Intervalle de fluctuation

Ce qui est affirmé sans preuve peut être nié sans preuve. *Euclide d'Alexandrie*

I- Intervalle de fluctuation

Lors d'une élection, deux listes, la rouge et la bleue, s'affrontent.

Un institut réalise un sondage pour connaître la **proportion** p des électeurs qui vont voter pour la liste rouge.

On constitue **un échantillon de n électeurs** et on calcule la proportion f des individus favorables à la liste rouge.

remarques :

On ne peut pas, dans la plupart des cas, connaître p a priori. On pourra seulement donner une " fourchette " probable.

f est la fréquence mesurée pour un échantillon (connu).

Dans le cas d 'une élection, un sondage donne la valeur de f.

On ne connaît p que lorsque l'élection est faite.

Savoir :

Les statisticiens ont démontré que dans les conditions suivantes:

*** p est compris entre 0,2 et 0,8

*** $n \geqslant 25$

il y a 95 % de chances pour que la valeur p appartienne à l'intervalle $\left[f-\frac{1}{\sqrt{n}}\,;f+\frac{1}{\sqrt{n}}\right]$.

Cet intervalle est appelé **intervalle de fluctuation** (ou fourchette) **au seuil de confiance** de 0,95.

Un petit calcul:

Montrer que $p \in \left[f-\frac{1}{\sqrt{n}}\,;f+\frac{1}{\sqrt{n}}\right]$ équivaut à $f \in \left[p-\frac{1}{\sqrt{n}}\,;p+\frac{1}{\sqrt{n}}\right]$

Démonstration :

$p \in \left[f-\frac{1}{\sqrt{n}}\,;f+\frac{1}{\sqrt{n}}\right]$ équivaut à $f-\frac{1}{\sqrt{n}} \leqslant p \leqslant f+\frac{1}{\sqrt{n}}$ (définition d'un intervalle fermé)

$f-\frac{1}{\sqrt{n}} \leqslant p \leqslant f+\frac{1}{\sqrt{n}}$ équivaut à $-\frac{1}{\sqrt{n}} \leqslant p-f \leqslant \frac{1}{\sqrt{n}}$ (retrancher ou ajouter le même nombre à chaque membre d'une inégalité ne change pas l'inégalité)

$-\frac{1}{\sqrt{n}} \leqslant p-f \leqslant \frac{1}{\sqrt{n}}$ équivaut à $-p-\frac{1}{\sqrt{n}} \leqslant -f \leqslant -p+\frac{1}{\sqrt{n}}$ (retrancher ou ajouter le même nombre à chaque membre d'une inégalité ne change pas l'inégalité)

$-p-\frac{1}{\sqrt{n}} \leqslant -f \leqslant -p+\frac{1}{\sqrt{n}}$ équivaut à $p+\frac{1}{\sqrt{n}} \geqslant f \geqslant p-\frac{1}{\sqrt{n}}$ (multiplier par (–1) ou prendre l'opposé de chaque membre d'une inégalité inverse l'inégalité)

$p+\frac{1}{\sqrt{n}} \geqslant f \geqslant p-\frac{1}{\sqrt{n}}$ équivaut à $f \in \left[p-\frac{1}{\sqrt{n}}\,;p+\frac{1}{\sqrt{n}}\right]$ (définition d'un intervalle fermé)

II- Utilisation du tableur pour vérifier cette propriété:

1) Réfléchir avant de " taper " n'importe quoi ...

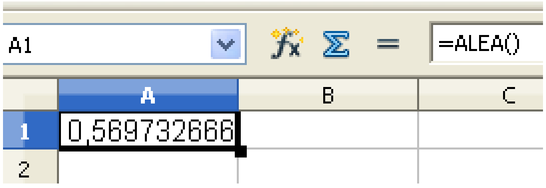
On suppose que la liste rouge a obtenu 35 % des voix. (On a donc $p = 0{,}35$)

On **simule** 100 échantillons de taille 100.

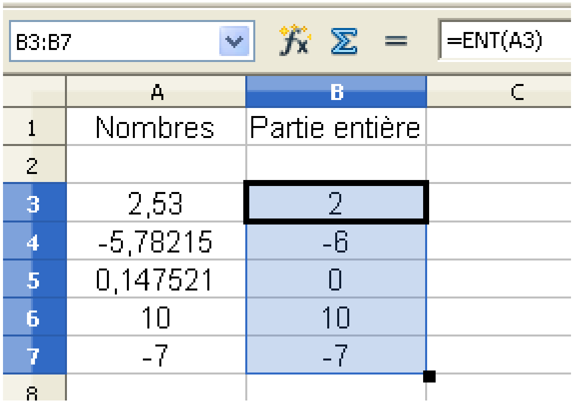
Déterminer l'intervalle $\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$.

2) Les fonctions utilisées sur le tableur.

a) La fonction ALEA() retourne AU HASARD un nombre x tel que $0 \leqslant x < 1$



b) La fonction ENT() retourne la partie entière d'un nombre.


