

Fiche Exercice Échantillonnage :

1) Fluctuation d'échantillonnage

Exercice 1 :

Dans le monde, la proportion des gauchers est 12 % .
On suppose qu' il en est de même dans votre lycée.



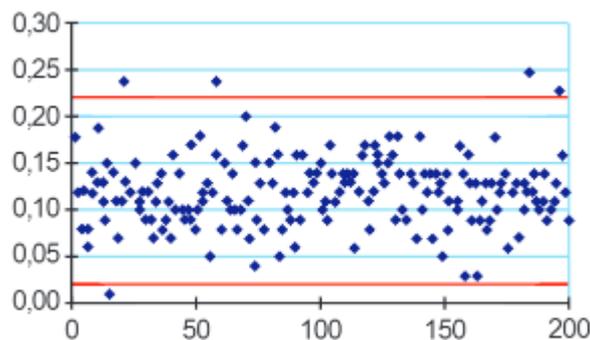
1) On simule des échantillons aléatoires de tailles 100 dans la population du lycée ; le graphique donne des fréquences des gauchers sur 200 échantillons.

a) Calculer les bornes de l'intervalle $I = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$

b) Quel est d'après le graphique le pourcentage d'échantillons fournissant une fréquence dans I ?

Ce résultat est il en accord avec la théorie ?

c) Si l'on simulait des échantillons dont la taille est celle de votre classe, quel serait l'intervalle de fluctuation au seuil 95 % ?



A ce seuil, entre quels entiers devrait se situer le nombre de gaucher de votre classe ? Vérifier.

Exercice 2 :

En France, on considère qu'environ 43 % de la population possède un groupe sanguin de type O. Un hôpital réalise une collecte de sang sur 750 personnes qui se présentent au hasard. Chaque personne se voit prélever 400 mL de sang.

1) Calculer à 10^{-2} près les bornes de l'intervalle de fluctuation au seuil 95 % de la fréquence des personnes du groupe O sur un tel échantillon.

2) En déduire, un intervalle de fluctuation à 95 % du nombre de donneurs de type O susceptibles d'y participer et du nombre de litres de sang de type O que l'hôpital peut s'attendre à collecter.

Exercice 3 :

En 1976, un accusé d'origine mexicaine condamné pour différents crimes au Texas attaqua le jugement sous le motif que la désignation des jurés était discriminatoire pour les américains d'origine mexicaine. Son argument était que ceux-ci n'étaient pas suffisamment représentés dans les jurys populaires.

A cette époque, dans l'état du Texas, 79,1 % de la population était mexico-américaine, et parmi les 870 personnes convoquées en tant que « grands jurés » (sur une période de 11 ans), 339 étaient d'origine mexicaine. La constitution des jurys pouvait-elle être considérée comme impartiale à l'époque dans cet état ?



2) Prise de décision

Exercice 4 :

On a lancé un dé 200 fois un dé cubique et on a obtenu les fréquences f_1, f_2, \dots, f_6 suivantes :

Face i	1	2	3	4	5	6
Fréquence f_i	0,195	0,1	0,19	0,205	0,16	0,15

1) Diriez vous que ce dé est bien équilibré ou pipé ?

2) On suppose le dé parfaitement équilibré.

a) Préciser la probabilité p de chaque face et calculer, à 10^{-2} près, les bornes de l'intervalle de fluctuation au seuil de 95 % de la fréquence f de chaque face du dé.



b) Diriez vous que les écarts entre les fréquences f_i observées et la probabilité théorique p s'expliquent par la seule fluctuation d'échantillonnage, due au hasard ? Êtes vous sûr d'avoir raison ?

Exercice 5 :

Deux entreprises A et B recrutent dans une région où il y a autant d'hommes que de femmes.

Dans l'entreprise A, il y a 100 employés dont 43 femmes.

Dans l'entreprise B, il y a 2500 employés dont 1150 femmes.



Quelle entreprise respecte le mieux la parité ?

Exercice 6 :

On réalise une analyse sanguine sur 1000 français choisis au hasard. On constate que 460 personnes sont du groupe sanguin O.

1) Comment peut-on estimer, à 95 %, la proportion de la population française qui est du groupe sanguin O ?

2) La population française est de 64 545 000 habitants. Grâce à cet intervalle trouvé au 1, donner un encadrement du nombre de français du groupe O dans toute la population française, avec un niveau de confiance de 95 %.

Exercice 7 :

Un horticulteur commande des bulbes de crocus. Le fournisseur garantit une composition de 50 % de crocus jaunes et 50 % de crocus violets.

L'horticulteur plante 625 bulbes ; il constate que tous les bulbes ont fleuri et que 290 fleurs sont jaunes.

1) Déterminer l'intervalle de fluctuation correspondant à cette situation.

2) Déterminer la fréquence observée.

