Université Claude Bernard Lyon 1

Tutorat Lyon Est

Unité d'Enseignement 3

BANQUE DE QCM

2013- 2022

STATISTIQUES DESCRIPTIVES
(Introduction)

QUESTIONS – REPONSES

2013-2021

Question 1 : Introduction aux probabilités :

Poids en Kg	Effectif
[45-50[	2
[50-55[	5
[55-60[	9
[60-65[	13
[65-70[	15
[70-75[	12
[75-80[	9
[80-85[	7
[85-90[	3

A. La variable « poids est une variable quantitative discrète.
B. La variable « poids en classes » est une variable quantitative discrète.
C. La variable « poids en classe » a pour mode la classe modale [65-70[.
D. Le troisième quartile de la distribution de la variable poids est dans la classe [70-75[.
E. Le troisième quartile de la distribution d'une variable est son 75ème percentile.

Question 1 : BCE

A FAUX Le poids est une variable quantitative continue.

B VRAI : C'est la conséquence du regroupement des valeurs des poids en classes. Contrairement à la variable « poids », la variable « poids en classe » ne peut prendre qu'un nombre fini de valeurs.

C VRAI : On remarque que toutes les classes ont la même amplitude. La classe [65-70[est celle qui contient le plus grand nombre d'étudiants, c'est donc la classe modale.

D FAUX : Pour trouver le troisième quartile on réalise le calcul suivant :

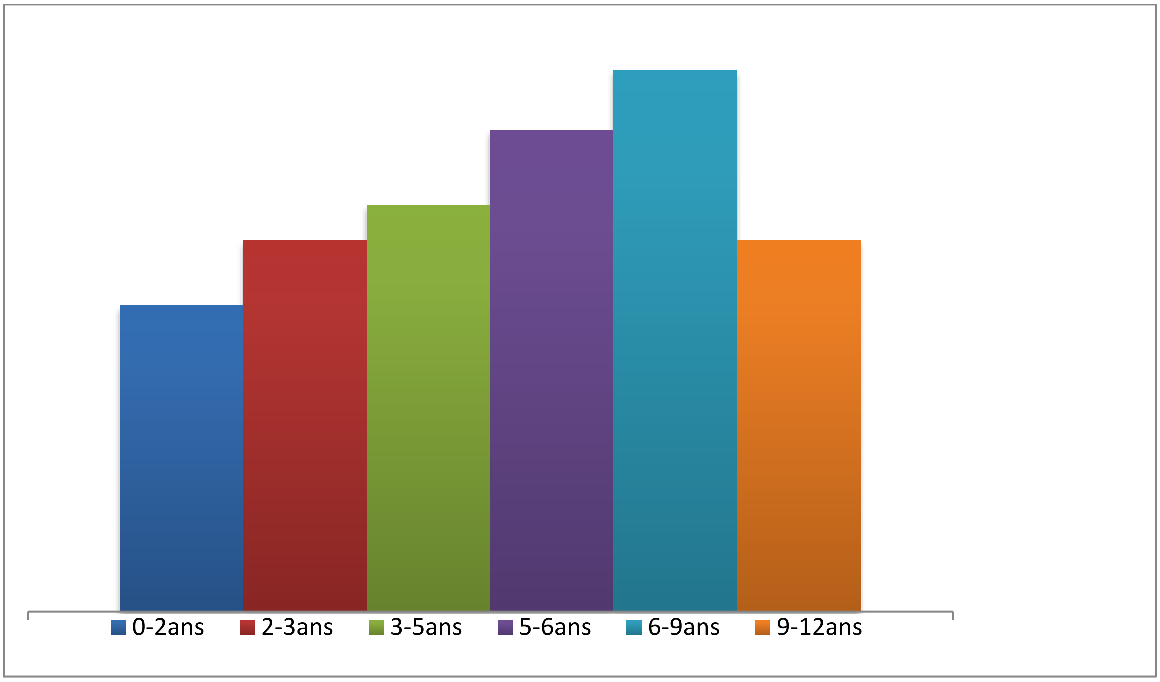
$$(75+1) \text{ X } \frac{3}{4} = 76 \text{ x } \frac{3}{4} = 57$$

Le troisième quartile de la distribution de la variable « poids » est la 57ème valeur des poids, la série des valeurs des poids ayant été préalablement triée. À partir des effectifs des « classes de poids », on observe que la 57ème valeur des poids est dans la classe [75-80[.

E VRAI : découle du calcul $\frac{3}{4}$ x $0{,}75$.

