TP 3

Statistiques descriptives univariées

Compétences attendues.

✓ Regrouper une série statistique par modalités ou par classes.

✓ Connaître les indicateurs de position (moyenne, médiane, quartiles) et les commandes associées.

✓ Connaître les indicateurs de dispersion (écart-type, étendue, distance inter-quartile) et les commandes associées.

✓ Représenter graphiquement une série statistique.

Liste des commandes `Python` exigibles aux concours.

- Dans la librairie `numpy` : `np.sum, np.min, np.max, np.cumsum, np.mean, np.median, np.var, np.std`.
- Dans la librairie `matplotlib.pyplot` : `plt.hist, plt.show`.

Objectifs. L'objet des statistiques descriptives univariées (ou unidimensionnelles) est de fournir des résumés synthétiques, graphiques et numériques, de séries de valeurs observées sur une population ou un échantillon. On présente ici les indicateurs les plus couramment employés pour décrire une série statistique.

Mathieu Mansuy

Professeur en ECG deuxième année spécialité mathématiques approfondies au Lycée Louis Pergaud (Besançon)

Page personnelle : `mathieu-mansuy.fr/`

E-mail : `mathieu.mansuy@ac-besancon.fr`

1 Principales notions en statistiques descriptives

1.1 Présentation des données

On considère un ensemble Ω appelé *population* en statistique descriptive. On appellera ses éléments ω des *individus*.

Exemple. $\Omega =$ l'ensemble de la population française, $\Omega =$ l'ensemble des voitures immatriculées en France.

On étudie un *caractère* de cette population.

Définition.

Un *caractère* (ou *variable*) sur la population Ω est une application $X : \Omega \to E$, où E désigne un ensemble quelconque.

Si E est un ensemble de nombres, on dit que X est un caractère *quantitatif*. Dans le cas contraire, on parle de caractère *qualitatif*.

Exemple. Un caractère possible sur la population française est la taille (caractère quantitatif) ou encore la couleur des yeux (caractère qualitatif).

Nous ne traiterons que du cas des caractères quantitatifs.

La connaissance complète d'un caractère X peut être rendue difficile, voir impossible, de part la taille de la population Ω. Afin de pouvoir l'étudier, on peut considérer ce caractère seulement pour une partie finie $\{\omega_1, \dots, \omega_n\}$ de Ω appelée *échantillon*. Son cardinal n est alors la *taille* ou l'*effectif* de l'échantillon.

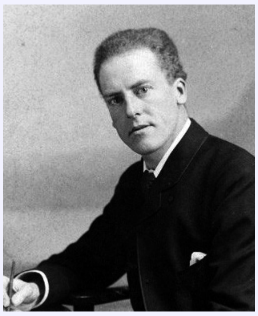
Définition.

- On appelle *série statistique* d'un échantillon $\{\omega_1, \dots, \omega_n\} \subset \Omega$ (ou *échantillon observé*) pour le caractère X la donnée de la liste $x = (x_1, \dots, x_n) = (X(\omega_1), \dots, X(\omega_n))$ des valeurs prises par X sur l'échantillon.
- Les valeurs prises par X sont appelées *modalités*.
- L'*effectif d'une modalité* m est le nombre n_m de fois où m apparait dans la série statistique.
- La *fréquence d'une modalité* m est le réel $f_m = \dfrac{\text{effectif de la modalité } m}{\text{effectif total de l'échantillon}} = \dfrac{n_m}{n}$.
- La *fréquence cumulée d'une modalité* m est le réel $p_m = \displaystyle\sum_{m' \leq m} f_{m'}$.

Le saviez-vous ?

Les statistiques sont nées en Angleterre, au début du 17-ème siècle pour décompter les décès lors d'une épidémie de peste. Ce n'était à l'époque que des données numériques, sans outil théorique pour les analyser. Il faut attendre le 19-ème siècle pour voir l'apparition de méthodes mathématiques pour l'étude de telles données. Ce n'est qu'à la fin du 19-ème siècle que la statistique devient une discipline à part entière des mathématiques sous l'impulsion des savants anglais Karl Pearson (1857-1936) et Udny Yule (1871-1951).

Karl Pearson est principalement connu pour avoir développé le coefficient de corrélation linéaire et le test du χ^2 (chi-deux).



