I. Contexte :

Dans une population on veut **estimer** la proportion p d'un caractère.

Exemple : Avant la commercialisation d'un médicament, on veut estimer la proportion de patients susceptibles de présenter des effets secondaires.

II. Procédé :

1. Dans un échantillon de la population, on calcule la fréquence f du caractère.

Exemple : Lors d'essais cliniques avant la commercialisation du médicament, on a observé que sur un échantillon de 3 000 personnes, 59 ont présenté des effets secondaires

La fréquence observée est donc : $f = \dfrac{59}{3\,000} \approx 0{,}02$

2. A partir de la fréquence f obtenue, on définit un **intervalle de confiance** I qui dépend de f et de la taille n de l'échantillon.

> L'intervalle de confiance donne une **estimation de p par encadrement**

Propriété :

> Dans un échantillon de taille n, on note f_n la fréquence des succès.
> Pour n assez grand, la probabilité que p appartienne à l'intervalle $\left[f_n - \dfrac{1}{\sqrt{n}} \;;\; f_n + \dfrac{1}{\sqrt{n}}\right]$ est supérieure à 0,95

Définition :

> Soit f une fréquence observée du caractère étudié sur un échantillon de taille n.
> $\left[f - \dfrac{1}{\sqrt{n}} \;;\; f + \dfrac{1}{\sqrt{n}}\right]$ est appelé **intervalle de confiance de la proportion p au seuil de confiance 95 %.**
>
> **On utilise cet intervalle dès que :** $n \geqslant 30$; $nf \geqslant 5$ et $n(1-f) \geqslant 5$

Dans notre exemple : A quelle proportion de patients présentant des effets secondaires peut-on s'attendre après commercialisation du médicament avec un seuil de confiance de 95 % ? $f = 0{,}02$

$n = 3000$; $nf = 3000 \times 0{,}02 = 60 \geqslant 5$ et $n(1-f) = 3000 \times 0{,}98 \geqslant 5$

L'intervalle de confiance au seuil de 95% est $\left[0{,}02 - \dfrac{1}{\sqrt{3000}} \;;\; 0{,}02 + \dfrac{1}{\sqrt{3000}}\right]$

Soit $I = [0{,}002 \;;\; 0{,}038]$

La proportion de patients présentant des effets secondaires notée p est estimée à : $0{,}002 \leq p \leq 0{,}038$

Soit entre 0,2% et 3,8% avec un seuil de confiance de 95%.

Soit inférieur à 4%.

Estimation TS

Ex 1

$f = \frac{99}{140}$

$n = 140 \geq 30$

$nf = 99 \geq 5$

$n(1-f) = 140 \times \frac{41}{140} = 41 \geq 5$

$I = \left[f - \frac{1}{\sqrt{n}} ; f + \frac{1}{\sqrt{n}}\right] = [0,623 ; 0,792]$

La proportion p de personnes satisfaites vérifie

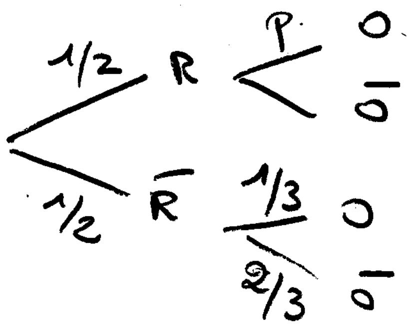
$0,623 \leq p \leq 0,792$

ou $\boxed{62,3\% \leq p \leq 79,2\%}$

Rq : plus de 60 % des personnes sont satisfaites.

Ex 2

1)



$q = p(O) = p(R \cap O) + p(\overline{R} \cap O)$

$q = \frac{1}{2}p + \frac{1}{2} \times \frac{1}{3}$

$\boxed{q = \frac{1}{2}p + \frac{1}{6}}$

2) $n = 1500 \geq 30$

a) $f = \frac{625}{1500}$

$nf = 625 \geq 5$

$n(1-f) = 875 \geq 5$

$I = [0,391 ; 0,442]$

la proportion q de jeunes répondant "oui" vérifie

$0,391 \leq q \leq 0,442$

ou $\boxed{39,1\% \leq q \leq 44,2\%}$

b) On a donc $0,391 \leq \frac{1}{2}p + \frac{1}{6} \leq 0,442$

$0,224 \leq \frac{1}{2}p \leq 0,275$

$\boxed{0,448 \leq p \leq 0,55}$

soit $\boxed{44,8\% \leq p \leq 55\%}$