Question 2 :

On s'intéresse à la proportion d'enfants atteints de rougeole dans une ville en fonction de l'âge. On obtient le graphique suivant :



On utilisera : 2/3 = 0,667

age	0-2ans	2-3ans	3-5ans	5-6ans	6-9ans	9-12ans	total
%/an	A	X	B	9,6	10,8	7,4	
ni	84	47	108	64	C	D	E

Remplacer les valeurs du tableau en cochant les items justes :

A. 5,9
B. 8,1
C. 216
D. 184
E. 834

Question 2 : BC

On notera : n_i le nombre de cas dans la $i^{ème}$ classe d'âge.

d_i la densité de fréquence de la $i^{ème}$ classe d'âge

On commence par calculer le nombre total de sujets grâce à la classe des 5-6 ans :

$$N_{tot} = \frac{ni}{di \times amplitude} = \frac{64}{0,096 \times 1} = \frac{32 \times 2}{32 \times 0,003} = \frac{2}{0,003} = \frac{2}{3} \times 10^3 = 667$$

Dans la suite de la correction, nous utiliserons $N_{tot} = \frac{2}{3} \times 10^3$ pour simplifier les calculs ; les résultats obtenus sont les mêmes qu'en utilisant 667 au dixième prêt.

A FAUX $d_i = \frac{ni}{Ntot \times amplitude} = \frac{84}{\frac{2}{3} \times 1000 \times 2} = \frac{84 \times 3}{4 \times 1000} = \frac{21 \times 3}{1000} = 0,063 = 6,3\%/an$

B VRAI $d_i = \frac{ni}{Ntot \ x \ amplitude} = \frac{108}{\frac{2}{3} x \ 1000 \ x \ 2} = \frac{108 \ x \ 3}{4 \ x \ 1000} = \frac{27 \ x \ 3}{1000} = \frac{81}{1000}$ = 0,081 = 8,1%/an

C VRAI $n_i = d_i$ x n_{tot} x amplitude = 0,108 x $\frac{2}{3}$ x 1000 x 3 = 108 x 2 = 216

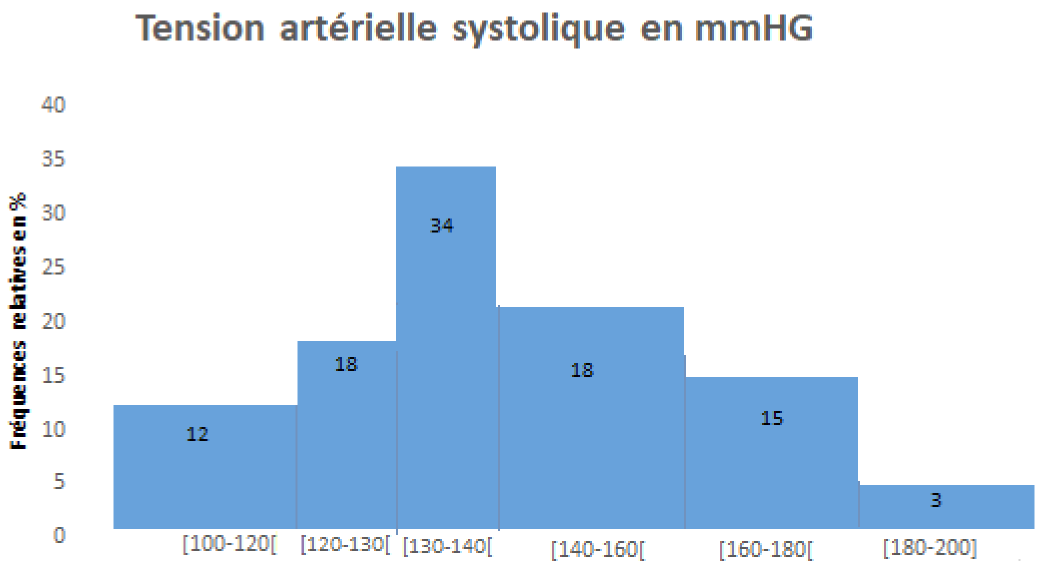
D FAUX $n_i = d_i$ x n_{tot} x amplitude = 0,074 x $\frac{2}{3}$ x 1000 x 3 = 74 x 2 = 148

E FAUX N_{tot} = 667

age	0-2ans	2-3ans	3-5ans	5-6ans	6-9ans	9-12ans	total
%/an	6,3	7	8,1	9,6	10,8	7,4	
n_i	84	47	108	64	216	148	667

Question 3 :

On s'intéresse à la tension artérielle systolique (mmHg) dans un échantillon de N personnes âgées traitées pour des problèmes d'hypertension.
On sait que parmi eux, 90 personnes ont une tension comprise entre 120 et 130 mmHg.
Les résultats obtenus sont présentés sous la forme de l'histogramme ci-dessous :



A. Le nombre de personnes ayant une tension entre 160 et 180 mm Hg est n160-180=15.
B. La densité de fréquence de la classe 120-130 est identique à celle de la classe 140-160.
C. Le nombre total de personnes interrogées est égal à 500.
D. L'histogramme ci-dessus présente les fréquences relatives obtenues à partir de l'échantillon avec des amplitudes égales pour chaque classe.
E. La médiane de cet échantillon vaut environ 150 = (100+200) /2, de plus on peut dire que 18% des personnes de l'échantillon ont une tension comprise dans la même classe que celle contenant la médiane.

