.1 Echantillonnage dans une population.

On connaît la proportion p d'apparition d'un caractère dans une population.

Alors, si on prélève des échantillons de taille n dans la population, environ 95 % des fréquences d'apparition (f) du caractère dans chaque échantillon seront dans l'intervalle $\boxed{\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}}\right]}$.

C'est ce qu'on appelle **l'intervalle de fluctuation asymptotique à au moins 95 % de la fréquence f**.

(Cela est valable si on a les 3 conditions : $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$.)

On dit "asymptotique" car : plus n (la taille des échantillons) est grand, plus l'intervalle est petit.

Exemple d'utilisation :
Dans une maternité, on sait qu'il nait en moyenne 51 % de garçons. On fait le point sur la proportion de garçons toutes les 100 naissances.

Déterminer l'intervalle de fluctuation asymptotique à au moins 95 % de la fréquence des garçons dans les échantillons de taille 100.

ici, les echantillons sont de taille 100 → $n = 100$
on connait la frequence de la population $p = 51\% = 0,51$
l'intervalle de fluctuation est donc $\left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right] = \left[0,51 - \frac{1}{\sqrt{100}} ; 0,51 + \frac{1}{\sqrt{100}}\right]$
$= [0,5 ; 0,52]$

.2 Test d'hypothèse, prise de décision à partir d'un échantillon.

On ne connaît pas la proportion p d'apparition d'un caractère dans une population.

mais on fait l'hypothèse que p a une certaine valeur, et
on se demande si on peut accepter cette hypothèse ou non.

On prélève un échantillon de taille n dans la population, et on calcule la fréquence d'apparition du caractère dans l'échantillon : f.
L'intervalle de fluctuation est $I = \boxed{\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}}\right]}$.

Alors **la règle de décision est la suivante** :

- Si $f \in I$, alors on accepte l'hypothèse que p a bien la valeur supposée, "au seuil de confiance de 95 %".
- Si $f \notin I$, alors rejette l'hypothèse que p a bien la valeur supposée, "au seuil de confiance de 95 %".

(Cette méthode est valable toujours sous les 3 conditions : $n \geqslant 30$, $np \geqslant 5$ et $n(1-p) \geqslant 5$).

Exemple d'utilisation : On fait l'hypothèse que 40 % des individus sont allergiques à un médicament. On réalise un échantillon de taille 200 et on observe que 58 personnes sont allergiques au médicament.

1. Déterminer un intervalle de fluctuation asymptotique I à au moins 95 % de la fréquence des personnes

n = taille de l'echantillon = 200
p supposee = $40\% = 0,4$

alors $I = \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$
$= \left[0,4 - \frac{1}{200} ; 0,4 + \frac{1}{200}\right] = [0,32 ; 0,48]$

allergiques dans les échantillons de taille 200.

2. Enoncer la règle de décision permettant d'accepter ou non l'hypothèse $p = 0,4$. Dire alors si on peut accepter l'hypothèse.

on a un échantillon qui donne $f = \frac{58}{200} = 0,29$

si $f \in I$, on accepte l'hypothèse $p = 0,4$, sinon, on la rejette.

ici $0,29 \notin [0,32 ; 0,48]$ donc on rejette l'hypothèse $p = 0,4$.

.3 Estimation d'une proportion p inconnue : intervalle de confiance.

On ne connaît pas la proportion p d'apparition d'un caractère dans une population.

mais on connaît sa fréquence f d'appartition dans un échantillon de taille n.

On souhaite alors donner un intervalle dans lequel p a 95% de chance de se trouver.

L'intervalle est alors $\boxed{\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right]}$.

C'est ce qu'on appelle **l'intervalle de confiance au niveau de confiance 95%.**

(Cette méthode est valable dès que n est suffisamment grand).

Ici encore, plus l'échantillon sera grand (n grand), plus l'intervalle sera petit.
On arrondit par défaut la borne inférieure de l'intervalle, et par excès la borne supérieure.

Exemple d'utilisation : Une entreprise souhaite estimer la proportion p de clients satisfaits dans l'ensemble de ses clients. Pour cela, elle fait un sondage auprès d'un échantillon de 500 clients. 410 d'entre eux se disent satisfaits.

$n = 500$

1. Estimer la proportion p de clients satisfaits dans l'ensemble de la clientèle par un intervalle de confiance au niveau de confiance 95 %. (arrondir les bornes à 0,01 près.)

on ne connait pas p. on connait la fréquence de l'échantillon

$f = \frac{410}{500} = 0,82$. L'intervalle de confiance est $\left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right]$

$= \left[0,82 - \frac{1}{\sqrt{500}} ; 0,82 + \frac{1}{\sqrt{500}}\right] = [0,77 ; 0,87]$

2. Est-il possible que $p = 0,75$?

$p = 0,75$ n'est pas dans cet intervalle.

Mais il est possible que $p = 0,75$, (à 5% de chances)

Car cet intervalle a 95% de chances de contenir p.