Karl Pearson (1857-1936)

Exemple. On considère la série statistique suivante :

$$x = (2, 11, 7, 2, 15, 4, 5, 5, 5, 13, 5, 15, 7, 7, 8, 10, 10, 10, 11, 13, 7, 2, 15, 15).$$

L'ensemble des modalités est $\{2, 4, 5, 7, 8, 10, 11, 13, 15\}$. L'effectif de la modalité $m = 5$ est $n_5 = 4$, sa fréquence est $f_5 = \dfrac{n_5}{n} = \dfrac{4}{24}$ et sa fréquence cumulée est $p_5 = \dfrac{8}{24}$.

Représentation informatique. Sous `Python`, nous représenterons une série statistique $(x_1, \dots, x_n)$ par un vecteur

```
x = np.array([x1,...,xn])
```

L'effectif de la série est obtenu à l'aide de la commande `np.shape(u)[0]`

Remarques.

- Si $x = (x_1, \dots, x_n)$ est une série statistique, $(m_1, \dots, m_p)$ ses modalités d'effectifs $(n_1, \dots, n_p)$ et de fréquences $(f_1, \dots, f_p)$, alors :

$$\sum_{i=1}^{p} n_i = n \qquad \text{et} \qquad \sum_{i=1}^{p} f_i = \sum_{i=1}^{p} \frac{n_i}{n} = \frac{\sum_{i=1}^{p} n_i}{n} = 1$$

- Les notions suivantes se correspondent entre probabilités et statistiques :

$$\begin{array}{rcl}
\text{Variable aléatoire } X & \leftrightarrow & \text{Caractère } X \\
\text{Ensemble image } X(\Omega) & \leftrightarrow & \text{Ensemble des modalités de } X \\
\text{Probabilités ponctuelles } P(X = x) & \leftrightarrow & \text{Fréquences } f_m \\
\text{Fonction de répartition } F_X : x \mapsto P(X \leq x) & \leftrightarrow & \text{Fréquences cumulées } p_m = \displaystyle\sum_{m' \leq m} f_{m'}
\end{array}$$

Une série statistique brute ne permettant pas une lecture efficace des données, on souhaite la présenter de manière synthétique. Pour cela, on procède de deux manières distinctes selon le nombre de ses modalités.

Regroupement par modalités

Dans le cas où le nombre de modalités de la série est raisonnable, on regroupe la série par *modalités - effectifs*, c'est-à-dire qu'on donne :

- la liste (m_i) des modalités du caractère X,
- les effectifs (n_i) correspondants.

On peut aussi choisir de présenter cette série regroupée par *modalité - fréquence*, en donnant les modalités (m_i) et les fréquences correspondantes (f_i).

Exemple. Le tri par modalités de la série statistique x donne :

2	3
4	1
5	4
7	4
8	1
10	3
11	2
13	2
15	4

Tri par modalités - effectifs de la série x.

2	0.125
4	0.0416667
5	0.1666667
7	0.1666667
8	0.0416667
10	0.125
11	0.0833333
13	0.0833333
15	0.1666667

Tri par modalités - fréquences de la série x.

Regroupement par classes

Dans le cas où le nombre de modalités est trop grand, plutôt que de conserver toutes les valeurs, il est plus intéressant de les regrouper par classes :

- on considère une suite de réels $c = (c_0 < \cdots < c_k)$ définissant les *classes* $I_1 = [c_0, c_1]$, $I_2 =]c_1, c_2]$, $\ldots$, $I_k =]c_{k-1}, c_k]$, l'*amplitude* de la classe I_i étant $c_i - c_{i-1}$;
- on note n_i le nombre d'éléments de X appartenant à l'intervalle I_i pour $1 \leq i \leq k$.

On se ramène ainsi à une série statistique de taille k, dont les modalités sont les milieux $y_i = \dfrac{c_{i-1} + c_i}{2}$ des classes et d'effectifs les n_i correspondants.

Commandes utiles

On rappelle les commandes suivantes qui pourront être utiles dans la suite.

Définition.

- Si `u` et `v` sont deux vecteurs de même format, l'instruction `u==v` renvoie un vecteur de même format que `u` dont les éléments sont `True` ou `False` selon que les coefficients correspondants de `u` et `v` à cette même place sont égaux ou non.
- Si tous les éléments de `v` sont égaux à un même réel `x`, on peut écrire simplement `u==x`.
- On définit de même les vecteurs booléens `u>v`, `u>=v`, `u<v`, `u<=v` et `u!=v`.
- On rappelle également les connecteurs logiques pour les booléens : `and` (et), `or` (ou), `not` (négation).

