Chapitre 13 : Échantillonnage avec la loi binomiale

I Intervalle de fluctuation vu en seconde

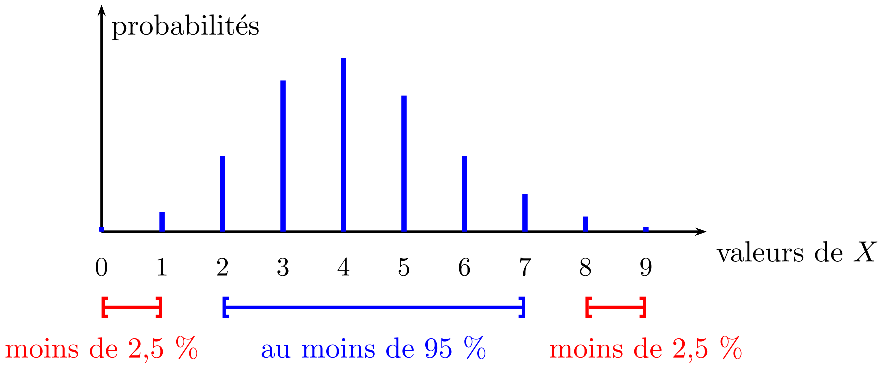
Propriété (intervalle de fluctuation d'une fréquence)
Soit un caractère dont la proportion dans une population donnée est p.
Lorsque $n \geqslant 25$ et $0,2 \leqslant p \leqslant 0,8$, au moins 95 % des échantillons de taille n issus de cette population sont tels que la fréquence f du caractère dans l'échantillon appartient à l'intervalle $\left[p - \dfrac{1}{\sqrt{n}}; p + \dfrac{1}{\sqrt{n}}\right]$.
L'intervalle $I = \left[p - \dfrac{1}{\sqrt{n}}; p + \dfrac{1}{\sqrt{n}}\right]$ est appelé intervalle de fluctuation au seuil de 95 %.

II Intervalle de fluctuation avec la loi binomiale

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.

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 paramétres n et p.

On partage l'intervalle $[0; n]$ en trois intervalles : $[0; a-1]$, $[a; b]$ et $[b+1; n]$ de façon que X prenne ses valeurs dans les deux intervalles extrêmes avec une probabilité proche de 2,5 % mais sans jamais la dépasser.
Ainsi, X prend ses valeurs dans $[a; b]$ avec une probabilité d'au moins 95 %.



Définition
Soit X une variable aléatoire suivant la loi binomiale $\mathscr{B}(n; p)$.
On note f la fréquence associée à un échantillon aléatoire de taille n de X.
L'intervalle de fluctuation au seuil de 95 % de la fréquence f est l'intervalle $\left[\dfrac{a}{n}; \dfrac{b}{n}\right]$, défini par :

- a est le plus petit entier tel que $P(X \leqslant a) > 0,025$,
- b est le plus petit entier tel que $P(X \leqslant b) \geqslant 0,975$.

Remarque
Comme les valeurs prises par X sont des nombres entiers (X suit $\mathscr{B}(n;p)$, donc les valeurs de X vont de 0 à n), on a aussi :

— a est le plus grand entier tel que $P(X < a) \leqslant 0,025$,

— b est le plus petit entier tel que $P(X > b) \leqslant 0,025$

On peut noter par exemple que $P(X \leqslant a) = P(X < a+1)$.

Exercice 1
Déterminer l'intervalle de fluctuation à 95 % d'une variable aléatoire suivant une loi binomiale : ressource 3943

Remarque
Cette propriété s'applique à des échantillons de variables aléatoires suivant une loi binomiale.
Elle est vraie quelles que soient les valeurs de n et p, contrairement à l'intervalle de fluctuation vu en seconde.
L'intervalle de fluctuation $\left[\frac{a}{n};\frac{b}{n}\right]$ n'est pas nécessairement centré sur p.
Il n'y a pas de formule donnant directement les bornes $\frac{a}{n}$ et $\frac{b}{n}$.

III Prise de décision avec la loi binomiale

Théorème (prise de décision)
On fait une hypothèse sur la valeur de p.

— Si la fréquence observée f n'appartient pas à l'intervalle de fluctuation au seuil de 95 % $\left[\frac{a}{n};\frac{b}{n}\right]$, alors on rejette l'hypothèse faite sur p, avec un risque de 5 % (de rejeter à tort une hypothèse vraie).

— si la fréquence observée f appartient à l'intervalle de fluctuation $\left[\frac{a}{n};\frac{b}{n}\right]$, alors on ne peut pas rejeter l'hypothèse faite sur p.

Remarque
Dans le cas où la fréquence f observée sur l'échantillon est dans l'intervalle de fluctuation, on ne peut pas rejeter l'hypothèse. Dans ce cas, on ne peut pas quantifier le risque d'erreur. L'hypothèse que l'on a acceptée est une hypothèse comme une autre.

Exercice 2
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 95 %, 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 $0,52$. 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 $P(X \leqslant k)$ où X suit la loi binomiale de paramètres $n = 100$ et $p = 0,52$.

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

Déterminer a et b tels que :
a est le plus petit entier tel que $P(X \leqslant a) > 0,025$;
b est le plus petit entier tel que $P(X \leqslant b) \geqslant 0,975$.

3. En déduire l'intervalle de fluctuation $\left[\frac{a}{n} ; \frac{b}{n}\right]$ au seuil de 95 % pour une fréquence f de personnes lui faisant confiance.
4. Comparer cet intervalle à l'intervalle de fluctuation $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$ vu en seconde.
5. Sur les 100 électeurs interrogés au hasard, 43 déclarent avoir confiance en Monsieur Z. Peut-on considérer, au seuil de 95 %, l'affirmation de Monsieur Z comme exacte ?

Exercice 3
Un laboratoire annonce qu'un médicament sauve 40 % des patients atteints d'une maladie rare.
Pour contrôler cette affirmation, on le teste sur 100 patients atteints de cette maladie.
Sur les 100 malades auxquels on a administré le médicament, on en a sauvé 30.
Au seuil de 5 %, que dire de l'annonce du laboratoire ?

Exercice 4 (sécurité au carrefour)
Un groupe de citoyens demande à la municipalité d'une ville la modification d'un carrefour en affirmant que 40 % des automobilistes tournent en utilisant une mauvaise file.
Un officier de police constate que sur 500 voitures prises au hasard, 190 prennent une mauvaise file.

1. Déterminer, en utilisant la loi binomiale sous l'hypothèse $p = 0,4$, l'intervalle de fluctuation au seuil de 95 %.
2. D'après l'échantillon, peut-on considérer, au seuil de 95 %, comme exacte l'affirmation du groupe de citoyens ?