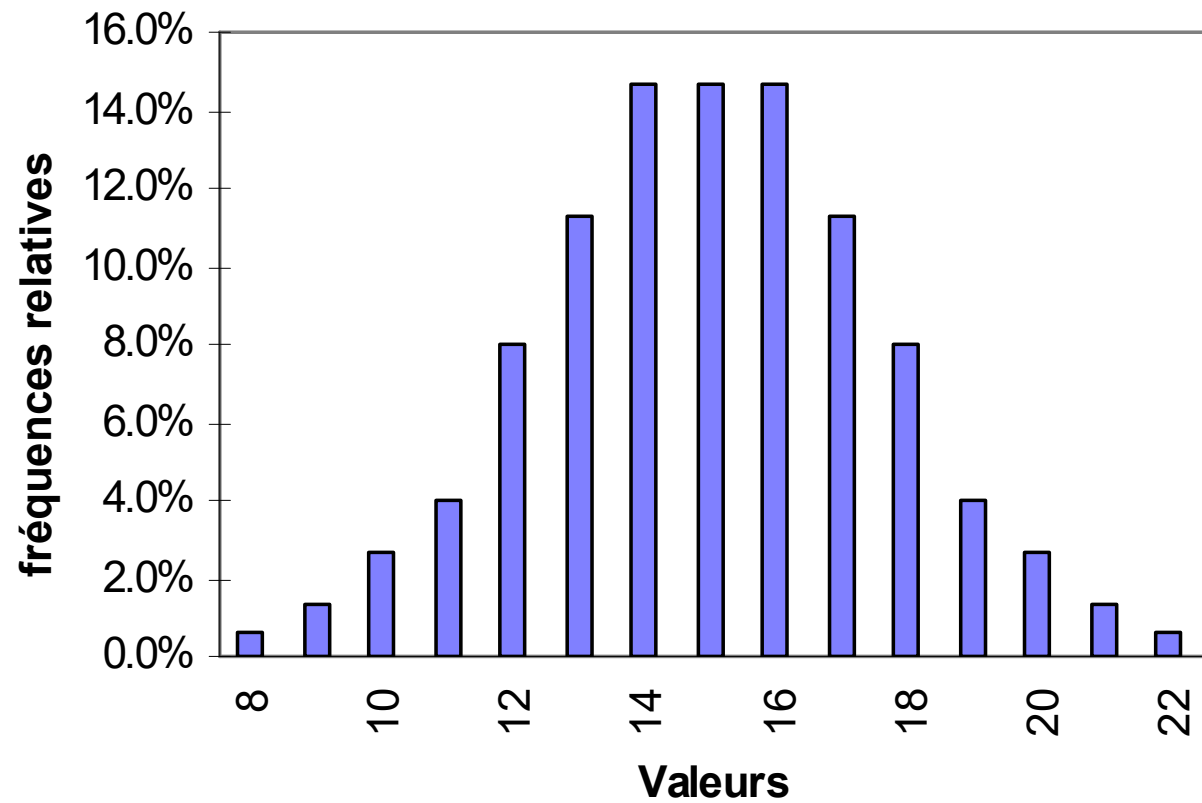
$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

X	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974

## ***Rappel distribution des durées d'incubation de la maladie A***

	<b>Valeurs</b>
<b>Moyenne</b>	<b>15 j</b>
<b>Ecart-type</b>	<b>2.62 j</b>
<b>Effectif</b>	<b>150</b>

## Distribution des valeurs de la variable étudiée



# Distribution à deux variables

## Deux variables qualitatives

## Distribution de fréquences à deux dimensions (fréquences absolues)

	Y	$y_1$	$y_2$	.....	$y_j$	.....	$y_J$	Totaux
X								
$X_1$		$n_{11}$	$n_{12}$	.....	$n_{1j}$	.....	$n_{1J}$	$n_{1.}$
$X_2$		$n_{21}$	$n_{22}$	.....	$n_{2j}$	.....	$n_{2J}$	$n_{2.}$
.								
$X_i$		$n_{i1}$	$n_{i2}$	.....	$n_{ij}$	.....	$n_{iJ}$	$n_{i.}$
.								
$X_I$		$n_{I1}$	$n_{I2}$	.....	$n_{Ij}$	.....	$n_{IJ}$	$n_{I.}$
Totaux		$n_{.1}$	$n_{.2}$	.....	$n_{.j}$	.....	$n_{.J}$	<b><math>n_{..}</math></b>