Question 3 : C

A FAUX 15% correspond à une fréquence de 0,15 car comme indiqué dans la légende de l'histogramme, c'est la fréquence relative qui nous ai donné. De plus $fi = \frac{ni}{Ni}$ soit $n_i = f_i$ x N

Or on ne connaît pas N, mais on sait que 90 personnes ont une tension entre 120 et 130 et que la fréquence correspondante est de 0,18.

Ainsi N = $\frac{n_{120-130}}{f_{120-130}} = \frac{90}{0{,}18} = \frac{90 \; x \; 100}{18}$ =5 x 100 = 500 donc on sait déjà que C VRAI

n_i= f_ix N= 0,18 x 500 = 90 donc il n'y a pas 15 personnes ayant une tension artérielle systolique comprise entre 120 et 130 mm Hg mais 90.

B FAUX **Attention, la densité de fréquence n'est pas la même chose que la fréquence !** En effet même si les fréquences des classes 120-130 et 140-160 sont identiques, leur densité de fréquence sont différentes car elles n'ont pas la même amplitude

$A_{120-130}$= 10 et $A_{140-160}$=20

De plus $d_i = \frac{f_i}{A_i}$ donc $d_{120-130} = \frac{0{,}18}{10}$ = 0,018 et $d_{140-160} = \frac{0{,}18}{20}$ = 0,009

C VRAI voir question A

D FAUX La définition donnée de l'histogramme présenté dans cet exercice est bien juste mais les amplitudes des classes sont différentes comme détaillé précédemment.

E FAUX Cela n'a aucun sens, on ne cherche pas la médiane en partant des valeurs que peuvent prendre les variables mais en partant des fréquences prises par ces dernières.

En effet, **la médiane sépare un échantillon en deux groupes de même effectif,** donc elle correspond à la valeur qui classe 50% des valeurs inférieures à la médiane d'un côté et 50% des valeurs supérieures à la médiane de l'autre.

Ici nous avons des classes donc on ne peut pas trouver la valeur précise de la médiane, cependant, grâce aux barres on voit que : 12% + 18% = 30% + 34% = 64%, on dépasse les 50% une fois qu'on ajoute la troisième classe (130-140), **on comprend alors que la médiane qui sépare le groupe en 2 parties égales, se trouve dans cette classe et qu'elle donc comprise entre 130 et 140 mm Hg.**

Question 4 :

On étudie la répartition des naissances (en SA = Semaine d'Aménorrhée) chez les bébés prématurés, dans un échantillon de N bébés.

On sait que 18 bébés sont nés entre la 26^{e} et la 27^{e} SA.

Les résultats de l'étude sont présentés sous la forme de l'histogramme ci-dessous :

