

Fluctuation d'échantillonnage

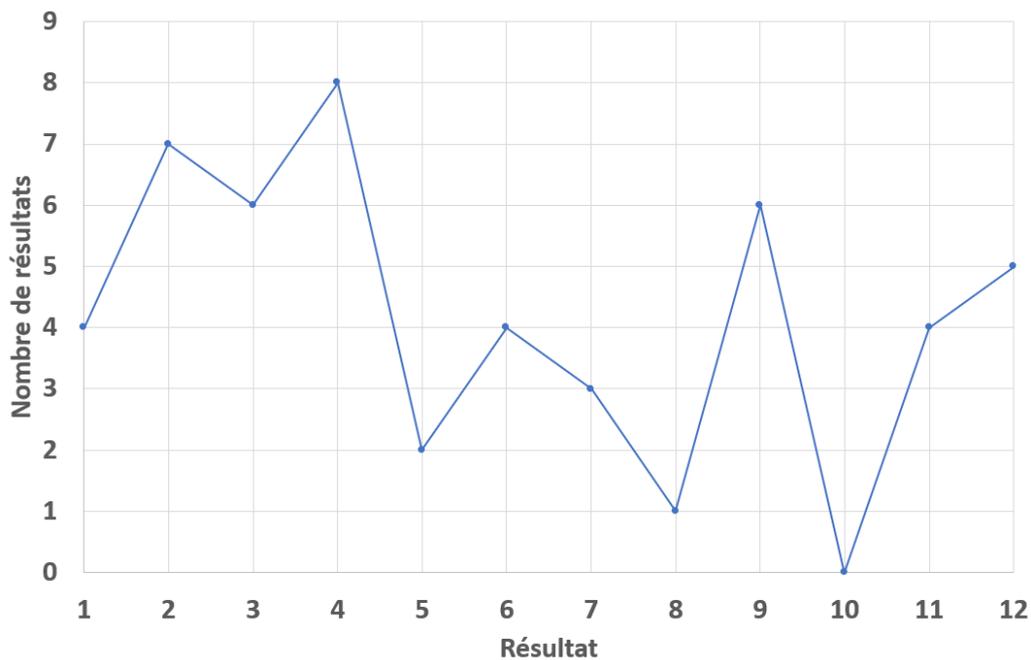
1er Gestion - Administration

Exercice 1

Pour la fête du lycée, Keylane organise une loterie avec une roue permettant de tirer un numéro de 1 à 12.



1. Pour tester la roue, il la lance un certain nombre de fois et reporte les résultats dans le graphique ci-dessous.



a) **Déterminer** la taille de l'échantillon ainsi réalisé.

.....
.....

b) **Calculer** la fréquence de sortie du nombre 12.

.....
.....

2. On simule 50 lancers de la roue et on obtient les résultats suivants :

7	1	7	11	11	7	11	7	11	1
7	11	9	9	9	11	10	1	3	4
5	1	11	5	1	10	5	10	10	5
4	3	8	7	11	12	3	10	9	12
2	11	8	8	6	9	11	11	9	7

a) Combien de fois le 12 apparait-il?

.....

b) Calculer la fréquence d'apparition du nombre 12.

.....

3. On refait 8 fois la simulation pour 50 lancers et on obtient la fréquence du nombre 12 pour chacune des simulations :

Échantillon	1	2	3	4	5	6	7	8
Fréquence	0.1	0.1	0.08	0.04	0.1	0.06	0.08	0.04

Calculer la moyenne des fréquences obtenues.

.....

.....

4. Si la roue est équilibrée, la probabilité de sortie du nombre 12 est $p = \frac{1}{12}$.

Comparer cette probabilité avec la moyenne des fréquences.

.....

.....

Exercice 2

Le tableau suivant indique, pour chaque président élu aux États-Unis depuis 1936, la proportion p en faveur de l'élu le jour du vote et la proportion f prévue par le sondage Gallup avec l'élection (source : www.gallup.com).

Année	Président élu	Vote p	Sondage f	Ecart
2008	Obama	0.55	0.53	0.02
2004	Bush	0.49	0.507	0.017
2000	Bush	0.48	0.479	0.001
1996	Clinton	0.52	0.492	0.028

Année	Président élu	Vote p	Sondage f	Ecart
1992	Clinton	0.49	0.433	0.057
1988	Bush	0.56	0.53	0.03
1984	Reagan	0.59	0.592	0.002
1980	Reagan	0.47	0.508	0.038
1976	Carter	0.48	0.501	0.021
1972	Nixon	0.62	0.618	0.002
1968	Nixon	0.43	0.435	0.005
1964	Johnson	0.64	0.613	0.027
1960	Kennedy	0.505	0.501	0.004
1956	Eisenhower	0.595	0.578	0.017
1952	Eisenhower	0.51	0.554	0.044
1948	Truman	0.445	0.495	0.05
1944	Roosevelt	0.515	0.538	0.023
1940	Roosevelt	0.52	0.55	0.03
1936	Roosevelt	0.557	0.625	0.068

1. Comment la dernière colonne est-elle calculée ?

.....

2. **Donner** les deux années où l'écart entre le sondage et le résultat du vote a été le plus grand.

.....

3. Pour les quatre premiers sondages (1936 à 1948), Gallup n'utilisait pas de tirage aléatoire (méthode des "quotas"), alors que depuis 1952 cet institut utilise un échantillonnage aléatoire.

a) **Calculer** la moyenne des écarts de 1936 à 1948.

.....

.....

b) **Vérifier** que l'écart moyen de 1952 à 2008 est plus de deux fois inférieur au précédent.

.....

.....

c) Que peut-on conclure pour la méthode aléatoire ?

.....