B3:B7 =ENT(A3)

	A	B	C
1	Nombres	Partie entière	
2			
3	2,53	2	
4	-5,78215	-6	
5	0,147521	0	
6	10	10	
7	-7	-7	
8			

c) On entre dans une cellule : =ENT(ALEA()). Quel est le résultat ?

0 puisque les nombres sont compris entre 0 et 1 (1 exclu)

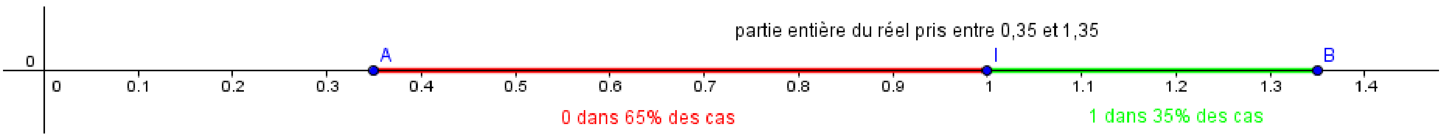
d) On entre dans une cellule : =ENT(ALEA()+0,35).

Quels sont les deux résultats possibles ? 0 ou 1

Donner la probabilité d'obtenir chacun des deux résultats :

La probabilité du résultat 0 est 0,65

La probabilité du résultat 1 est 0,35



e) Lorsqu'on recopie des formules dans la feuille de calcul, les noms des cellules de la formule augmentent de 1 colonne au fur et à mesure vers la droite et d'une ligne au fur et à mesure vers le bas.

Si on veut " fixer " le nom de la colonne ou de la ligne, on place le caractère " \$ " devant la colonne ou la ligne.

En faisant : " Maj+F4 ", les " \$ " s'écrivent devant la colonne et la ligne.

3) Au tableur

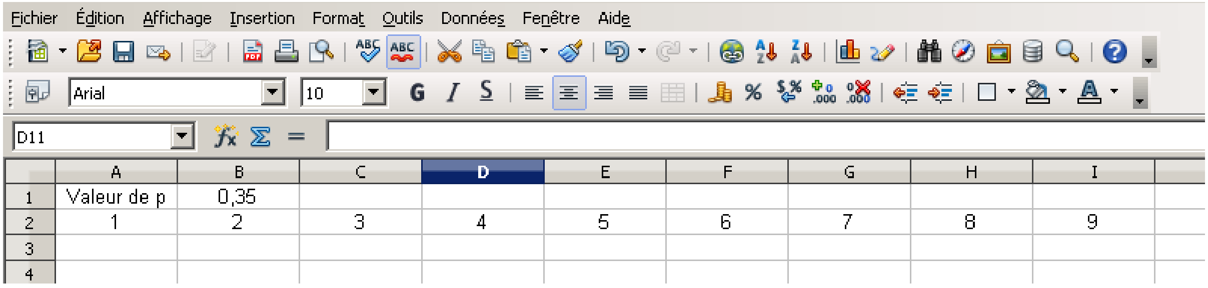
a) Ouvrir une feuille de calcul du tableur.

Intervalle de fluctuation

Ce qui est affirmé sans preuve peut être nié sans preuve. *Euclide d'Alexandrie*

Dans la cellule **B1**, entrer la valeur de *p (on pourra ainsi faire différents essais en ne modifiant que la valeur dans la cellule* **B1.**

La ligne 2 donne le numéro d'un individu dans un échantillon (Ici, il y a 100 individus: il suffit de taper 1 dans la cellule **A2** et de tirer vers la droite jusqu'à la colonne **CV.**

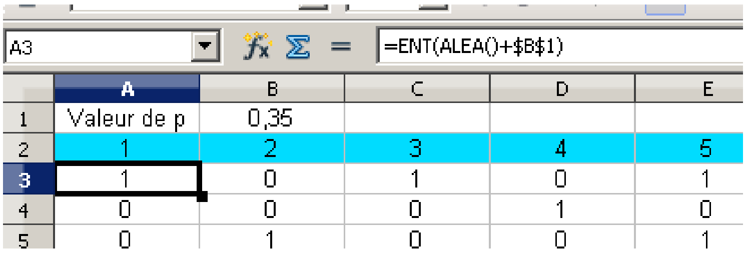


2) Dans la cellule **A3**, on entre la formule

= ENT(ALEA()+B1)

Quel est le résultat obtenu? 0 ou 1

Que représente ce résultat? le vote de l'individu n°1

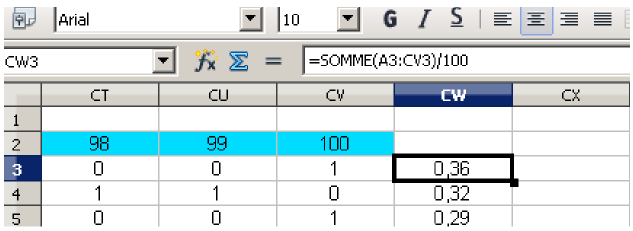


Si le résultat est 0, alors " l'individu " simulé est contre la liste rouge.

Si le résultat est 1, alors " l'individu " simulé est pour la liste rouge.

Recopier cette formule vers la droite jusqu'à la colonne **CV**.

