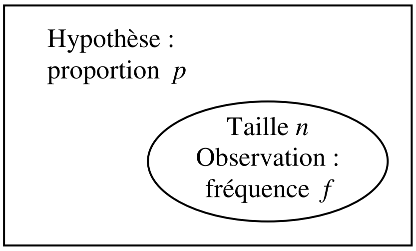
Intervalle de fluctuation à 95 % d'une fréquence et loi binomiale

Ce texte précise le contenu « *Utilisation de la loi binomiale pour une prise de décision à partir d'une fréquence* » et la capacité correspondante, « *Exploiter l'intervalle de fluctuation à un seuil donné, déterminé à l'aide de la loi binomiale, pour rejeter ou non une hypothèse sur une proportion* », des programmes du lycée.



On considère une population dans laquelle on suppose que la proportion d'un certain caractère est p. Pour juger de cette hypothèse, on y prélève, au hasard et avec remise, un échantillon de taille n sur lequel on observe une fréquence f du caractère.
On rejette l'hypothèse selon laquelle la proportion dans la population est p lorsque la fréquence f observée est trop éloignée de p, dans un sens ou dans l'autre. On choisit de fixer le seuil de décision de sorte que la probabilité de rejeter l'hypothèse, alors qu'elle est vraie, soit inférieure à 5 %.

Lorsque la proportion dans la population vaut p, la variable aléatoire X correspondant au nombre de fois où le caractère est observé dans un échantillon aléatoire de taille n, suit la loi binomiale de paramètres n et p. On cherche à partager l'intervalle $[0, n]$, où X prend ses valeurs, en trois intervalles $[0, a - 1]$, $[a, b]$ et $[b + 1, n]$ de sorte que X prenne ses valeurs dans chacun des intervalles extrêmes avec une probabilité proche de 0,025, sans dépasser cette valeur.
En tabulant les probabilités cumulées $P(X \leq k)$, pour k allant de 0 à n, il suffit de déterminer le plus petit entier a tel que $P(X \leq a) > 0{,}025$ et le plus petit entier b tel que $P(X \leq b) \geq 0{,}975$.
La règle de décision est la suivante : si la fréquence observée f appartient à l'intervalle de fluctuation à 95 % $\left[\frac{a}{n}, \frac{b}{n}\right]$, on considère que l'hypothèse selon laquelle la proportion est p dans la population n'est pas remise en question et on l'accepte ; sinon, on rejette l'hypothèse selon laquelle cette proportion vaut p.

Définition : l'intervalle de fluctuation à 95 % d'une fréquence correspondant à la réalisation, sur un échantillon aléatoire de taille n, d'une variable aléatoire X de loi binomiale, est l'intervalle $\left[\frac{a}{n}, \frac{b}{n}\right]$ défini par :

- a est le plus petit entier tel que $P(X \leq a) > 0{,}025$;
- b est le plus petit entier tel que $P(X \leq b) \geq 0{,}975$.

Pour $n \geq 30$, $n \times p \geq 5$ et $n \times (1 - p) \geq 5$, on observe que l'intervalle de fluctuation $\left[\frac{a}{n}, \frac{b}{n}\right]$ est sensiblement le même que l'intervalle $\left[p - \frac{1}{\sqrt{n}},\ p + \frac{1}{\sqrt{n}}\right]$ proposé dans le programme de seconde.

Exemple d'exercice

Monsieur Z, chef du gouvernement d'un pays lointain, affirme que 52 % des électeurs lui font confiance. On interroge 100 électeurs au hasard (la population est suffisamment grande pour considérer qu'il s'agit de tirages avec remise) et on souhaite savoir à partir de quelles fréquences, au seuil de 5 %, on peut mettre en doute le pourcentage annoncé par Monsieur Z, dans un sens, ou dans l'autre.

1. On fait l'hypothèse que Monsieur Z dit vrai et que la proportion des électeurs qui lui font confiance dans la population est $p = 0{,}52$. Montrer que la variable aléatoire X, correspondant au nombre d'électeurs lui faisant confiance dans un échantillon de 100 électeurs, suit la loi binomiale de paramètres $n = 100$ et $p = 0{,}52$.

2. On donne ci-contre un extrait de la table des probabilités cumulées $P(X \leq k)$ où X suit la loi binomiale de paramètres $n = 100$ et $p = 0{,}52$.

k	$P(X \leq k) \approx$
40	0,0106
41	0,0177
42	0,0286
43	0,0444
...	...
61	0,9719
62	0,9827
63	0,9897
64	0,9941

a. Déterminer a et b tels que :

- a est le plus petit entier tel que $P(X \leq a) > 0{,}025$;
- b est le plus petit entier tel que $P(X \leq b) \geq 0{,}975$.

b. Comparer l'intervalle de fluctuation à 95 %, $\left[\frac{a}{n}, \frac{b}{n}\right]$, ainsi obtenu grâce à la loi binomiale, avec l'intervalle $\left[p - \frac{1}{\sqrt{n}},\ p + \frac{1}{\sqrt{n}}\right]$.

3. Énoncer la règle décision permettant de rejeter ou non l'hypothèse $p = 0{,}52$, selon la valeur de la fréquence f des électeurs favorables à Monsieur Z obtenue sur l'échantillon.

4. Sur les 100 électeurs interrogés au hasard, 43 déclarent avoir confiance en Monsieur Z. Peut-on considérer, au seuil de 5 %, l'affirmation de Monsieur Z comme exacte ?

Éléments de réponse

2. a. On lit $a = 42$ et $b = 62$.

b. Les intervalles sont identiques.

3. Si f appartient à l'intervalle [0,42 ; 0,62], l'hypothèse $p = 0{,}52$ est acceptable, sinon, l'hypothèse $p = 0{,}52$ est rejetée, au seuil de 5 %.

4. On considère que l'affirmation de Monsieur Z est exacte.

Remarque : la recherche de l'intervalle de fluctuation peut-être illustrée par le diagramme en bâton de la loi binomiale de paramètres $n = 100$ et $p = 0{,}52$.