

FLUCTUATION D'ÉCHANTILLONNAGE

Objectif :

Observation des fluctuations dans un Tirage de N pile ou face simulé avec le Tableur Excel

Application : développer l'esprit critique face aux sondages d'opinion.

(Statistiques en classe de Seconde)

Partie I : Construction de la simulation au moyen du Tableur Excel

- ① Simuler le jet d'une pièce avec Excel
- ② Étendre la simulation à 10 jets consécutifs de la même pièce :
- ③ Pour tirer 10 nouvelles valeurs au hasard, appuyer sur la touche **F9**.
- ④ Préparation du tableau des effectifs :
- ⑤ Préparation du graphique :
Le graphique en forme de camembert traduit bien visuellement les proportions relatives de Pile et Face
- ⑥ Observation des fluctuations des fréquences à taille d'échantillon constante :
Retirer 10 nouvelles valeurs et observer la nouvelle répartition de Pile et Face :
Que constate-t-on ?
- ⑦ Observation de l'influence de la taille de l'échantillon :
Créer un tirage de $N = 20$ valeurs de Pile ou Face
Observer alors les fluctuations comme précédemment.
Puis refaire la même chose avec $N = 50$; $N = 100$; jusqu'à $N = 200$.
Que peut-on dire des variations de fréquences quand la taille N de l'échantillon augmente ?

Partie II : Exploitation de la simulation comme aide à l'interprétation rationnelle des sondages d'opinion

Dans une élection présidentielle, traditionnellement très serrée, deux candidats P et F s'affrontent au 2^{ème} tour. Un sondage tire N électeurs au hasard dans les listes électorales, et leur demande pour qui ils vont voter.

On fait l'hypothèse suivante : P et F ont la même chance d'être choisis (Hypothèse du Match Nul : le pire cas de figure pour l'institut de sondage !).

Dans ce cas, ce sondage revient à tirer à Pile ou Face le choix de l'électeur.

① Simuler alors 20 sondages avec un échantillon représentatif d'effectif $N = 10$ en prenant soin de relever le résultat des 2 sondages donnant le plus grand écart = [score du perdant ; score du gagnant]. Le plus petit de ces 2 intervalles sera nommé Fourchette.

Quel est le pourcentage des sondages à l'intérieur de la fourchette ? En dehors ?

② Si un sondage réel est logé à la même enseigne que les 20 simulations (c'est-à-dire : satisfait aussi l'Hypothèse du Match Nul), quel pourcentage de chances a-t-il de tomber dans la fourchette ?

③ Un sondage réel avec $N = 10$ a donné le résultat suivant :

P a 80 % des voix et F a le reste : 20 %.

Peut-on alors dire que P va gagner l'élection ? Pourquoi ?

④ Simuler 20 nouveaux sondages (procéder comme en 1.), mais avec $N = 100$ cette fois-ci.

Un sondage réel avec $N = 100$ a pour résultat : P a 62 % des voix, et F a les 38 % restants.

Peut-on affirmer raisonnablement que P va gagner l'élection ?

⑤ Quel risque (en % de chances) a-t-on alors de rejeter à tort l'Hypothèse du Match Nul sur la foi des 20 simulations ? (voir résultat du 2.)

⑥ A l'aide de la formule théorique de l'Intervalle de Confiance à 95 % admise dans le cours de 2^{nde}, proposer une taille d'échantillon N qui réduise la fourchette aux environs de [47 % ; 53%].

Vérifier avec 20 simulations.

Indications

Partie I : Construction de la simulation au moyen du Tableur Excel

- ① Simuler le jet d'une pièce avec Excel
 Dans la case **A1** (comme à la bataille navale) écrire la formule : **=SI(ALEA()<0,5;"Pile";"Face")**
 ALEA() est un nombre aléatoire compris entre 0 et 0,99999999 que l'ordinateur tire à notre place.
 Il y a donc 1 chance sur 2 que ce nombre soit plus petit que 0,5. Dans ce cas, la réponse affichée dans la case est Pile, et dans le cas contraire Face.
- ② Etendre la simulation à 10 jets consécutifs de la même pièce :
 Recopier dans les cellules **A2 à A10**
 Les 10 premières cases de la colonne prennent des valeurs aléatoires équiprobables parmi Pile ou Face.
- ③ Pour tirer 10 nouvelles valeurs au hasard, appuyer sur la touche **F9**.
 Les 10 nouvelles valeurs tirées sont affichées à la place des anciennes.
- ④ Préparation du tableau des effectifs :
 Ecrire Pile en case **D1**, Face en case **E1**, et Effectifs en case **C2**
 Dans la case **D2** : écrire la formule : **=NB.SI(\$A\$1:\$A\$1000;"=PILE")**
 Dans la case **E2** : écrire la formule : **=NB.SI(\$A\$1:\$A\$1000;"=FACE")**
 La fonction NB.SI parcourt toutes les 1000 cases désignées, y compris les cases vides.
 Elle compte le nombre des cases dont le résultat est exactement Pile (resp. Face) et l'affiche.
- ⑤ Préparation du graphique :
 Passer la zone rectangulaire **D1:E2** en surbrillance
 Cliquer en haut l'icône de l'Assistant graphique 
 Une petite croix avec un dessin en bas à droite apparaît : Il s'agit alors de sélectionner sur la feuille une zone pour représenter le dessin.

Etape 1 sur 5 : plage de données Les données sont déjà sélectionnées. Cliquer **Suivant**.

Etape 2 sur 5 : Type de graphique Choisir **Secteurs** puis cliquer **Suivant**.

Etape 3 sur 5 : Format de graphique Choisir **7** puis cliquer **Suivant**.

Etape 4 sur 5 Série de données en : **ligne**
 Utiliser les **1** premières colonnes pour les étiquettes de l'axe des catégories (X) "
 Cliquer **Suivant** .

Etape 5 sur 5 Titre du graphique : **Résultats des tirages**
 Cliquer **Fin** .
 Le graphique en forme de camembert apparaît.
 Il traduit bien visuellement les proportions relatives de Pile et Face