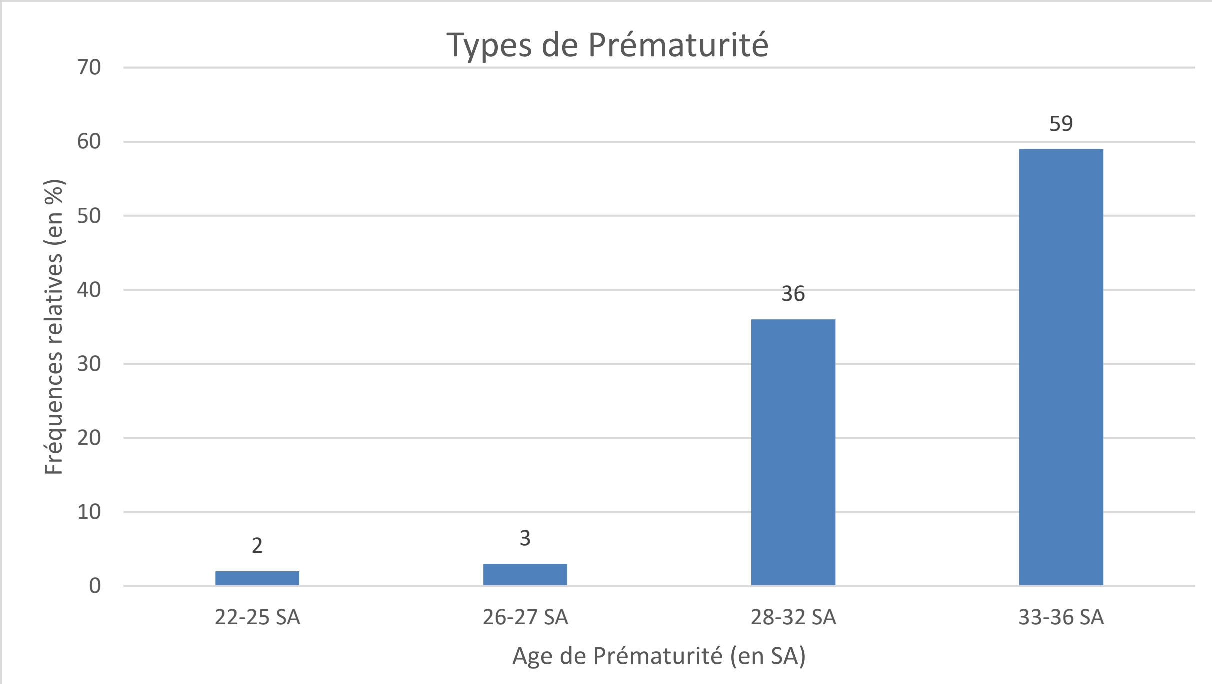


Figure 1 : Répartition des prématurés par degré de prématurité

A. L'échantillon compte N=600 bébés : on parle d'effectif total.
B. Il y a n=36 bébés nés entre 28 et 32 SA.
C. 50% des bébés prématurés naissent avant la 32e SA (la 32e SA étant incluse).
D. La densité de fréquence est identique entre les classes [22-25]SA et [33-36]SA.
E. La classe avec la plus grande densité de fréquence est la classe [33-36]SA.

Question 4 : AE

METHODE : Compréhension du graphique

L'axe des ordonnées donne des **fréquences relatives**, c'est-à-dire des proportions : $f_i = \frac{n_i}{N}$ avec f_i , la fréquence ; n_i, l'effectif d'une catégorie et N, l'effectif total.
Ainsi, les chiffres situés au-dessus des barres de l'histogramme sont des proportions. On a donc :
- 2% des bébés prématurés naissent entre 22 et 25 SA incluse
- 3% des bébés prématurés naissent entre 26 et 27 SA incluse
- 36% des bébés prématurés naissent entre 28 et 32 SA incluse
- 59% des bébés prématurés naissent entre 33 et 36 SA incluse

A VRAI On donne dans l'énoncé « 18 bébés sont nés entre la 26e et la 27e SA » cela revient à dire que : 18 bébés représentent 3% des bébés prématurés soit que $n_{[26-27]} = 18$ et $f_{[26-27]} = \frac{3}{100}$

Par produit en croix on obtient :

$$f_i = \frac{n_i}{N} \Leftrightarrow N = \frac{n_i}{f_i} \Leftrightarrow N = \frac{18}{\frac{3}{100}} = 18.\frac{100}{3} = 600$$

Autre méthode (cf item B) : calculer les n_i pour chaque classes et les additionner (en effet, la somme des fréquences relatives = 2% + 3% + 36% + 59% = 100% donc la somme des n_i de chaque classe donnera l'effectif total N).

B FAUX ATTENTION ne pas confondre fréquence et effectif. Le chiffre 36 signifie « 36% des bébés prématurés naissent entre 28 et 32 SA incluse », donc c'est une fréquence.

On cherche combien de bébés sont nées entre 28 et 32 SA soit l'effectif correspondant à 36%.
On sait que 18 bébés représentent 3% des bébés prématurés
Soit que $n_{[26-27]} = 18$ et $f_{[26-27]} = \frac{3}{100}$
Et $n_{[28-32]}$ =? et $f_{[28-32]} = \frac{36}{100}$
Par produit en croix on obtient $n_{[28-32]} = \frac{36}{100}.18.\frac{100}{3} = 36.6 = 216$

Ainsi, 216 bébés dans l'échantillon sont nés entre 28 et 32 SA (32 incluse).

NB. On peut refaire ce calcul pour toutes les classes et vérifier ainsi que N l'effectif total vaut 600.
$n_{[22-25]} = \frac{2}{100}.18.\frac{100}{3} = 6.2 = 12$
$n_{[26-27]} = 18$
$n_{[28-32]} = \frac{36}{100}.18.\frac{100}{3} = 36.6 = 216$
$n_{[33-36]} = \frac{59}{100}.18.\frac{100}{3} = 59.6 = 354$
N= 12 + 18 + 216 + 354 = 30 + 570 = 600

C FAUX Ici on s'intéresse à la médiane (50% des bébés). **La médiane est une valeur qui coupe un échantillon en 2 groupes de même effectif**. Donc, 50% des valeurs de l'échantillon se situent au-dessus de cette médiane et 50% en-dessous.

Si on additionne les fréquences relatives jusqu'à la classes [28-32]SA incluse, on atteint : 2% + 3% + 36% = 41%
Pour arriver aux 50%, il est nécessaire de prendre en compte les valeurs de la classe [33-36]SA.
Donc, moins de 50% des bébés prématurés naissent avant 32SA (32SA incluse).