(On a ainsi constitué notre premier échantillon de taille 100)



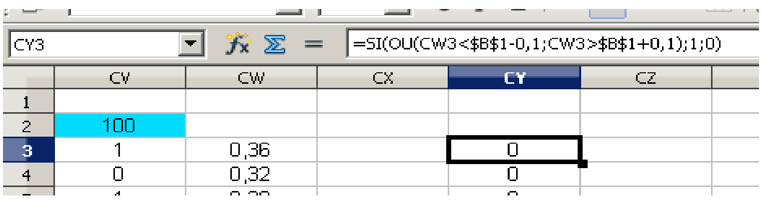
Dans la cellule **CW3**, écrire la formule =SOMME(A3:CV3)/100

Que représente le nombre obtenu dans cette cellule **CW3**?

La proportion dans l'échantillon de 100 individus des " POUR ".

3) Constituer les 100 échantillons en recopiant vers le bas la ligne 3 (jusqu'à la ligne 102)

4) Dans la cellule **CY3**, écrire la formule =SI(OU(CW3<B1–0,1;CW3>B1+0,1);1;0)



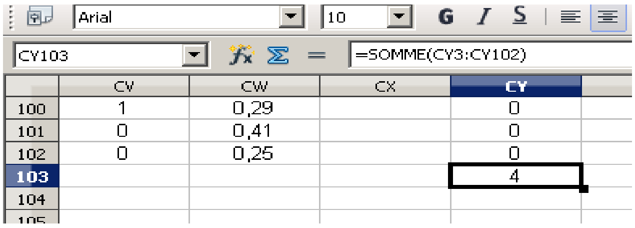
Que représente le résultat obtenu dans la cellule **CY3**?

Si le résultat est 0, la proportion de l'échantillon est dans l'intervalle de fluctuation.

Si le résultat est 1, la proportion de l'échantillon n'est pas dans l'intervalle de fluctuation.

Recopier cette formule vers le bas pour les 100 échantillons.

Dans la cellule **CY103** , écrire la formule =SOMME(CY3:CY102)



Que représente le résultat de cette cellule **CY103**?

Le nombre d'échantillons sur les 100 échantillons qui donnent une proportion en-dehors de l'intervalle de fluctuation.

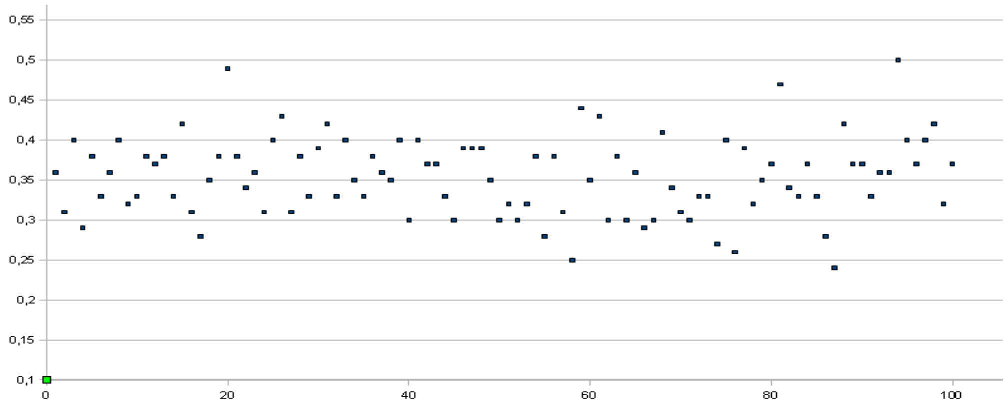
5) Recalculer en faisant Ctl+Maj+F9.

L'affirmation du I- est-elle vérifiée à chaque fois?

Il est possible mais peu probable que le nombre d'échantillons en-dehors de l'intervalle de fluctuation soit supérieur à 5.

III- Nuage de points

Sélectionner la plage **CW3:CW102** puis représenter les résultats par un nuage de points.



Comment utiliser ce diagramme pour vérifier que 95 % des résultats sont dans l'intervalle de fluctuation ?

IV- Un exercice:

Lors d'un sondage, sur 1068 personnes interrogées, 550 ont répondu qu'elles voteraient pour M. Trucmuche.

M. Trucmuche déclare alors qu'il sera certainement élu.

Que pensez-vous de la déclaration de M. Trucmuche?

Calcul de la fréquence observée f_{obs} lors du sondage :

$f_{obs} = \dfrac{550}{1068} \approx \dfrac{51{,}50}{100}$ en arrondissant à 0,000 1 près.

Or, il y a 95 % de chances pour que la valeur p appartienne à l'**intervalle de fluctuation** $\left[f-\dfrac{1}{\sqrt{n}}\,;f+\dfrac{1}{\sqrt{n}}\right]$.

Ici, n = 1 068.

L'intervalle de fluctuation est : $I_f = \left[\frac{550}{1068} - \frac{1}{\sqrt{1068}} ; \frac{550}{1068} + \frac{1}{\sqrt{1068}}\right] \approx [0{,}48 ; 0{,}55]$

M. Trucmuche ne peut pas assurer qu'il sera certainement élu.