Définition.

Si `u` est un vecteur dont les composantes sont des booléens, alors :

- la commande `np.sum(u)` renvoie le nombre de booléens qui ont pris la valeur `True` ;
- la commande `np.mean(u)` renvoie la proportion de booléens qui ont pris la valeur `True`.

Exemple. Supposons avoir représenté la série statistique x à l'aide d'un vecteur `x`. L'instruction `np.sum(x==7)` renvoie l'effectif de la modalité 7, `np.mean(x==7)` renvoie sa fréquence, et `np.mean(x<=7)` sa fréquence cumulée.

1.2 Indicateurs de position

Définition.

On appelle *moyenne empirique* de la série statistique $x = (x_1, \ldots, x_n)$ le réel :

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Remarque. Si la série statistique x est groupée par modalités - effectifs, avec les modalités $(m_1, \ldots, m_p)$ d'effectifs $(n_1, \ldots, n_p)$ et de fréquences $(f_1, \ldots, f_p)$, alors :

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{p} m_i \cdot n_i = \sum_{i=1}^{p} m_i \cdot \frac{n_i}{n} = \sum_{i=1}^{p} m_i \cdot f_i.$$

On notera sur cette dernière formule la correspondance entre les notions d'espérance en probabilité et de moyenne en statistique :

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x) \quad \leftrightarrow \quad \overline{x} = \sum_{i=1}^{p} m_i \cdot f_i.$$

Définition.

La *médiane* d'une série statistique ordonnée est un réel m partageant la série en deux séries d'effectifs égaux. Si $(x_1 \leq x_2 \leq \cdots \leq x_n)$ est la série statistique ordonnée, m est défini par :

- si $n = 2p - 1$ est impaire, $m = x_p$ (la valeur du milieu) ;
- si $n = 2p$ est paire, $m = \dfrac{x_p + x_{p+1}}{2}$ (la moyenne des deux termes du milieu).

Propriété 1 (Moyenne et médiane en `Python`)

- `np.mean(x)` donne la moyenne du vecteur `x`.
- `np.median(x)` donne une médiane du vecteur `x` (non nécessairement ordonné).

Définition.

Soit $x = (x_1, \ldots, x_n)$ une série statistique.

- Le *premier quartile* de x est la plus petite modalité q_1 de x pour laquelle 25 % des éléments de x lui sont inférieures ou égales.
- Le *troisième quartile* de x est la plus petite modalité q_3 de x pour laquelle 75 % des éléments de x lui sont inférieures ou égales.

Remarque. De même, on définit les *déciles* et les *centiles* d'une série statistique :

- Pour $k \in [\![1, 99]\!]$, le k-ième centile est la modalité c_k de la série pour laquelle moins de k % de la population prend des valeurs strictement inférieures à c_k et moins de $(100 - k)$ % de la population prend des valeurs strictement supérieures à c_k.
- Pour $k \in [\![1, 9]\!]$, le k-ième décile est la modalité d_k de la série pour laquelle moins de k % de la population prend des valeurs strictement inférieures à d_k et moins des $(10 - k)$ dixièmes de la population prend des valeurs strictement supérieures à d_k.

Définition.

On appelle *mode* d'une série statistique toute modalité pour laquelle l'effectif est maximal (il peut y en avoir plusieurs).

Exemple. Reprenons l'exemple de la série statistique x, qu'on trie par modalités - fréquences et par modalités - fréquences cumulées :

2	0.125
4	0.0416667
5	0.1666667
7	0.1666667
8	0.0416667
10	0.125
11	0.0833333
13	0.0833333
15	0.1666667

Tri par modalités - fréquences de x.

2	0.125
4	0.1666667
5	0.3333333
7	0.5
8	0.5416667
10	0.6666667
11	0.75
13	0.8333333
15	1

Tri par modalités - fréquences cumulées de x.

Déterminer le premier quartile, le troisième quartile et le huitième décile, ainsi que le(s) mode(s) de la série x.

1.3 Indicateurs de dispersion

Définition.

Soit $x = (x_1, \dots, x_n)$ une série statistique.

