

D.M. n°9 : Échantillonnage Estimation

2^{nde}

A rendre le lundi 13 mai 2013, Ce sujet est à rendre avec la copie.

Nom :	Communication : - ± +	Signature des parents : <i>Vu</i>	Note : <u>5</u>
Prénom :	Technique : - ± +	Raisonnement : - ± +	

Il faut toujours prouver vos affirmations (sauf mention contraire de l'énoncé).

Exercice 1. Estimation Source Transmath

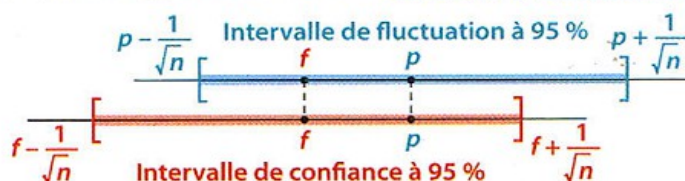
OBJECTIF. Estimer une proportion p inconnue à partir d'un échantillon de taille n .

1 Un encadrement de p

Dans une population, la proportion d'un caractère donné est égale à p . Nous avons dit que, pour un échantillon de taille $n \geq 25$ et pour p comprise entre 0,2 et 0,8, la fréquence observée f de ce caractère appartient à l'intervalle $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$ avec une probabilité d'au moins 0,95.

Puisque $f \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}} \right]$, alors $p - \frac{1}{\sqrt{n}} \leq f \leq p + \frac{1}{\sqrt{n}}$.

On en déduit que, lorsque $n \geq 25$, la proportion p est dans l'intervalle $\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right]$ avec une probabilité d'au moins 0,95. On appelle cet intervalle **l'intervalle de confiance** lié à la fréquence f .



2 Autour d'une élection

Lors du premier tour d'une élection, trois candidats A, B et C s'affrontaient. Les deux candidats placés en tête au 1^{er} tour participaient à un éventuel second tour. Deux jours avant le scrutin, était paru le sondage suivant :

Candidat	A	B	C
Résultat	30 %	28 %	24 %

Sur un échantillon de 1 000 personnes de 18 ans et plus

La simple lecture du sondage pouvait laisser présager un duel entre les candidats A et B au second tour.

a) Déterminez l'intervalle de confiance à 95 % pour chacun des candidats. (Les bornes seront données à 0,001 près.)

b) Représentez sur une même droite graduée ces trois intervalles.



c) Le résultat du 1^{er} tour ne confirma pas le sondage :

Candidat	A	B	C
Résultat	29,9 %	26,2 %	26,9 %

Selon vous, faut-il remettre en question la validité du sondage réalisé avant l'élection ? Argumentez votre réponse à l'aide des questions précédentes.

SYNTHÈSE. Dans une population, on désire estimer la proportion inconnue p d'un caractère donné.

On étudie un échantillon de taille n , $n \geq 25$. Le caractère étudié apparaît avec la fréquence f ($0,2 \leq f \leq 0,8$).

On peut estimer que la proportion p du caractère dans la population totale est dans **l'intervalle de confiance** $\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right]$ avec une probabilité d'au moins 0,95.

Attention. Dans 5 % environ des cas, l'intervalle de confiance ne contient peut-être pas p .

Exercice 2. Simulation pour obtenir un intervalle de confiance

Partie I : Simulation

Contexte : En avril 2002 le candidat Jean-Marie Le Pen a surpris tout le monde en battant Lionel Jospin au premier tour des élections présidentielles en France. Une semaine avant le premier tour, un sondage donnait 14% d'intentions de vote pour J-M Le Pen et 18% pour L. Jospin. Soient p_1 et p_2 les proportions de votants dans l'ensemble de l'électorat ayant l'intention de voter pour Le Pen et Jospin respectivement. p_1 et p_2 sont inconnues. On cherche à les évaluer par un sondage sur 1000 personnes. La méthode qui nous permet habituellement de déterminer par le calcul un intervalle de confiance au seuil de 95% pour une proportion p n'étant pas applicable dans ce cas, on souhaite réaliser une simulation pour pouvoir déterminer un intervalle de confiance pour chacune des proportions p_1 et p_2 .

1) Dans ce but, écrire (sur votre copie) un programme qui simule 1000 expériences d'un jeu où on a 14% de chances de gagner à chaque expérience et qui compte (puis affiche) la proportion de fois où l'on a gagné sur ces 1000 parties.

Le programme peut être écrit en langage naturel ou en langage Algobox ou en langage TI ou en langage Casio.

2) Grâce à votre programme, vous allez alors simuler 10 fois de suite une série de 1000 expériences et noter la fréquence d'apparition de pile dans le tableau :

Série de 1000 expériences	1	2	3	4	5	6	7	8	9	10
f										

Il faudrait en fait non pas 10 mais 100 séries de 1000 expériences. Cependant, à moins de se répartir la tâche, ce serait un peu long.

On pourrait aussi améliorer le programme pour qu'il refasse automatiquement 100 fois ces 1000 expériences et note les résultats dans une liste mais c'est peut-être encore un peu compliqué pour certains d'entre vous.

Bien sûr si vous vous sentez de taille à obtenir les résultats pour 100 séries de 1000 expériences par l'une ou l'autre de ces méthodes, faites-le !

Les autres se contenteront d'utiliser les résultats de ces 100 simulations qui figurent dans l'exercice 60 p 155.

Partie II : Utilisation de la simulation : Ex 60 p155.