D FAUX J'ai mis cet item pour que vous fassiez la distinction entre fréquence et amplitude.
La densité de fréquence se définit par la formule $d_i = \frac{f_i}{A}$ avec d_i, la densité ; f_i, la fréquence et A, l'amplitude.

ATTENTION ici les classes [22-25] et [33-36] ont la même amplitude. Or leur fréquence est différente (respectivement 2% et 59% donc leurs densités seront forcément différentes).

Par le calcul ça donne :

$$d_{[22-25]} = \frac{\frac{2}{100}}{3} = 0{,}67.10^{-2}$$
$$d_{[33-36]} = \frac{\frac{59}{100}}{3} \approx \frac{\frac{60}{100}}{3} = 0{,}2$$

E VRAI Pour répondre à cet item il faut calculer toutes les densités de fréquence de toutes les classes (pour que vous manipuliez la formule ;-))

$$d_{[22-25]} = \frac{\frac{2}{100}}{3} = 0{,}67.10^{-2}$$
$$d_{[26-27]} = \frac{\frac{3}{100}}{1} = 0{,}03$$
$$d_{[28-32]} = \frac{\frac{36}{100}}{4} = \frac{36}{400} = 0{,}09$$

$$d_{[33-36]} = \frac{\frac{59}{100}}{3} \approx \frac{\frac{60}{100}}{3} = 0{,}2$$
Donc la classe avec la plus grande densité de fréquence est la classe [33-36]SA.

Question 5 – L'intro de l'UE3, Késako ?? :

On étudie la taille des étudiants en P1 de Lyon Est.
Pour cela, on étudie un échantillon de 400 P1.
Les résultats de l'étude sont présentés sous la forme du tableau ci-dessous :

Taille (en cm)	Fréquence cumulée (en %)
[140-150[	8
[150-160[	30
[160-170[	64
[170-180[	96
[180-195[	100

On trace ensuite la fonction de répartition empirique de la taille de P1 de Lyon-Est, en utilisant la taille centrale de chaque classe, comme taille moyenne.

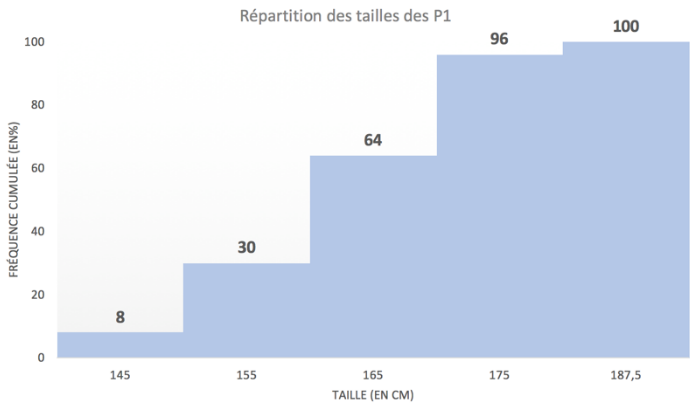


Figure : Graphique de la taille des P1 de Lyon-Est

A. 64% des P1 mesurent entre 160 et 170 cm (170cm exclus).
B. La proportion de P1 mesurant entre 180 et 195 cm (195cm exclus) est plus importante que celle de P1 mesurant entre 135 et 150 cm (150 cm exclus).
C. 88 P1 mesurent entre 150 et 160 cm (160 exclus)
D. 352 P1 mesurent entre 150 et 180 cm (180 exclus)

Pour l'item E, on s'intéresse aux deux histogrammes des densités de fréquences ci-dessous, arrondi au dixième. Un seul des deux est correct .

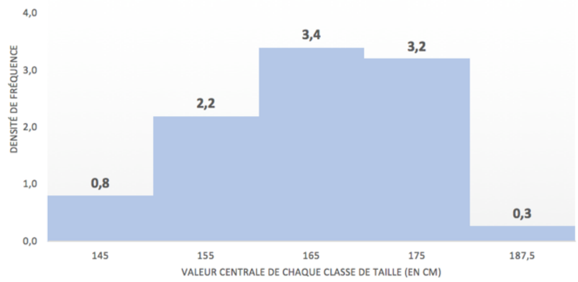


Figure a

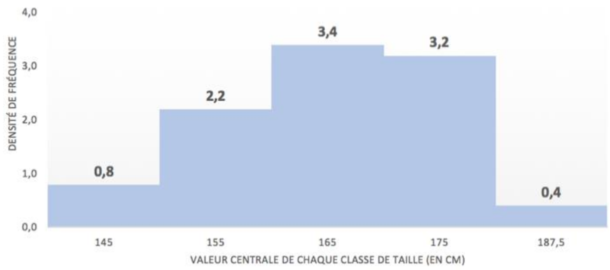


Figure b

E. La **figure a** est la figure qui représente correctement la densité de fréquence pour chacune des classes.

Question 5 : CDE

ATTENTION : ce graphique présente des fréquences cumulées !!!! C'est-à-dire qu'à la proportion de chaque classe, on ajoute la proportion de la classe précédente. On CUMULE les proportions.
C'est-à-dire que, comme le graphique présente des fréquences cumulées, il faut calculer les proportions pour chaque classe, en retranchant à chaque fois la valeur de la classe précédente.