- On appelle *variance* de x le nombre réel positif : $v = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2$.
- On appelle *écart-type* de x le réel $\sigma = \sqrt{v}$.

Remarques.

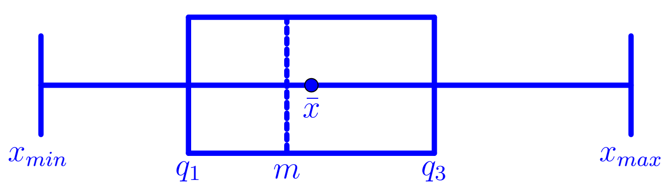
- La variance mesure la dispersion de la série statistique autour de sa moyenne.
- Comme en probabilités, la formule de Koenig-Huygens est valable :

$$v = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\overline{x} + \overline{x}^2) = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2\times\overline{x}\times\underbrace{\frac{1}{n}\sum_{i=1}^{n}x_i}_{=\overline{x}} + \overline{x}^2 = \frac{1}{n}\left(\sum_{i=1}^{n}x_i^2\right) - \overline{x}^2.$$

Définition.

- On appelle *étendue* d'une série statistique la différence entre la plus grande et la plus petite modalité.
- On appelle *distance inter-quartile* le réel $q_3 - q_1$.

Remarque. La distance inter-quartile est un indicateur de dispersion : c'est la longueur de l'*intervalle inter-quartile* $[q_1, q_3]$, lequel contient la moitié des valeurs de la série, réparties autour de la médiane m. Pour rendre compte de l'étendue et de la distance inter-quartile d'une série statistique, on représente parfois sa *boîte à moustache* :



Propriété 2 (Variance, écart-type et étendue sur `Python`)

- `np.var(x)` donne la variance du vecteur `x`.
- `np.std(x)` (pour standard deviation) donne l'écart-type du vecteur `x`.
- `np.max(x)-np.min(x)` donne l'étendue du vecteur `x`.

Exemple. Déterminer l'écart-type de la série statistique x, et représenter son diagramme à moustache.

2 Représentations graphiques

On suppose avoir importé la bibliothèque `matplotlib.pyplot` à l'aide de l'instruction :

```
import matplotlib.pyplot as plt
```

2.1 Diagrammes en bâtons

Définition.

On représente une série statistique **groupée par modalités** à l'aide d'un *diagramme en bâtons*, en plaçant sur l'axe des abscisses les modalités et en dressant à la verticale de chacune d'elles un bâton de hauteur égale à son effectif ou sa fréquence.

On dispose de la commande `Python` suivante (non exigible).

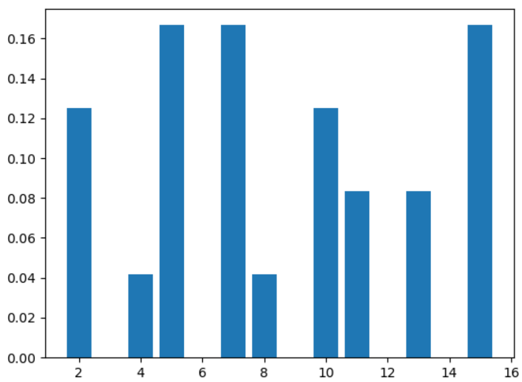
Définition.

Si `x` et `y` sont des vecteurs de même taille, `plt.bar(x,y)` trace le diagramme en bâtons d'abscisses contenues dans `x` et d'ordonnées dans `y`.

Exemple. Reprenons la série statistique x. En entrant les instructions suivantes dans la console :

```
>>> m = np.array([2.,4.,5.,7.,8.,10.,11.,13.,15.])
>>> f = np.array([0.125,0.04167,0.16667,0.16667,0.04167,0.125,0.08333,0.08333,0.16667])
>>> plt.bar(m,f)
>>> plt.show()
```

on obtient le diagramme en bâtons des effectifs de la série x :



Remarque. La commande `plt.bar` nécessite d'avoir trié au préalable la série statistique par modalités - effectifs, ce que nous ne pourrons pas toujours faire. En effet, nous n'avons malheureusement pas de commande `Python` pour effectuer ce tri. Nous expliquerons ci-dessous comment tracer le diagramme en bâtons d'une série statistique brute (i.e. sans tri préalable) à l'aide de la commande `plt.hist`.

2.2 Histogrammes

Définition.

