

Intervalle de fluctuation

1 Loi binomiale

Au sein d'une population, on suppose que la proportion d'un certain caractère est p . On souhaite juger cette hypothèse. Pour cela on prélève dans la population au hasard et avec remise, un échantillon de taille n sur lequel on observe une fréquence f de ce caractère.

Définition:

L'intervalle de fluctuation au seuil de $1 - \alpha$ d'une fréquence, sur un échantillon aléatoire de taille n , selon la loi binomiale de paramètres n et p est :

$$\left[\frac{a}{n}; \frac{b}{n} \right]$$

avec a le plus petit entier tel que $P(X \leq a) > \frac{\alpha}{2}$ et b le plus petit entier tel que $P(X \leq b) \geq 1 - \frac{\alpha}{2}$.

Règle:

- Si $f \in \left[\frac{a}{n}; \frac{b}{n} \right]$, l'hypothèse est acceptée au seuil de $1 - \alpha$;
- Si $f \notin \left[\frac{a}{n}; \frac{b}{n} \right]$, l'hypothèse est rejetée au seuil de $1 - \alpha$.

Exercice 1:

Voici les résultats du premier tour de l'élection présidentielle de 2012 (le 22 avril 2012).

Nom	HOLLANDE	SARKOZY	LE PEN	MELENCHON
Prénom	Francois	Nicolas	Marine	Jean-Luc
Partie	PS	UMP	FN	FG
% des voix exprimés	28,63%	27,18%	17,9%	11,1%

Voici les résultats de sondages effectués avant le premier tour de l'élection présidentiel sur un panel représentatif de 1.000 électeurs. Chaque sondage correspond ainsi à un échantillon de 1.000 personnes.

Nom	HOLLANDE	SARKOZY	LE PEN	MELENCHON
Prénom	François	Nicolas	Marine	Jean-Luc
13 mars - Sofres	30%	26%	16%	11,5%
16 mars - OPINIONWAY	27,5%	27,5%	16%	13%
16 mars - IFOP	26,5%	27,5%	17%	13%

1. Déterminer, en utilisant la loi binomiale, l'intervalle de fluctuation au seuil de 95% pour chaque candidat.
2. Déterminer les échantillons que l'on peut valider au seuil de 95%.
3. Conclure.

2 Théorème de Moivre-Laplace

On rappelle ci-dessous l'énoncé du théorème de Moivre-Laplace :

Théorème:

Soit, pour tout entier n , une variable aléatoire X_n qui suit une loi binomiale $\mathcal{B}(n, p)$ et $Z_n = \frac{X_n - np}{\sqrt{np(1-p)}}$, la variable centrée réduite associée à X_n . Alors pour tous nombres a et b tels que $a \leq b$,

$$\lim_{n \rightarrow +\infty} P(a \leq Z_n \leq b) = P(a \leq Z \leq b)$$

où Z suit la loi normale centrée réduite $\mathcal{N}(0, 1)$.

On considère à présent la variable aléatoire Z qui suit la loi normale centrée réduite $\mathcal{N}(0, 1)$ et $\alpha \in]0; 1[$. On a démontré précédemment qu'il existe un unique réel u_α tel que $P(-u_\alpha \leq Z \leq u_\alpha) = 1 - \alpha$.

Exercice 2:

Démontrer que $\lim_{n \rightarrow +\infty} P\left(\frac{X_n}{n} \in I_n\right) = 1 - \alpha$ où $I_n = \left[p - u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}; p + u_\alpha \frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$

Définition:

Sous les conditions énoncées dans le théorème ci-dessus, on dit que l'intervalle I_n est un intervalle de fluctuation asymptotique de la variable aléatoire fréquence $F_n = \frac{X_n}{n}$ au seuil de $1 - \alpha$. Ainsi, pour n grand, la variable aléatoire F_n prend ses valeurs dans l'intervalle I_n avec une probabilité proche de $1 - \alpha$. En pratique, on utilise cet intervalle de fluctuation de F_n dès que $n \geq 30$; $np \geq 5$ et $n(1-p) \geq 5$.

Exercice 3:

Dans l'exercice 1, déterminer les intervalles de fluctuation asymptotique au seuil de 0,95 et au seuil de 0,99 pour chaque candidat. Conclure.

Exercice 4:

Une compagnie aérienne possède des avions d'une capacité de 300 places. Cette compagnie a vendu n billets pour le vol 747 en direction de Sydney. La probabilité pour qu'un acheteur se présente à l'embarquement est p et les comportements des acheteurs sont indépendants les uns des autres. On note X_n la variable aléatoire désignant le nombre d'acheteurs d'un billet se présentant à l'embarquement. La compagnie cherche à optimiser le remplissage de l'avion en vendant éventuellement plus de places que la capacité totale de l'avion (surréservation ou surbooking) soit ici $n > 300$. Comme il y a évidemment un risque que le nombre de passagers munis d'un billet se présentant à l'embarquement excède 300, la compagnie veut maîtriser ce risque.

1. Déterminer la loi de X_n .

2. On suppose $p \in [0, 5; 0, 95]$. Écrire l'intervalle de fluctuation asymptotique I_n de $\frac{X_n}{n}$ au seuil de 0,95.

3. Montrer que si $I_n \subset \left[0; \frac{300}{n}\right]$ alors la probabilité que le nombre de passagers se présentant à l'embarquement excède 300 est proche de 0,05.

4. On cherche à déterminer la valeur de n maximale permettant de satisfaire la condition de l'inclusion $I_n \subset \left[0; \frac{300}{n}\right]$.

a. Montrer que $I_n \subset \left[0; \frac{300}{n}\right] \implies pn + 1,96\sqrt{n}\sqrt{p(1-p)} - 300 \leq 0$.

b. On pose $f(x) = px + 1,96\sqrt{x}\sqrt{p(1-p)} - 300$. Montrer qu'il existe un entier n_0 tel que si $n \leq n_0$ alors $f(n) \leq 0$ et si $n > n_0$ alors $f(n) > 0$.

c. Tracer la courbe représentative de f pour les valeurs $p = 0,85$; $p = 0,9$; $p = 0,95$.

d. Déterminer à la calculatrice les valeurs de n_0 pour $p = 0,85$; $p = 0,9$; $p = 0,95$.