

Stage statistique – probabilités

TI graphiques (83 Premium CE & Advanced)

Intervalles de confiance

Énoncé : Lors des élections présidentielles de 2012, une nouvelle polémique sur la fiabilité des sondages est apparue, rappelant celle des présidentielles de 2002.

En effet, de nombreux observateurs se sont interrogés sur les résultats des derniers sondages avant le premier tour des élections et ceux du vote.

Ainsi, considérons le sondage IPSOS pour France Télévisions, France Inter et Le Monde du 20 avril 2012. Pour ce sondage, 1 021 personnes ont été interrogées par la méthode des quotas¹.

Si les résultats du sondage et du vote pour François Hollande, Eva Joly, François Bayrou et les « petits candidats » semblent cohérents, on note des différences importantes pour Nicolas Sarkozy, Marine Le Pen et Jean-Luc Mélenchon :

| | Nicolas Sarkozy | Marine Le Pen | Jean-Luc Mélenchon |
|------------------|-----------------|---------------|--------------------|
| Sondage IPSOS | 25,5 % | 16 % | 14 % |
| Vote du 22 avril | 27,2 % | 17,9 % | 11,1 % |

À la télévision, le représentant d'IPSOS a justifié ces différences par la « marge d'erreur » des sondages. Sa réponse est-elle acceptable ?

Nous allons déterminer des intervalles de confiance au seuil de confiance de 95 % des résultats de ce sondage.

Partie 1 : Des programmes pour les intervalles de confiance

Rappels : Déterminer un intervalle de confiance, c'est estimer par intervalle une proportion inconnue p à partir d'un échantillon.

- Si f désigne la fréquence observée sur un échantillon de taille n , avec $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$, alors p est élément de l'intervalle $\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}} \right]$ avec un niveau de confiance de plus de 95 %.

Pour une meilleure précision, sous les mêmes conditions, on estime que p appartient à l'intervalle² $\left[f - 1,96 \frac{\sqrt{f(1-f)}}{\sqrt{n}} ; f + 1,96 \frac{\sqrt{f(1-f)}}{\sqrt{n}} \right]$ avec un niveau de confiance de plus de 95 %.

a) En utilisant la première formule, écrire un programme, nommé **CONFIANC**, permettant de calculer les extrémités de l'intervalle de confiance et représenter cet intervalle.

b) Opérer, de même, avec la deuxième formule. On nommera le programme **CONFIAN1**.

Partie 2 : Une formule simple pour l'intervalle de confiance

¹ La méthode des quotas s'appuie sur la répartition connue de la population des électeurs (âge, sexe, situation géographique, catégorie socio-professionnelle, etc.).

² Cet intervalle de confiance est construit à partir de l'intervalle de fluctuation asymptotique (voir programme).

a) Vérifier que les conditions du sondage vérifient les conditions usuelles de précision.

.....

b) Déterminer, avec le programme **CONFIANC**, les encadrements de chacune des proportions concernant les votes des trois candidats cités, à partir des fréquences observées dans le sondage IPSOS.

.....

.....

.....

.....

c) Ces résultats permettent-ils d'affirmer que les « marges d'erreur » justifient les différences entre le sondage et la réalité du vote ?

.....

.....

Partie 3 : Une autre formule pour l'intervalle de confiance

a) Déterminer, avec le programme **CONFIAN1**, les encadrements de chacune des proportions concernant les votes des trois candidats cités, à partir des fréquences observées dans le sondage IPSOS.

.....

.....

.....

.....

b) Ces résultats permettent-ils d'affirmer que les « marges d'erreur » justifient les différences entre le sondage et la réalité du vote ?

.....

.....

c) Comment apprécier ces différences ?

.....

.....

.....

.....

.....

.....