A FAUX $f_{c[140-150[} = 0{,}08$ avec f_c la fréquence cumulée. Or il n'y a pas de valeurs avant l'intervalle [140-150[donc $f_{[140-150[} = 0{,}08$
Donc 8% des P1 mesurent entre 140 et 150 cm (150 exclus)

$f_{c[150-160[} = f_{[140-150[} + f_{[150-160[}$
soit $f_{[150-160[} = f_{c[150-160[} - f_{[140-150[} = 0{,}3 - 0{,}08 = 0{,}22$
Donc 22% des P1 mesurent entre 150 et 160 cm (160 exclus)

Ainsi de suite pour toutes les fréquences. On obtient donc :
$f_{c[160-170[} = f_{[140-150[} + f_{[150-160[} + f_{[160-170[}$
soit $f_{[160-170[} = f_{c[160-170[} - f_{[150-160[} - f_{[140-150[} = 0{,}64 - 0{,}22 - 0{,}08 = 0{,}34$

Donc 34% des P1 mesurent entre 160 et 170 cm (170 exclus)

$f_{c[170-180[} = f_{[140-150[} + f_{[150-160[} + f_{[160-170[} + f_{[170-180[}$
soit $f_{[170-180[} = f_{c[170-180[} - f_{[160-170[} - f_{[150-160[} - f_{[140-150[} = 0{,}96 - 0{,}34 - 0{,}22 - 0{,}08 = 0{,}32$

Donc 32% des P1 mesurent entre 170 et 180 cm (180 exclus)

$f_{c[180-195[} = f_{[140-150[} + f_{[150-160[} + f_{[160-170[} + f_{[170-180[} + f_{[180-195[}$
soit $f_{[180-195[} = f_{c[180-195[} - f_{[170-180[} - f_{[160-170[} - f_{[150-160[} - f_{[140-150[} = 1 - 0{,}32 - 0{,}34 - 0{,}22 - 0{,}08 = 0{,}04$

Donc 4% des P1 mesurent entre 180 et 195 cm (195 exclus)
On vérifie que le total fasse bien 1 (soit 100%) : 0,08+0,22+0,34+0,32+0,04 = 1.

Ainsi, 64% au-dessus de la classe [160-170[signifie que 64% des P1 mesurent 170cm ou moins, 170cm exclus. Pour connaître le nombre de P1 mesurant entre 160 et 170 cm (170 cm exclus), il faut aux 64% retrancher les proportions des classes précédentes.
On a ainsi $n_{[160-170[} = 64\% - 30\% = 34\%$
Donc 34% des P1 mesurent entre 160 et 170 cm.
NB : on ne retranche pas le 8% car il est déjà inclus dans le 30% (le 30% étant le cumul des classes [140-150[et [150-160[).

B FAUX Le piège ici c'est d'oublier que les fréquences sont cumulées et de croire que 100% correspond à la proportion de P1 mesurant entre 180 et 195 cm (195 cm exclu). **CECI EST FAUX.**
100 % correspond au nombre de P1 mesurant 195 cm ou moins (195 exclus).
D'après la correction de l'item A on a :

4% des P1 mesurent entre 180 et 195 cm (195cm exclus)
8% des P1 mesurent entre 140 et 150 cm (150cm exclus)

On conclut qu'il y a plus de P1 qui mesurent entre 140 et 150 cm (150cm exclus) que de P1 mesurent entre 180 et 195 cm (195cm exclus) dans notre échantillon.

C VRAI D'après la correction de l'item A, on connaît la proportion de P1 mesurant entre 150 et 160 cm (160 cm exclus) : 22%
Comme l'effectif total est, d'après l'énoncé 400, on utilise la formule suivante :

$f_i = \frac{n_i}{N}$ avec $N = 400\ et\ f_i = 22\%$
Donc : $n_i = f_i.N = 0{,}22.400 = 88$

D VRAI On cherche le nombre de P1 mesurant entre 150 et 180 cm (180 exclus).
2 méthodes :

- D'après les proportions de chaque classe trouvées en A

22% des P1 mesurent entre 150 et 160 cm (160 exclus)
34% des P1 mesurent entre 160 et 170 cm (170 exclus)
32% des P1 mesurent entre 170 et 180 cm (180 exclus)

TOTAL : 22%+34%+32% = 88% des P1 mesurent entre 150 et 180 cm (180 exclus)
Calculons l'effectif :

$$n_i = f_i.N = 0{,}88.400 = 352$$

- D'après le graphique

On lit sur le graphique au-dessus de la classe [170-180[la valeur 96%. Cela signifie que 96% des P1 mesurent moins de 180 cm (180 cm exclus)

On lit sur le graphique au-dessus de la classe [135-150[la valeur 8%. Cela signifie que 8% des P1 mesurent moins de 150 cm (150 cm exclus).

