Échantillonnage

Classe de première STMG - Lycée Saint-Charles

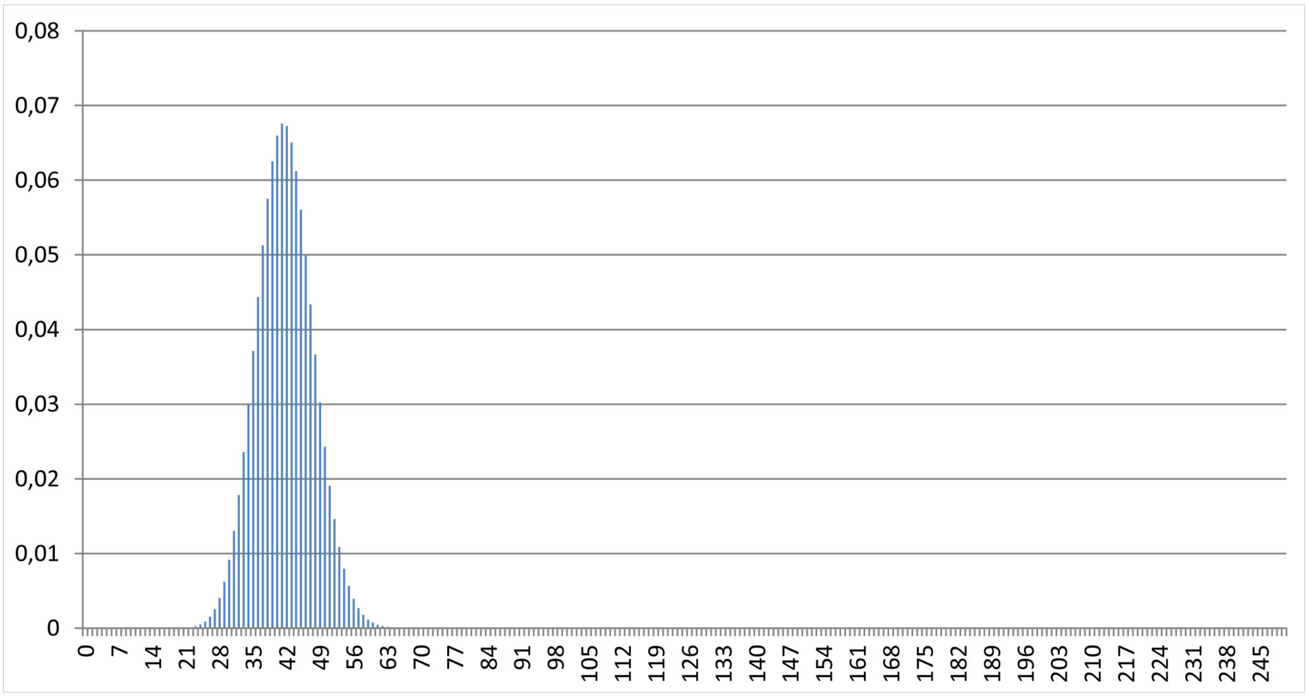
Patrice Jacquet - www.mathxy.fr - 2015

Objectif :

- Savoir décider, avec une marge d'erreur de 5%, à partir d'un échantillon d'une population, si une hypothèse sur une proportion est acceptable ou non.

Exemple – En lançant un dé 250 fois, on obtient 31 fois la valeur 6. Peut-on dire (avec une marge d'erreur de 5%) que le dé est bien équilibré ?

Remarque – La loi de probabilité de la variable aléatoire X qui compte le nombre de 6 est une loi binomiale de paramètres 250 et $\frac{1}{6}$.
La loi binomiale $\mathcal{B}\left(250; \frac{1}{6}\right)$ est représentée par le diagramme en bâtons ci-dessous (obtenu avec un tableur) :



Avec le tableur on calcule :
$P(X \leqslant 31) \approx 0,0385 = 3,85\%$ $\quad P(X \leqslant 30) \approx 0,0255 = 2,55\%$ $\quad P(X \leqslant 29) \approx 0,0163 = 1,63\%$
30 est le plus petit nombre a tel que $P(X \leqslant a) > 0,025$
$P(X \geqslant 53) \approx 0,0360 = 3,60\%$ $\quad P(X \geqslant 54) \approx 0,0251 = 2,51\%$ $\quad P(X \geqslant 55) \approx 0,0171 = 1,71\%$
54 est le plus petit nombre b tel que $P(X \geqslant b) > 0,025$

Conclusion : le nombre de 6 appartient à l'intervalle $[30; 54]$ dans au moins 95% des cas. Donc, on peut affirmer (avec une marge d'erreur de 5%) que le dé est bien équilibré.

Définition 1 – intervalle de fluctuation

Soit une loi binomiale de paramètres n et p.
Pour cette valeur de p, on appelle intervalle de fluctuation à 95% de la fréquence f pour un échantillon de taille n, l'intervalle $\left[\frac{a}{n}; \frac{b}{n}\right]$, où a est le plus petit nombre entier tel que $P(X \leqslant a) > 0,025$ et b est le plus petit nombre entier tel que $P(X \geqslant b) > 0,025$.

Propriété 1

Lorsque $f \in \left[\frac{a}{n}; \frac{b}{n}\right]$ on accepte, au risque 5%, l'hypothèse « la proportion a pour valeur p ».

Exemple – En lançant un dé 250 fois, on obtient 31 fois la valeur 6.

Dans notre exemple, on a :

$n = 250 \qquad p = \frac{1}{6}$

$a = 30 \qquad b = 54 \qquad$ (calculé avec le tableur).

$\frac{a}{n} = 0,12 \qquad \frac{b}{n} = 0,216$

$f = \frac{31}{250} = 0,124$

$0,124 \in [0,12; 0,216]$ donc on accepte l'hypothèse, au risque 5%.