3) Que peut en penser l'horticulteur ?

4) Si la proportion réelle de bulbes jaunes avait été de 46 % quel aurait été l'intervalle de fluctuation correspondant à cette nouvelle proportion théorique ? Que constate-t-on ?

Exercice 8 :

En 2006, un rapport de l' Inserm a déterminé que sur l'ensemble des personnes hospitalisées cette année-là, 4,97 % avaient contracté une maladie nosocomiale.

En 2007, dans l'hôpital d'une ville donnée, il y a eu 3742 personnes admises.

1) Si on suppose que le pourcentage des personnes hospitalisées déterminé en 2006 reste le même en 2007, calculer l'intervalle de fluctuation au seuil de 95 % correspondant à cet hôpital.

2) Sachant que 191 patients de cet hôpital ont contracté une maladie nosocomiale en 2007, l'hôpital doit-il prendre des mesures supplémentaires de prévention ?

Exercice Bonus :

[0,246 ; 0,557] est un intervalle de fluctuation au seuil de 95 %.

Quelle est la taille de l'échantillon ?

3) Intervalle de confiance (fourchette de sondage)

Exercice 9 :

Dans une entreprise de prêt à porter, les pièces sont découpées dans l'atelier de découpe et elles alimentent l'atelier de couture où elles sont assemblées.

Les articles ainsi obtenus subissent un contrôle à la sortie de cet atelier. Pour apprécier la qualité de la production, le contrôleur cherche à évaluer la proportion p d'articles non commercialisables. Pour cela il prélève au hasard des échantillons de taille n . Cette taille étant petite devant celle de la production totale, on assimile ce prélèvement à un tirage avec remise.

Le contrôleur prélève un échantillon de 250 articles et constate que 46 ne sont pas commercialisables.

Donner une estimation de p à l'aide de l'intervalle de confiance de niveau 0,95.

Exercice 10 :

Un institut de sondage réalise une étude auprès d'un échantillon de personnes pour estimer les intentions de votes pour 3 candidats A, B et C à une élection.

Malheureusement , une mauvaise manipulation informatique a fait disparaître un certain nombre de données juste avant leur publication. Il ne reste à l'institut que les chiffres suivants :

→ 39 % des personnes de l'échantillon déclarent vouloir voter pour A.

L'institut avait calculé l'intervalle de confiance à 95 % des intentions de votes pour A et avait trouvé [0,365 ; 0,415]

→ 28,5 % avaient déclaré vouloir voter pour B.

→ Aucun chiffre pour le candidat C.

1) Retrouver la taille de l'échantillon ainsi que les intervalles de confiance pour les candidats B et C.

2) Peut-on faire un classement de ces trois candidats ?

Exercice 11 : Problème ouvert

Au 1^{er} tour de l'élection présidentielle de 2002, les résultats sont : Jacques Chirac : 19,88 %, Lionel Jospin : 16,18 % et Jean Marie Le Pen : 16,86 %.

Le lendemain, 22 Avril 2002, on pouvait lire :

« L'éviction de Jospin laisse Le Pen face à Chirac ! Un séisme. Le duel entre Jacques Chirac et Lionel Jospin n'aura pas lieu ». (Source : Libération).

Le dernier sondage, effectué sur environ 1000 personnes, publié par BVA le 19/04/2002 donnait comme prévisions :

Jacques Chirac : 19 % , L. Jospin : 18 % et J.-M. Le Pen : 14 % !

Unanimement, les médias mettent en cause les sondages.

Mais ,à contre courant, un statisticien s'exprime ainsi dans le journal Le Monde :

« Pour les rares scientifiques qui savent comment sont produites les estimations, il est clair que l'écart des intentions de vote entre les candidats Le Pen et Jospin rendait tout à fait plausible le scénario qui s'est réalisé. »



Alors qui croire ?

Exercice 12 :

Un institut de sondage interroge 1052 personnes entre les deux tours de l'élection présidentielle sur leur intention de vote.

614 déclarent avoir l'intention de voter pour Martine Phinon.

En supposant que les votes seront conformes aux intentions, la candidate a-t-elle raison de croire qu'elle sera élue ?

Exercice 13 :

Un échantillon de 10 000 personnes sur une population étant donné, on sait que le taux moyen de personnes à soigner pour un problème de cholestérol élevé est de 7,5%. Donner un intervalle dans lequel on soit «sûr» à 95%, de trouver le nombre exact de personnes à soigner sur les 10 000.

Exercice 14 :

Un constructeur automobile fait appel à un institut de sondage afin de mesurer le degré de satisfaction du service après-vente.

L'institut souhaite estimer la proportion de clients satisfaits au niveau de confiance 0,95 avec une amplitude d'au plus 5 centièmes.

Combien de personnes au minimum faut-il interroger ?