On représente une série statistique **groupée par classes** à l'aide d'un *histogramme*, en plaçant les c_i sur un axe horizontal et en traçant à la verticale un rectangle de base $[c_i, c_{i+1}]$ et d'aire égale à la fréquence de la classe correspondante.

Définition.

Soit `x` un vecteur.

- L'instruction `plt.hist(x,n)` trace l'histogramme associé à la série `x` en `n` classes équiréparties entre la plus petite valeur de `x` et la plus grande (par défaut, `n` vaut 10).
- L'instruction `plt.hist(x,c)` trace l'histogramme associé à la série `x` dont les classes sont définies par le vecteur aux composantes strictement croissantes `c`.

Méthode. Comment tracer un diagramme en bâtons d'une série statistique brute ?

Pour tracer le diagramme en bâtons d'une série statistique brute (non triée) `x` *à valeurs entières, on procède ainsi :*

(i) on détermine les modalités $m_1 < m_2 < \cdots < m_k$ de la série statistique `x` *;*

(ii) on définit les classes $c = (m_1 - 0,5 < m_1 + 0,5 < m_2 - 0,5 < m_2 + 0,5 < \cdots < m_k - 0,5 < m_k + 0,5)$;

(iii) on dessine l'histogramme (le « diagramme en bâtons ») des effectifs à l'aide de la commande :

```
plt.hist(x,c,edgecolor='k',color='...', label="...")
```

et l'histogramme (le « diagramme en bâtons ») des fréquences à l'aide de la commande :

```
plt.hist(x,c,density='True',edgecolor='k',color='...', label="...")
```

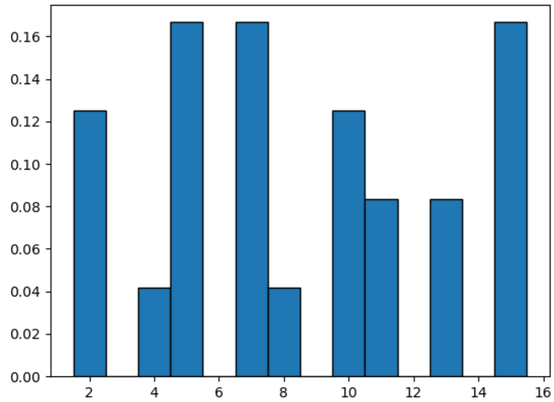
où l'on a ajouté les options de tracé suivantes (non exigibles) :

- *normalisation des rectangles (la surface totale vaut 1) :* `density='True'`
- *contours des rectangles en noir :* `edgecolor='k'`
- *couleur des rectangles :* `color='...'` *(indiquer la couleur en anglais)*
- *légende associée à chaque histogramme :* `label="..."` *(indiquer la légende choisie)*

Exemple. En entrant les instructions suivantes dans la console :

```
>>> c = np.arange(np.min(x),np.max(x)+2)-0.5
>>> plt.hist(x, c, density = 'True', edgecolor = 'k')
>>> plt.show()
```

on obtient l'histogramme (le « diagramme en bâtons ») des fréquences de la série x :



Notons qu'ici, nous n'avons pas eu au préalable à trier la série statistique x par modalités - fréquences, c'est fait par `Python` en exécutant `plt.hist`.

3 Exercices

Dans les deux exercices suivants, nous aurons besoin de fonctions de la bibliothèque `numpy.random`. On va donc l'importer à l'aide de l'instruction :

```
import numpy.random as rd
```

Exercice 1 (★★)

La commande `rd.binomial(n,p,r)` renvoie un vecteur contenant r simulations de la loi binomiale $\mathscr{B}(n, p)$.

1. À l'aide de la commande `rd.binomial`, simuler 10000 nombres suivant la loi $\mathscr{B}(10, 0.5)$. On notera `x` le vecteur contenant cette série statistique.

2. Déterminer l'effectif, la fréquence et la fréquence cumulée de la modalité 5.

3. Déterminer la moyenne, la médiane et l'écart-type de `x`. Était-ce prévisible ?

4. Créer un vecteur `m` de taille 11 tel que `m[k]` contient l'effectif de la modalité k. Déterminer le(s) mode(s) de la série `x`.

5. Représenter à l'aide de la commande `plt.bar` les diagrammes en bâtons des effectifs, des fréquences et des fréquences cumulées de la série `x`.

