

↔ Activités du chapitre XIII

# Échantillonnage

## Activité d'introduction

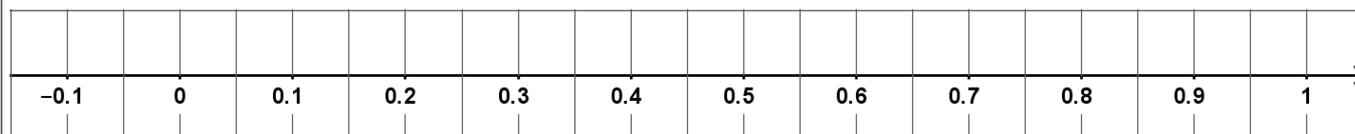
l'année scolaire 2013–2014, dans la classe de **Seconde 7** du lycée Europe, il y avait 9 garçons et 26 filles, ce qui paraît disproportionné par rapport à la moyenne nationale (env. 50% de filles et 50% de garçons).

### A. Simulations et observations

**Problème :** *Lorsqu'on choisit 35 personnes au hasard dans la population nationale, est-il fréquent d'obtenir la même répartition Filles – Garçons que dans la Seconde 7 ?*

- Calculer la fréquence des filles dans la classe de **Seconde 7** :  $f =$
- À l'aide de la calculatrice, simuler le choix de 35 personnes au hasard dans une population constituée d'une moitié de filles et d'une moitié de garçons : MENU 1 OPTN PROB > Ran#
  - Nombre** de fille obtenu dans cette simulation :
  - Fréquence** des filles dans cette simulation (arrondie au centième) :
- Rentrer cette fréquence sur la feuille de calcul commune et noter les résultats des simulations de toute la classe dans le tableau ci-dessous :


- Pour cette série statistique :
  - déterminer les **valeurs extrêmes**, les premier et troisième **quartiles**, la **médiane** et la **moyenne** :
    - $x_{min} =$                        $x_{max} =$                        $Q_1 =$                        $M_e =$                        $Q_3 =$
  - Représenter ci-dessous le **diagramme en boîte** :



- Donner l'**intervalle inter-quartile** :  $I =$
- D'après ces résultats, peut-il arriver que le hasard produise une distribution comparable à celle de la **Seconde 7** ? Si oui, est-ce fréquent ?

### B. Étude du cours [Onglet n° 16 : Probabilité]

### C. Test de représentativité

**Problème :** *Peut-on avancer la proposition suivante, avec un risque de 5 % de se tromper :*

« L'**échantillon** (la classe de **Seconde 7**) est **représentatif** de la population nationale (constituée d'une moitié de filles et d'une moitié de garçons). »

1. Dans cette question, on utilise les mêmes notations que dans le cours. Compléter :

- (a) Caractère étudié :  
Proportion du caractère au sein de la population de référence (population nationale) :  $p =$
- (b) Taille de l'échantillon :  $n =$   
Fréquence d'observation du caractère observée dans l'échantillon :  $f =$
- (c) Intervalle de fluctuation :  $I_n =$
- (d) Répondre au problème posé :

2. Et si le choix de  $p = 0,5$  n'était pas pertinent.

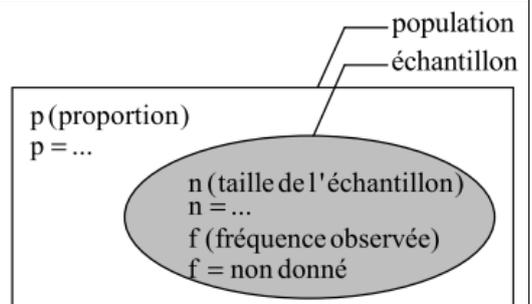
À l'administration du lycée, on pouvait obtenir l'information suivante : « Au lycée Europe, pour l'année scolaire 2013–2014, il y a 524 élèves en Seconde, dont 329 filles et 195 garçons ».