## Distribution de fréquences à deux dimensions (fréquences relatives)

	Y	$y_1$	$y_2$	.....	$y_j$	.....	$y_J$	Totaux
X								
$x_1$		$f_{11}$	$f_{12}$	.....	$f_{1j}$	.....	$f_{1J}$	$f_{1.}$
$x_2$		$f_{21}$	$f_{22}$	.....	$f_{2j}$	.....	$f_{2J}$	$f_{2.}$
.								
$x_i$		$f_{i1}$	$f_{i2}$	.....	$f_{ij}$	.....	$f_{iJ}$	$f_{i.}$
.								
$x_I$		$f_{I1}$	$f_{I2}$	.....	$f_{Ij}$	.....	$f_{IJ}$	$f_{I.}$
Totaux		$f_{.1}$	$f_{.2}$	.....	$f_{.j}$	.....	$f_{.J}$	1

$$f_{jk} = n_{jk} / n$$



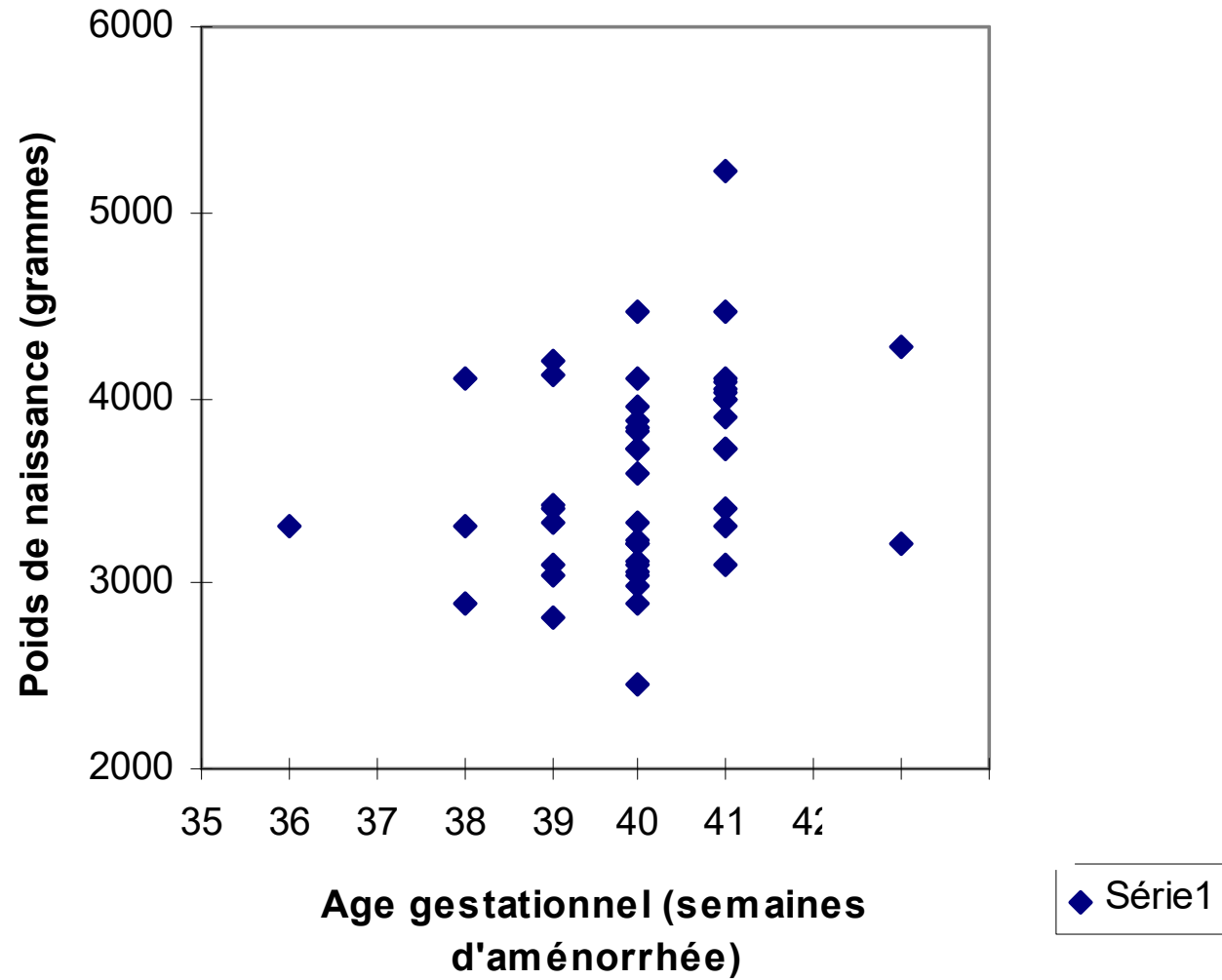
Evolution de la proportion des cancers de l'ampoule rectale réséqués  
de stade A ou B de Dukes en fonction du temps.

Registre Bourguignon des Cancers Digestifs.

Période de diagnostic	Stade (classification de Dukes)*		
	A ou B	Autres stades	Totaux
1976-80	35.8%	64.2%	100%
1981-85	44.6%	55.4%	100%
1986-90	52.5%	47.5%	100%
Totaux	43.8%	56.2%	100%

\*stade inconnu chez 7 individus.

### Age gestationnel et poids de naissance chez 48 enfants: diagramme de dispersion



# ***En résumé***

## **Présentation des données, tableaux et graphes**

- ***Tableau des effectifs et des proportions de la distribution***
- ***diagrammes - Histogrammes***

## **Résumés quantitatifs des données (données en nombre fini) :**

### ***- Paramètres de position :***

- \* Moyenne arithmétique
- \* Médiane
- \* Mode: Mo ou Valeur la plus fréquente

### ***- Paramètres de dispersion ou variabilité :***

- \* Extrêmes
- \* Variance
- \* Ecart-type