6. Représenter de nouveau le diagramme en bâtons des effectifs et des fréquences de la série `x`, cette fois à l'aide de la commande `plt.hist`.

Exercice 2 (★★)

La commande `rd.random(n)` permet de simuler un vecteur de taille n dont chaque coefficient est un nombre réel choisi aléatoirement entre 0 et 1.

1. Créer un vecteur `x` contenant 10000 nombres réels choisis aléatoirement entre 1 et 5.

2. Calculer la moyenne, la médiane, l'écart-type et l'étendue de la série statistique `x`.

3. Faut-il mieux regrouper cette série statistique par modalités ou par classes ? Pourquoi ?

4. Tracer l'histogramme associé à cette série statistique en la regroupant par classes (choisir 100 classes de même amplitude). Que remarque-t-on ?

---

4 Étude de la population mondiale

À partir de mon site `mathieu-mansuy.fr/ecg2`, ouvrir le document `population-mondiale`, permettant de définir plusieurs vecteurs :

- `pays` contient les noms des pays ;
- `superficie` contient la surface terrestre en milliers de km^2 de chaque pays ;
- `population` contient le nombre d'habitants en millions de chaque pays ;
- `naissance` contient le nombre de naissances sur 1000 habitants ;
- `mort` contient le nombre de décès sur 1000 habitants ;
- `homme` contient l'espérance de vie des hommes ;
- `femme` contient l'espérance de vie des femmes ;

à partir des données contenues dans l'étude 2017 de l'Institut National d'Études Démographiques (disponible également sur mon site). Pour les exercices qui suivent, vous trouverez en annexe de ce TP la liste des pays et de leurs index.

Exercice 3 (Interrogation de la base de données)
Écrire une fonction `donnees(n)` qui affiche le nom, la superficie, le nombre d'habitants et la densité de population du pays d'index n.

Exercice 4

1. Calculer la surface terrestre mondiale, le nombre d'habitants mondial et la densité moyenne d'habitants au km^2.
2. Calculer la surface terrestre, le nombre d'habitants et la densité moyenne d'habitants au km^2 pour chaque continent.
3. Représenter la densité moyenne d'habitants au km^2 pour chaque continent en utilisant un diagramme en bâtons (on mettra en abscisse des entiers de 1 à 5).
4. Faire de même pour la répartition de la surface terrestre par continent, puis du nombre d'habitants par continent.

Exercice 5
On considère l'espérance de vie des hommes (ou des femmes) par pays.

1. Calculer la moyenne sur l'ensemble des pays. Ce résultat correspond-il à l'espérance de vie mondiale des hommes (ou des femmes) ?
2. Calculer l'écart-type et la médiane.
3. Calculer les espérances de vie minimale et maximale en précisant les pays correspondant à ces valeurs extrémales.
4. Représenter l'histogramme de l'espérance de vie des hommes sur l'intervalle $[0, 100]$ avec 20 classes. Quelle est la classe modale de l'espérance de vie des hommes ?
5. On s'intéresse au tableau `homme`.

   (a) Vérifier que ses premier et troisième quartiles valent respectivement 64 et 76.

(b) Vérifier que ses premier et neuvième déciles valent respectivement 59 et 80. Donner la liste des pays dont l'espérance de vie des hommes est inférieure au premier décile ou supérieure au neuvième décile.

Exercice 6

On rappelle que le taux d'accroissement naturel est la différence entre la natalité et la mortalité.

1. Quels sont les accroissements minimaux et maximaux ? Préciser les pays.
2. Faire afficher la liste des pays pour lesquels l'accroissement est négatif.
3. Déterminer l'accroissement mondial moyen.
4. Dans ses projections, l'INED prévoit une population mondiale de 9731 millions d'habitants en 2050. Cela est-il conforme à l'hypothèse d'un taux d'accroissement constant ?