- (a) Déterminer l'intervalle de fluctuation correspondant, pour un échantillon de taille 35.
- (b) La fréquence des filles de la Seconde 7 appartient-elle à cet intervalle ? Conclure.

**Exo méthodo n° 1: Caractéristiques d'un échantillon**

La proportion  $p$  d'un caractère dans une population est égale à 0,65.

Compléter le schéma ci-contre puis déterminer l'intervalle de fluctuation au seuil de 95% de la fréquence d'un échantillon de taille 1000.

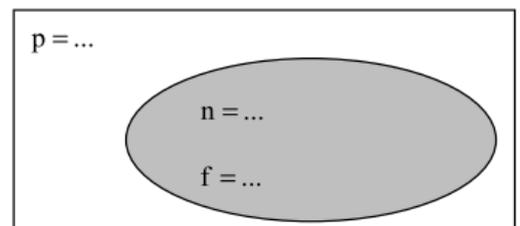


**Exo méthodo n° 2: Sondage pour une élection**

Un candidat à un poste de député a été élu avec 51% des voix.

Dans un village de sa circonscription, après dépouillement sur 1253 votes, 49% étaient en sa faveur.

Ce village est-il représentatif de la couleur politique de la circonscription ?



**Exercice n° 1: Expérience aléatoire et intervalle de fluctuation**

Une urne contient 10 boules : cinq rouges, trois noires et deux blanches. On tire une boule et on note sa couleur et on la remet dans l'urne.

1. Avec la table de nombres aléatoires entiers de 0 à 9 donnée ci-dessous, simuler 25 tirages en expliquant la méthode adoptée.

8	4	5	7	1	4	4	7	3	9	1	5	3	8	6	8	7	4	7	0
0	8	3	8	4	9	3	7	5	6	9	1	6	3	8	9	7	6	7	2
2	3	6	6	8	9	3	6	6	0	5	6	4	9	6	3	6	7	1	4
8	2	4	6	4	8	3	5	6	1	4	6	4	4	2	9	0	9	4	6
5	7	8	3	2	3	8	0	0	0	6	9	3	2	8	0	7	3	5	2

- 2. Calculer les fréquences obtenues pour chaque couleur.
- 3. Déterminer pour chacune des couleurs l'intervalle de fluctuation pour un échantillon de taille 25. Les fréquences obtenues à la question 2 sont-elles dans ces intervalles ? Conclure.

### Exercice n° 2: Comosition d'un jury

*L'ensemble des faits évoqués ci-dessous est réel.*

En novembre 1976, dans un comté du sud du Texas, RODRIGO PARTIDA était condamné à huit ans de prison pour cambriolage d'une résidence.

Il attaqua ce jugement au motif que la désignation des jurés de ce comté était discriminante à l'égard des Américains d'origine mexicaine. Alors que 79,1 % de la population de comté était d'origine mexicaine, sur les 870 personnes convoquées pour être jurés lors d'une certaine période de référence, il n'y eût que 339 personnes d'origine mexicaine.

1. Déterminer l'intervalle de fluctuation correspondant à la proportion d'origine mexicaine pour un échantillon de taille 870.
2. La fréquence des personnes d'origine mexicaine dans les personnes convoquées est-elle dans cet intervalle ?
3. Que peut-on en conclure ?

### Exercice n° 3: Élection présidentielle de 2007

Au premier tour de l'élection présidentielle française de mai 2007, parmi les suffrages exprimés, les proportions, en pourcentage, pour les candidats ayant obtenu pour de 2 % des suffrages, étaient les suivantes :

Bayrou	Besancenot	De Villiers	Le Pen	Royal	Sarkozy
18,57	4,08	2,23	10,44	25,87	31,18

Cinq mois plus tôt, le 13 décembre 2006, l'institut de sondage BVA faisait paraître un sondage effectué sur un échantillon de 797 personnes dont voici les résultats, en pourcentage, concernant les candidats précédemment cités :

Bayrou	Besancenot	De Villiers	Le Pen	Royal	Sarkozy
7	4	2	10	34	32

1. Pour quels candidats peut-on appliquer les intervalles de fluctuation parmi ceux présents au premier tour ?
2. Pour ces candidats déterminer les intervalles de fluctuation pour un échantillon de taille 797.
3. Les résultats du sondage donnent-ils des fréquences appartenant à ces intervalles ?
4. Qu'en conclure ?

