Simulation - Échantillonnage

Christophe ROSSIGNOL*

Année scolaire 2018/2019

Table des matières

Liste des tableaux

1 Simulation d'expérience

Activité : Activité 1 page 133[1] [TransMath]

1.1 Définition

Définitions :
- Une expérience aléatoire est une expérience que l'on peut reproduire dans les mêmes conditions et dont on connaît à priori toutes les résultats possibles, sans pouvoir dire avec certitude le résultat qui se produira.
- Simuler une expérience, c'est la remplacer par une autre, plus simple à réaliser, et qui permettra d'obtenir des résultats similaires.

Remarque : Pour cela, on utilise généralement un générateur de nombres au hasard. Il y en a de plusieurs types :
- les tables de nombres au hasard (très peu utilisées de nos jours) ;
- la commande `ALEA.ENTRE.BORNES` d'un tableur (voir 1.2) ;
- la commande `RANDOM` des calculatrices.

1.2 Utilisation d'un tableur

Pour simuler un tirage au sort dans un ensemble à p éléments à l'aide d'un tableur, on peut utiliser la méthode suivante :
- On attribue à chaque élément un nombre entre 1 et p.
- On simule le tirage aléatoire en utilisant la commande `=ALEA.ENTRE.BORNES(1,p)`
- Pour calculer la fréquence d'une issue pour N tirages au sort, on utilise la commande `=NB.SI(plage;critère)/N` où la plage est l'ensemble des cellules où ont été effectués les simulations.

Exercices : 1, 2, 3, 4, 6, 7 page 139[2] – 8, 9, 11 page 139[3] [TransMath]

2 Fluctuation d'échantillonnage

Activité : Activité 2 page 134[4] [TransMath]

2.1 Échantillon statistique

Définitions :
1. On appelle échantillon de taille n la série statistique formée des résultats obtenus lorsqu'on répète n fois une expérience dans les mêmes conditions.
2. La distribution de fréquences associée à l'échantillon est le tableau des fréquences issues de cette échantillon.

Exemple : On lance dix fois de suite une pièce bien équilibrée. On obtient P (Pile), F (Face), F, F, P, P, F, P, F, F.
Il s'agit d'un échantillon de taille 10 (on a répété l'expérience 10 fois).
La distribution de fréquences est :

Résultat	Pile	Face
Effectif	4	6
Fréquence	**0,4**	**0,6**

1. Simuler un tirage dans une urne.
2. Utilisation du tableur.
3. Simulations à l'aide d'un tableur.
4. Découvrir l'intervalle de fluctuation.

2.2 Fluctuation d'échantillonnage

— Les distributions de fréquences varient d'un échantillon à l'autre pour la même expérience. C'est ce qu'on appelle la fluctuation d'échantillonnage.
— Même pour des échantillon de même taille, la distribution de fréquences peut varier.
— Lorsque la taille de l'échantillon augmente, les distributions de fréquences ont tendance à se stabiliser. Plus précisément, on a la propriété suivante :

Propriété : (admise)
Soit un caractère dont la proportion dans une population donnée est p.
Si $0,2 \leq p \leq 0{,}8$ et si $n \geq 25$, alors 95 % des échantillons de taille n auront une distribution de fréquence dans la quelle la fréquence du caractère sera dans l'intervalle :

$$\left[p - \frac{1}{\sqrt{n}}\,;\,p + \frac{1}{\sqrt{n}}\right]$$

Cet intervalle est appelé intervalle de fluctuation au seuil de 95 %.

Exercices : 15, 16 page 140[5] [TransMath]

Activité : Activité 3 page 136[6] et Activité 4 page 137[7] [TransMath]

Remarques : 1. Connaissant l'intervalle de fluctuation d'un caractère dans un échantillon, on va pouvoir valider ou rejeter cet échantillon :
— Si la fréquence observée de ce caractère est dans l'intervalle de fluctuation, on valide l'échantillon au seuil de 95 %.
— Si la fréquence observée de ce caractère n'est pas dans l'intervalle de fluctuation, on rejette l'échantillon au seuil de 95 %.

L'expression « au seuil de 95 % » signifie que, dans environ 5 % des cas, la décision prise risque d'être incorrecte.

2. Si l'on étudie des échantillons de tailles de plus en plus grandes, on constate que l'amplitude des fluctuations diminuent. Les fréquences ont tendance à se stabiliser autour d'une valeur, que l'on peut utiliser comme une estimation de la proportion p.

Exercices : 17, 18 page 140 et 20 page 141[8] [TransMath]

2.3 Intervalle de confiance

Activité : Activité 5 page 138[9] [TransMath]

Définition : si c'est la proportion p d'un caractère qui est inconnue, on peut procéder de la façon suivante :
On étudie un échantillon de taille n, avec $n \geq 25$. Le caractère étudié apparaît alors avec une fréquence f.
Si $0,2 \leq f \leq 0,8$, on peut estimer que la proportion p du caractère dans la population totale vérifie :

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

avec une probabilité d'au moins 0,95.
Cet intervalle est appelé intervalle de confiance au seuil des 0,95.

Remarque : Dans environ 5 % des cas, l'intervalle de confiance peut ne pas contenir la proportion p...

Exercices : 21 page 141 ; 23, 24, 25 page 142 et 28, 29 page 143[10] [TransMath]

5. Déterminer un intervalle de fluctuation.
6. Utiliser l'intervalle de fluctuation.
7. Estimer une proportion p inconnue.
8. Utiliser l'intervalle de fluctuation.
9. Estimer une proportion inconnue par l'intervalle de confiance.
10. Intervalle de confiance.

Références

[TransMath] Transmath Seconde, Nathan (édition 2010).

2, 3