

## Contrôle statistiques

### Exercice 1

1) Complétez le tableau suivant

valeurs	2	3	4	5	7	8	9	10	Total
effectif	14	19	17	27	23	46	44	10	200
effectif cumulés croissants									XXXX
fréquence									1

- 2) La calculatrice donne  $Q_1=4,5$ , expliquez pourquoi on est sûr que notre recherche à la main donnera un résultat différent
- 3) Déterminer à la main la médiane,  $Q_1$ ,  $Q_3$  et la moyenne
- 4) Expliquez pourquoi la calculatrice donne une valeur de  $Q_1$  différente de celle que vous avez trouvé.

### Exercice 2

Déterminer la moyenne de la série suivante

valeurs	12	24	57	112	13	Total
fréquence	0,25	0,15	0,35	0,2	0,05	1

### Exercice 3

Un professeur a deux classes de seconde, la première de 27 élèves a eu 13,5 de moyenne au contrôle et la seconde de 32 élèves a eu 11,2 à ce même contrôle. Quelle est la moyenne sur l'ensemble des 59 élèves des deux classes ?

### Exercice 4

valeurs	1	2	3	4	5
effectif	4	7	1	13	15

- 1) Indiquez les lignes de commandes à rentrer dans votre calculatrice pour obtenir moyenne, médiane, quartiles
- 2) Recopiez les valeurs données par la calculatrice.

### Exercice 5

Au Etats Unis la prévalence du syndrome de down est de 92 pour 100 000 naissances (donc  $p = \dots\dots\dots$ ).

Un médicament contre l'acné le « roatane » a été longtemps décrié par les associations de consommateurs, elles ont observé en étudiant 9000 naissances par des femmes ayant pris le médicament durant les 2 ans précédent leur grossesse que 22 d'entre elles ont eu un enfant présentant le syndrome de down.

Les associations soutiennent que le médicament fait plus que doubler la prévalence de la maladie.

Le laboratoire lui affirme qu'on peut être sûr à 95% que le médicament ne favorise pas cette maladie génétique.

- 1) Vérifier les dires des associations
- 2) Vérifier les dires du laboratoire
- 3) Qu'en pensez-vous ?

### Exercice 6

Lors des élections présidentielles M. X et M. Y sont les deux favoris, lors d'un sondage sur 225 personnes ils sont crédités respectivement de 30% et 34% d'intention de vote.

- 1) Donnez l'intervalle de confiance à 95% pour ces deux candidats
- 2) Au final c'est M. X qui l'emporte avec 33% des votes, contre 32% des votes pour M. Y, est ce que cela contredit les intervalles de confiance trouvés à la question précédente.

## Correction

### Exercice 1

valeurs	2	3	4	5	7	8	9	10	Total
effectif	14	19	17	27	23	46	44	10	200
effectif cumulés croissants	14	33	50	77	100	146	190	200	x
fréquence	0,07	0,095	0,085	0,135	0,115	0,23	0,22	0,05	1

On sait que l'on ne peut trouver un  $Q_1$  valant 4,5, car nous utilisons une méthode qui fait que  $Q_1$  est nécessairement une des valeurs du tableau, or 4,5 n'en n'est pas une.

La moyenne est de 6,565, la médiane de 7,5,  $Q_1$  et  $Q_3$  seront respectivement 4 et 9.

La calculatrice utilise une méthode différente de la notre pour déterminer  $Q_1$  donc il arrive que les valeurs diffèrent un peu.

### Exercice 2

$$\bar{x} = 12 \times 0.25 + 24 \times 0.15 + 57 \times 0.35 + 112 \times 0.2 + 13 \times 0.05 = 49.6$$

### Exercice 3

$$\bar{x} = \frac{27 \times 13,5 + 32 \times 11,2}{27 + 32} = \frac{722,9}{59} \approx 12,25$$

### Exercice 4

valeurs	1	2	3	4	5	Total
effectif	4	7	1	13	15	40
effectif cumulés croissants	4	11	12	25	40	x
fréquence	0,1	0,175	0,025	0,325	0,375	1

On a  $\{1,2,3,4,5\} \rightarrow L_1$

$\{4,7,1,13,15\} \rightarrow L_2$

Stat 1-var  $L_1, L_2$

On obtient alors un moyenne de 3,7, une médiane de 4,  $Q_1$  et  $Q_3$  seront respectivement 2 et 5.

### Exercice 5

- Dans la population la proportion est  $p = \frac{92}{100000} = 0,00092$  dans l'échantillon elle est de  $\frac{22}{9000} \approx 0,0024$  or 0,0024 est plus de deux fois plus grand que 0,00092, donc les associations semblent avoir raison.
- On va regarder si l'échantillon est représentatif de la population, et pour cela nous allons utiliser l'intervalle de fluctuation au seuil de 95% et voir si la fréquence de l'échantillon est dedans ou non.  $\left[ p - \frac{1}{\sqrt{n}}, p + \frac{1}{\sqrt{n}} \right] = \left[ \frac{92}{100000} - \frac{1}{\sqrt{9000}}; \frac{92}{100000} + \frac{1}{\sqrt{9000}} \right] \cdot \frac{92}{100000} - \frac{1}{\sqrt{9000}} \approx -0,0096$  et  $\frac{92}{100000} + \frac{1}{\sqrt{9000}} \approx 0,0115$  et  $f = \frac{22}{9000} \approx 0,0024$  donc la fréquence est bien dans l'intervalle de fluctuation et donc l'échantillon est représentatif de la population, on peut penser que le médicament n'a pas d'effet significatif sur la conception

### Exercice 6

- Pour M. X, p devrait être dans l'intervalle de confiance suivant :  $\left[ f - \frac{1}{\sqrt{n}}, f + \frac{1}{\sqrt{n}} \right] = \left[ 0,23 - \frac{1}{\sqrt{225}}; 0,23 + \frac{1}{\sqrt{225}} \right]$  donc on est sur à 95% que  $0,23 < p < 0,37$ . Pour M. Y, p devrait être dans l'intervalle de confiance suivant :  $\left[ f - \frac{1}{\sqrt{n}}, f + \frac{1}{\sqrt{n}} \right] = \left[ 0,34 - \frac{1}{\sqrt{225}}; 0,34 + \frac{1}{\sqrt{225}} \right]$  donc on est sur à 95% que  $0,27 < p < 0,41$
- Pour M. X on a  $p = 0,33$  et donc on a bien  $0,23 < p < 0,37$  et pour M. Y on a  $p = 0,32$  et donc on a bien  $0,27 < p < 0,41$  et donc il n'y a pas de contradiction entre les résultats des élections et les intervalles de confiance trouvés à la question précédente.