# Rappels sur les intervalles de fluctuation et intervalles de confiance

## I Échantillon statistique

**Définition**

Un échantillon de taille n est constitué des résultats de n répétitions indépendantes de la même expérience sur l'ensemble des personnes ou objets sur lesquels porte l'étude statistique (la population). Un échantillon issu d'une population est donc l'ensemble de quelques éléments de cette population.

## II Intervalle de fluctuation

### II.1 Pile ou Face

On lance une pièce de monnaie un certain nombre de fois et on note la fréquence de « Pile »; on attend une fréquence théorique $f = 0,5$.
En pratique, on obtient des fréquences différentes de 0,5, qui sont néanmoins « proches » de 0,5.
On observe que plus le nombre de cancers est grand, plus les fréquences sont roches de 0,5.

Voir simulation sur Excel.

**Exemple** : On suppose que 22 % des cartes à puce produites par une entreprise sont défectueuses. La proportion théorique $p$ de cartes défectueuses est donc égale à 22 %.
On prélève un échantillon de taille 200 parmi cette production et on compte le nombre de cartes à puce défectueuses parmi cet échantillon. Ce nombre est égal à 41.
Dans ce cas, la fréquence **observée** $f$ est égale à $\dfrac{41}{200} = 0,205$.
Pour un échantillon de taille 200, l'intervalle de fluctuation de la fréquence p des cartes à puce défectueuses au seuil de 95 %, est un intervalle de centre 0,22 tel que les fréquences observées se trouvent dans cet intervalle pour 95 % % des échantillons de taille 200.

**Définition**

L'intervalle de fluctuation au seuil de 95 % d'une fréquence d'un échantillon de taille $n$ est l'intervalle centré autour de la proportion théorique $p$ tel que la fréquence observée $f$ se trouve dans l'intervalle avec une probabilité égale à 0,95.

**Propriété**

Pour $0,2 < p < 0,8$ et $n > 25$, l'intervalle de fluctuation au seuil de 95 % de $f$ est l'intervalle
$\left[p - \dfrac{1}{\sqrt{n}}\ ;\ p + \dfrac{1}{\sqrt{n}}\right]$.
L'amplitude de cet intervalle est $\dfrac{2}{\sqrt{n}}$

**Interprétation**

Cela signifie qu'on a une probabilité de 0,95 pour que la fréquence observée se trouve dans l'intervalle $\left[p-\frac{1}{\sqrt{n}}\ ;\ p+\frac{1}{\sqrt{n}}\right]$

Dans l'exemple précédent, l'intervalle de fluctuation au seuil de 95 % de $p=0,22$ est $\left[0,22-\frac{1}{\sqrt{200}}\ ;\ 0,22+\frac{1}{\sqrt{200}}\right]$ soit de façon approchée [0,15 ; 0,29].

## III Estimation d'une proportion, intervalle de confiance

On suppose désormais que l'on ne connaît pas la proportion réelle $p$ du caractère et l'on cherche à l'estimer.

**Propriété**

On suppose que $p \in [0,2\ ;\ 0,8]$ (mais on ne connaît pas sa valeur) et que $n > 25$.
Parmi tous les échantillons de taille $n$ possibles ayant comme fréquence observée $f$, 95 % des intervalles associés $\left[f-\frac{1}{\sqrt{n}}\ ;\ f+\frac{1}{\sqrt{n}}\right]$ contiennent $p$.

**Définition**

L'intervalle $\left[f-\frac{1}{\sqrt{n}}\ ;\ f+\frac{1}{\sqrt{n}}\right]$ est dit intervalle de confiance de $p$ au niveau 0,95

**Exemple :** Le parti d'un candidat commande un sondage réalisé à partir de 1 600 personnes à l'issue duquel il est donné gagnant avec 52 % des voix.
A-t-il des raisons d'être confiant ?

**Réponse :**
Soit $p$ la proportion de gens votant pour lui ; il est élu si $p \geqslant 0,5$.
D'près l'intervalle de confiance, $p \in \left[f-\frac{1}{\sqrt{1\,600}}\ ;\ f+\frac{1}{\sqrt{1\,600}}\right] = [0,495\ ;\ 0,545]$ au seuil de 95%, donc il n'est pas sûr d'être élu.