### Exercice n° 4: Fluctuations d'échantillonnage

*Les questions 1 et 2 sont indépendantes.*

1. Dans un lycée, pour la session 2014 du baccalauréat général, il y a eu 290 reçus pour 320 candidats se présentant à l'épreuve.

Les fréquences des reçus en Série L, ES et S étaient, respectivement, 0,766, 0,896 et 0,963.

Déterminer si les différences de réussite entre les filières peuvent être dues aux fluctuations d'échantillonnage.

2. Dans le village chinois de Xicun en 2000, il est né 20 enfants dont 16 garçons.

On suppose que la proportion de garçons et de filles est la même à la naissance dans toute l'espèce humaine.

Déterminer si la fréquence des naissances de garçons dans le village de Xicun en 2009 peut être due aux fluctuations d'échantillonnage.

**Exercice n° 5: Parité hommes/femmes**

Les questions 1 et 2 sont indépendantes.

1. On suppose que la proportion de femmes dans la population française est 0,5.  
 À l'assemblée nationale, il y a 577 députés, dont 108 femmes.  
 Peut-on considérer que cette répartition est un effet de la fluctuation d'échantillonnage ou bien dire que la parité des sexes n'est pas respectée à l'assemblée nationale?
2. En 1990, les employés et ouvriers constituaient 58,7 % de la population française (d'après le recensement de l'INSEE).  
 Suite à l'élection législative de 1993 on recensait 1,6 % de députés dont l'ancien métier était employé ou ouvrier.  
 Peut-on considérer que cette répartition est un effet de la fluctuation d'échantillonnage ?

**Problème n° 1.**

Dans une région où il y a autant de femmes que d'homme, les entreprises sont tenues de respecter la parité.

L'entreprise A a un effectif de 100 personnes dont 43 femmes.

L'entreprise B a un effectif de 2 500 personnes dont 1 150 femmes.

1. Calculer le pourcentage de femmes dans ces deux entreprises. Qu'en conclure ?
2. Si respecter la parité revient à ne pas tenir compte du caractère homme-femme, on peut alors considérer l'ensemble des salariés d'une entreprise comme un échantillon prélevé au hasard dans la population de la région.
  - (b) Déterminer les intervalles de fluctuation relatifs aux deux échantillons.
  - (b) Les résultats confirment-ils la conclusion de la première question ?

**Problème n° 2.**

Les données du tableau ci-dessous sont celles de l'année scolaire pour les Premières générales d'un lycée de pour l'année scolaire 2014–2015 :

	1 ES	1 S	1 L	Total
Filles	76	92	50	218
Garçons	43	76	13	132
Total	119	168	63	350

Les résultats seront donnés au centième.

1. On s'intéresse d'abord à la proportion de garçons et de filles dans l'établissement.
  - (a) Déterminer les proportions de garçons et de filles dans le lycée cette année là.  
 Peut-on utiliser les intervalles de fluctuations dans le cas des filles et des garçons ?
  - (a) Déterminer les intervalles de fluctuations pour des échantillons de tailles respectives 119, 168 et 63.
  - (a) Calculer les fréquences de garçons et de filles dans chacune des trois filières.
  - (a) Dans quelles filières peut-on dire, au seuil de 95 %, que la fréquence des filles et des garçons peut être due aux fluctuations d'échantillonnage ?
2. En s'inspirant de la question précédente, déterminer pour chaque sexe si l'on peut dire, au seuil de 95 %, que la fréquence des ES, S et L peut être due aux fluctuations d'échantillonnage.