Donc :
$f_{c[170-180[} - f_{c[135-150[} = f_{[150-180[}$
$f_{c[170-180[} - f_{c[135-150[} = 0.96 - 0{,}08 = 0{,}88$

De même que précédemment, on calcule l'effectif :

$$n_i = f_i.N = 0{,}88.400 = 352$$

E VRAI La seule différence entre les deux graphiques correspond à la densité de fréquence pour la classe [180-195[. D'après la formule, on a $d_i = \frac{f_i}{A}$ avec di la densité, fi la fréquence et A l'amplitude.
Donc pour notre classe [180-195[, cela donne : $d_i = \frac{f_i}{A} = \frac{4}{15} = 0{,}2667 = 0{,}3$ par arrondi au dixième.
Ainsi la figure a est correct. (la b correspond à une amplitude de 10 comme pour les autres classes).

Question 6 – « Oeufsreusement » on a tout nettoyé :

Vos tuteurs, à qui les soirées avaient manqué durant leur difficile première année, ont commencé à bien profiter. Lors d'un after trop arrosé, les voisins n'ont pas trop apprécié la musicalité de la soirée (ou nos cris ?) et se sont mis à jeter des œufs sur le balcon de Bob le chondrocyte (notre star à tous). Nous allons ici étudier le nombre d'œufs envoyés par les 15 voisins de la résidence.

Indication : vous arrondirez vos calculs à la première décimale.

Nombre d'œufs envoyés	0	1	2	3	4	5
Nombre de voisins ayant envoyé ce nombre	5	1	3	1	3	2

A. La variable étudiée « nombre d'œufs lancés et reçus sur le balcon » est une variable qualitative nominale.
B. La médiane de cette série de valeurs est de 2 œufs lancés.
C. Le mode de cette série est de 0 œufs lancés.
D. En moyenne, chaque voisin a lancé 1 œuf.

E. La variance de cette série est de 3,7 œufs.

Question 6 – « Oeufsreusement » on a tout nettoyé : BC

A FAUX La variable étudiée est une variable **quantitative**, et on peut même ajouter que c'est une variable **discrète** car, sur un intervalle de valeurs à bornes finies, cette variable prend un nombre fini de valeurs (on peut lancer 1 œuf, 2 œufs, mais pas 1,65 œuf par exemple).

B VRAI La médiane partage la série en deux parties de même effectif. Ici, l'effectif n=15 ce qui est impair. Ainsi, la valeur qu'on cherche est la (15+1)/2=8^e valeur. [La formule théorique est (n+1)/2]. Si on compte grâce au tableau on voit que la 8^e valeur correspond à 2 œufs lancés.
Pour trouver qui est la 8^e valeur, on peut faire un **tableau de fréquences cumulées** (c'est normalement ce qu'il se passe un peu dans votre tête aussi) :

Nombre d'œufs envoyés	0	1	2	3	4	5
Fréquence cumulée du nombre de voisins ayant envoyé ce nombre	5	6	9	10	13	15

C VRAI Le mode est la valeur observée avec la plus grande fréquence. En l'occurrence, 6 voisins n'ont pas lancé d'œufs donc le mode est bien de 0.

D FAUX Ici, on nous demande simplement de calculer une moyenne. Pour ça, on divise le nombre d'œufs envoyés au total par le nombre de voisins. Ainsi, on a :

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

Avec $\sum x_i$ la somme des différentes valeurs de X et n le nombre total.

Moyenne = (0 x 5 + 1 x 1 + 2 x 2 + 3 x 1 + 4 x 3 + 5 x 2) / 15 = 32 / 15 = 2,1333… = 2
Je ne vous ai pas mis d'aide au calcul car si vous pensiez bien à multiplier par le nombre de voisins qui ont envoyé le nombre d'œufs vous vous rendiez vite compte que le résultat était supérieur à celui proposé. Par contre l'approximation était importante pour la suite.