Annexe : Liste des pays et de leurs index

Afrique

Afrique septentrionale

1. Algérie
2. Égypte
3. Libye
4. Maroc
5. Sahara occidental
6. Soudan
7. Tunisie

Afrique occidentale

8. Bénin
9. Burkina Faso
10. Cap-Vert
11. Côte d'Ivoire
12. Gambie
13. Ghana
14. Guinée
15. Guinée-Bissau
16. Liberia
17. Mali
18. Mauritanie
19. Niger
20. Nigeria
21. Sénégal
22. Sierra Leone
23. Togo

Afrique orientale

24. Burundi
25. Comores
26. Djibouti
27. Érythrée
28. Éthiopie
29. Kenya
30. Madagascar
31. Malawi
32. Maurice
33. Mayotte
34. Mozambique
35. Ouganda
36. Réunion
37. Rwanda
38. Seychelles
39. Somalie
40. Sud-Soudan
41. Tanzanie
42. Zambie
43. Zimbabwe

Afrique centrale

44. Angola
45. Cameroun
46. Centrafricaine (Rép.)
47. Congo
48. Congo(Rép. dém.)
49. Gabon
50. Guinée équatoriale
51. Sao Tomé-et-Principe
52. Tchad

Afrique australe

53. Afrique du Sud
54. Botswana
55. Lesotho
56. Namibie
57. Swaziland

Amérique du Nord

Amérique septentrionale

58. Canada
59. États Unis

Amérique centrale

60. Belize
61. Costa Rica
62. Guatemala
63. Honduras
64. Mexique
65. Nicaragua
66. Panama
67. Salvador

Caraïbes

68. Antigua-et-Barbuda
69. Aruba
70. Bahamas
71. Barbade
72. Cuba
73. Curaçao
74. Dominicaine (Rép.)
75. Dominique
76. Grenade
77. Guadeloupe
78. Haïti
79. Jamaïque
80. Martinique
81. Porto Rico
82. Sainte Lucie
83. St Vincent Grenadines
84. St.Kitts-et-Nevis
85. Trinité-et-Tobago
86. Vierges (Îles)

Amérique du Sud

87. Argentine
88. Bolivie
89. Brésil
90. Chili
91. Colombie
92. Équateur
93. Guyana
94. Guyane (française)
95. Paraguay
96. Pérou
97. Surinam
98. Uruguay
99. Venezuela

Asie

Asie occidentale

100. Arabie saoudite
101. Arménie
102. Azerbaïdjan
103. Bahreïn
104. Chypre
105. Émirats arabes unis
106. Géorgie
107. Irak
108. Israël
109. Jordanie
110. Koweït
111. Liban
112. Oman
113. Palestine (Territoires)
114. Qatar
115. Syrie
116. Turquie
117. Yémen

Asie centrale

118. Kazakhstan
119. Kirghizistan
120. Tadjikistan
121. Turkménistan
122. Ouzbékistan

Asie du sud

123. Afghanistan
124. Bangladesh
125. Bhoutan
126. Pakistan
127. Inde
128. Iran
129. Maldives
130. Népal
131. Sri Lanka

Asie du sud-ouest

132. Brunei
133. Cambodge
134. Indonésie
135. Laos
136. Malaisie
137. Myanmar (Birmanie)
138. Philippines
139. Singapour
140. Thaïlande
141. Timor-Est
142. Viêt Nam

Asie orientale

143. Chine
144. Chine-Hong Kong
145. Chine-Macao
146. Corée du Nord
147. Corée du Sud
148. Japon
149. Mongolie
150. Taïwan

Europe

Europe septentrionale

151. Danemark
152. Estonie
153. Finlande
154. Irlande
155. Islande
156. Lettonie
157. Lituanie
158. Norvège
159. Royaume-Uni
160. Suède

Europe occidentale

161. Allemagne
162. Autriche
163. Belgique
164. France (métropolitaine)
165. Liechtenstein
166. Luxembourg

167. Monaco
168. Pays-Bas
169. Suisse

Europe orientale

170. Biélorussie
171. Bulgarie
172. Hongrie
173. Moldavie
174. Pologne
175. Roumanie
176. Russie
177. Slovaquie
178. Tchèque (République)
179. Ukraine

Europe méridionale

180. Albanie
181. Andorre
182. Bosnie-Herzégovine
183. Croatie
184. Espagne
185. Grèce
186. Italie
187. Kosovo
188. Macédoine
189. Malte
190. Monténégro
191. Portugal
192. Saint-Marin
193. Serbie
194. Slovénie

Océanie

195. Australie
196. Fidji
197. Guam
198. Kiribati
199. Marshall (Îles)
200. Micronésie (États fédérés de)
201. Nouvelle-Calédonie
202. Nouvelle-Zélande
203. Papouasie-Nouvelle Guinée
204. Polynésie française
205. Salomon (Îles)
206. Samoa occidentales
207. Tonga
208. Vanuatu