4. Si l'on compare les sondages Gallup à des échantillons aléatoire de taille 1000, la fluc-

tuation d'échantillonnage prévoit, dans environ 95% des cas, un écart inférieur à $\frac{1}{\sqrt{1000}}$.

a) **Donner** une valeur approchée, à 0.001 près, de $\frac{1}{\sqrt{1000}}$.

.....

b) Combien de sondages ont-ils un écart supérieur à la valeur précédente ?

.....

Exercice 3

Partie A : au milieu du XXe siècle

Vers les années 1950, face à l'augmentation de la consommation de tabac et du nombre de cancers du poumon, débutent des études statistiques sur le sujet. À l'hôpital Bellevue, en 1952, la fréquence des "grands fumeurs" (plus de 15 cigarettes par jour) parmi les malades est 44%. Parmi les 1357 malades soignés pour un cancer du poumon, 806 sont de grands fumeurs (source : Zeiseil et Kaye, *Prove it with Figure*).

1. **Calculer**, à 0.01 près, les bornes $p - \frac{1}{\sqrt{n}}$ et $p + \frac{1}{\sqrt{n}}$ de l'intervalle de fluctuation de 95% des fréquences des échantillons aléatoires de tailles $n = 1357$, lorsqu'on suppose que la proportion de grands fumeurs est $p = 0.44$.

.....
.....

2. **Calculer** la fréquence f des grands fumeurs parmi les 1357 malades atteints de cancer du poumon.

.....
.....

3. Est-il raisonnable de penser que la différence entre f et p est uniquement due au hasard ? **Expliquer**.

.....
.....
.....

Partie B : à la fin du XXe siècle

Lors d'un sondage aléatoire effectué aux États-Unis en 1995, sur 737 fumeurs quotidiens, seuls 295 estimèrent courir un risque de cancer supérieur à celui des non-fumeurs

de leur âge (source : *Journal of the American Medical Association*, 1999).

1. Calculer la fréquence f des fumeurs interrogés pensant prendre un risque.

.....

2. Si l'on suppose que 50% des fumeurs aux États-Unis pensent prendre un risque, **calculer** l'intervalle de fluctuation $\left[0.5 - \frac{1}{\sqrt{737}}; 0.5 + \frac{1}{\sqrt{737}}\right]$ des fréquences de plus de 95% des échantillons de taille 737.

.....

3. D'après le sondage effectué, peut-on estimer que moins de 50% des fumeurs aux États-Unis pensent prendre un risque ? **Expliquer.**

.....

.....

.....

Exercice 4

Pour décider la construction d'un grand stade, la municipalité d'une ville importante voudrait sonder la population pour estimer si plus de 50% des électeurs sont favorables. Si cette proportion est $p = 0.5$, plus de 95% des sondages aléatoires de taille n fournissent une fréquence dans l'intervalle $I = \left[0.5 - \frac{1}{\sqrt{n}}; 0.5 + \frac{1}{\sqrt{n}}\right]$. La municipalité ne décidera de construire le stade que si la fréquence obtenue lors du sondage est supérieur à $0.5 + \frac{1}{\sqrt{n}}$.

1. Calculer les bornes de l'intervalle I pour $n = 100$.

.....

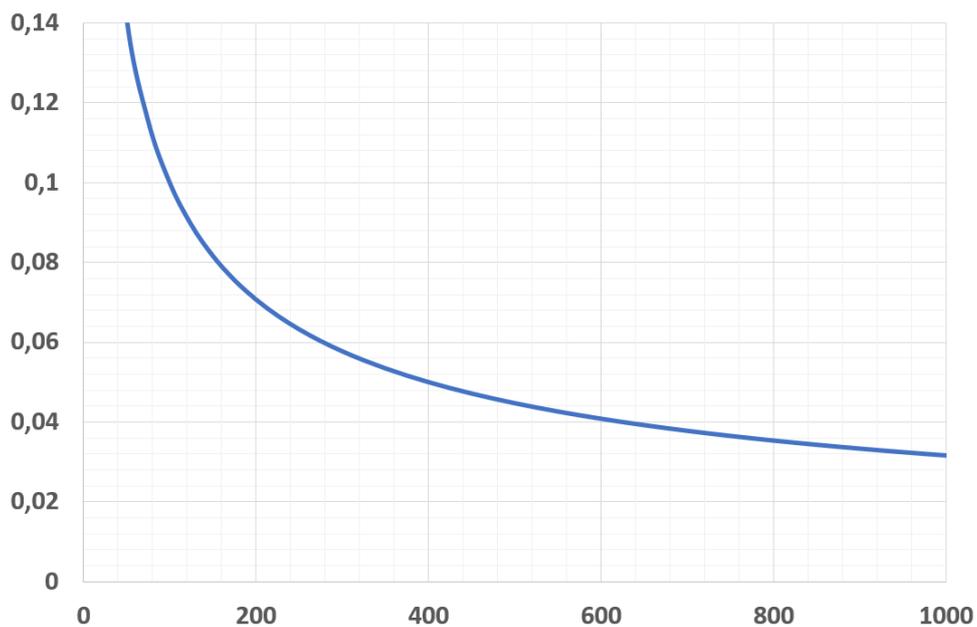
.....

2. Sur 100 électeurs interrogés au hasard, 54 sont favorables à la construction du stade. La municipalité décide-t-elle de construire le stade ?

.....

.....

3. La courbe ci-dessous a pour équation $f(x) = \frac{1}{\sqrt{x}}$



Lire approximativement la plus petite valeur de x pour laquelle on a $\frac{1}{\sqrt{x}} \leq 0.04$.

.....

4. Sur 650 électeurs interrogés au hasard, 54% sont favorables à la construction du stade. La municipalité décide-t-elle de construire le stade ?

.....

.....