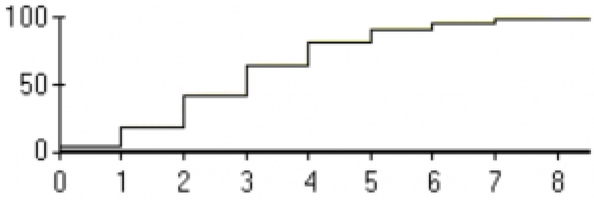
E FAUX La variance est la moyenne des carrés des écarts à la moyenne (oui ça fait peur) mais autrement dit c'est :

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

Var = [(0-2)²x 5 + (1-2)² x 1 + (2-2)² x 2 + (3-2)² x 1 + (4-2)² x 3 + (5-2)² x 2] / 15
= (20 + 1 + 0 + 1 + 12 + 18)/15
= 52 / 15
≈ 3,5

Question 7 – L'instant détente :

Diagramme 1 :



A. Le diagramme 1 est appelé un histogramme des fréquences cumulées.
B. L'aspect en forme d'escalier de la courbe du diagramme 1 est dû au fait que nous étudions une variable aléatoire continue.
C. La médiane est un paramètre de dispersion.
D. L'écart-type correspond au carré de la variance.
E. Le troisième quartile est la valeur en dessous de laquelle se trouvent 75% des valeurs de la série.

Question 7 – L'instant détente : AE

A VRAI C'est bien comme ça qu'on l'appelle, assez facile à reconnaitre, avec en ordonnée soit 100% au maximum soit 1, et en abscisse on a les valeurs de notre série.

B FAUX Il a ici une forme d'escalier ce qui signifie que l'on étudie une variable aléatoire discrète (il n'existe pas de valeur qui soit à 2,5 par exemple, on a donc un nombre fini de valeurs dans notre liste).

C FAUX La médiane est un paramètre de position : c'est le milieu de la série.

D FAUX L'écart-type (paramètre de dispersion) est la racine carrée de la variance. Par contre, la variance est le carré de l'écart type. En bref :

$$Var = \sigma^2 \text{ qui est équivalent à } \sigma = \sqrt{Var}$$

E VRAI C'est en effet la définition du 3è quartile. Pour le premier quartile, il faut 25% des valeurs en dessous. Le second quartile correspond à la médiane.

Question 8– Intro :

A. On peut calculer la moyenne d'une variable aléatoire qualitative ordinale.
B. Le sexe (féminin ou masculin) pourrait être décrit par une variable aléatoire qualitative ordinale.
C. Une variable quantitative est forcément discrète.
D. Le stade d'une maladie sera décrit par une variable quantitative.
E. Le groupe sanguin sera décrit par une variable qualitative nominale.

Question 8– Intro : E

A FAUX Une **VA qualitative ordinale** induit une **notion d'ordre** : par exemple, on pourra décrire l'intensité d'une douleur. Ainsi, nous ne pouvons pas faire une moyenne, par contre, on peut quand même décrire quel niveau d'intensité de douleur est choisi par le plus de personnes.

B FAUX Le sexe ne donne pas de notion d'ordre : ainsi, on le décrira plutôt par une VA qualitative nominale.

C FAUX Une VA quantitative peut être discrète (nombre de fois où l'on a mangé dans une journée) mais aussi continue (taux d'hémoglobine dans le sang).

D FAUX Le stade d'une maladie n'est pas une quantité : il sera donc plutôt décrit par une variable qualitative (ordinale car il y a une notion d'ordre).

E VRAI Le groupe sanguin n'est pas une quantité et n'induit pas de notion d'ordre donc il sera décrit par une VA qualitative nominale.

Question 9 –Introduction : avez-vous lu le cours ? :

A. La moyenne est un paramètre de position.
B. L'unité statistique peut aussi être appelée individu.
C. L'inférence statistique permet de tirer des conclusions générales à partir de cas particuliers.
D. Le mode est un paramètre de position.
E. Le coefficient de variation vaut : $\frac{\sigma}{\bar{x}}$

Question 9 –Introduction : avez-vous lu le cours ? : ABCDE

A VRAI

La moyenne vaut $\bar{x} = \frac{\sum x_i}{n}$. Elle est un paramètre de position.

Paramètres de position	Paramètres de dispersion
Moyenne	Variance
Médiane	Ecart-type
Mode ou classe modale	Quartile
	Extrêmes et étendue
	Coefficient de variation

B VRAI L'unité statistique est un élément de l'échantillon. L'ensemble des unités statistique constituent l'échantillon. Ainsi, on peut remplacer le terme unité statistique par celui d'individu lorsqu'on est dans une population.

C VRAI C'est la phrase qui est dans le cours. Le verbe **« inférer »** veut d'ailleurs dire « Tirer une conséquence de quelque chose, conclure, induire quelque chose de quelque chose ».
L'inférence statistique correspond donc à une interprétation, ce n'est pas une simple analyse descriptive ; nous tirons des conclusions générales à partir de cas particuliers

D VRAI On le voit dans le tableau de l'item A. 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. On peut aussi parler de la **classe modale** lorsqu'on travaille avec des classes, c'est celle qui correspond au plus grand effectif si toutes les classes ont la même amplitude.

E VRAI C'est la bonne formule : écart-type divisé par la moyenne. C'est une valeur de **dispersion relative** : plus il est élevé plus les valeurs sont dispersées autour de la moyenne.