CH-18-Intervalle de fluctuation et Estimation

I- Intervalle de fluctuation (sert à accepter ou à rejeter une hypothèse)

1) Echantillon et intervalle de fluctuation asymptotique

- Un « échantillon de taille n » correspond aux résultats obtenus après avoir répété n fois une expérience aléatoire de manière indépendante.

Exemple : On lance 80 fois une pièce de monnaie et on obtient 35 fois le résultat « pile ».

Remarque : Cet échantillon donne donc une proportion (fréquence), $f = \frac{35}{80} = 0{,}4375$ pour le nombre de « pile » obtenu.

- On s'intéresse ensuite à la probabilité théorique p d'obtenir le résultat étudié.

Exemple : Si on suppose que la pièce est bien équilibré (non truqué), alors la probabilité d'obtenir « pile » est théoriquement de $p = 0{,}5$.

- Dans l'objectif de valider ou non cette hypothèse à l'aide de notre échantillon expérimentale, on calcule « l'intervalle de fluctuation asymptotique au seuil de confiance 95% de la proportion étudié dans un échantillon de taille n » qui correspond à :

$$\left[p - 1{,}96 \times \frac{\sqrt{p(1-p)}}{\sqrt{n}} \; ; \; p + 1{,}96 \times \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$$

Remarques :

- Les bornes de cet intervalle dépendent de la taille n de l'échantillon et il est centré autour de la probabilité théorique p.
- Plus n est grand et plus l'intervalle se « resserre » autour de la valeur p.
- Conditions : intervalle valable si : $n \geq 30$, $n \times p \geq 5$ et $n \times (1-p) \geq 5$

Exemple : Dans la situation précédente, l'intervalle de fluctuation au seuil de 95% de la proportion de « pile » obtenu dans l'échantillon de taille 80 est :

$$\left[0{,}5 - 1{,}96 \times \frac{\sqrt{0{,}5 \times (1-0{,}5)}}{\sqrt{80}} \; ; \; 0{,}5 + 1{,}96 \times \frac{\sqrt{0{,}5 \times (1-0{,}5)}}{\sqrt{80}}\right] \approx [0{,}390 \; ; \; 0{,}610]$$

(à arrondir par DÉFAUT) (par EXCÈS)

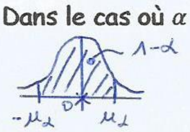
- On peut être amené à faire des études avec un seuil de confiance différent de 95%. De manière générale, on a donc la définition suivante :

> Définition : L'intervalle $\left[p - u_\alpha \times \frac{\sqrt{p(1-p)}}{\sqrt{n}} \; ; \; p + u_\alpha \times \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right]$ est un intervalle de fluctuation asymptotique au seuil de confiance $(1-\alpha)$ de la proportion étudié dans un échantillon de taille n.

Exemple : Dans le cas où $\alpha = 0{,}05$, alors $1 - \alpha = 0{,}95$ soit 95% , et on retrouve le fait que : $u_{0{,}05} = 1{,}96$

(voir chapitre loi normale)

Rappel :



2) Prise de décision (accepté ou rejeté une hypothèse)

- Il faut faire une hypothèse pour proposer une probabilité théorique p.

Exemple : On a supposé que la pièce était bien équilibré, ce qui a donné $p = 0,5$.

- On va accepter ou rejeter cette hypothèse en analysant notre échantillon via l'intervalle de fluctuation asymptotique au seuil de confiance $(1-\alpha)$ que l'on notera I.

> **La règle de décision est la suivante :**
> - Si $f \in I$, alors on accepte l'hypothèse avec un niveau de confiance de $(1-\alpha)$ (on dit aussi avec un risque d'erreur de α)
> - Si $f \notin I$, alors on rejette l'hypothèse avec un niveau de confiance de $(1-\alpha)$ (on dit aussi avec un risque d'erreur de α)

Exemple : Dans l'étude précédente, on a : $f = 0,4375$

Or, $0,4375 \in [0,390 ; 0,610]$

Donc on ACCEPTE l'hypothèse selon laquelle la pièce est bien équilibré avec un niveau de confiance de 95% (ou un risque d'erreur de 5%)

Exercice : On prend une autre pièce du porte monnaie et on la lance 400 fois. On obtient cette fois-ci 175 « pile ».

Peut-on accepter ou non l'hypothèse $p = 0,5$ au seuil de risque 5% ?

- D'après l'échantillon, on a : $n = 400$ et $p = 0,5$
- Conditions : $n = 400 \geq 30$, $n \times p = 200 \geq 5$ et $n \times (1-p) = 200 \geq 5$.
- L'intervalle de fluctuation asymptotique au seuil de confiance de 95% de la proportion de "pile" obtenu dans l'échantillon de taille 400 est :

$$\left[0,5 - 1,96 \times \frac{\sqrt{0,5 \times (1-0,5)}}{\sqrt{400}} ; 0,5 + 1,96 \times \frac{\sqrt{0,5 \times (1-0,5)}}{\sqrt{400}}\right]$$

$$\simeq [0,451 ; 0,549]$$

- De plus, $f = \frac{175}{400} = 0,4375$

Or, $0,4375 \notin [0,451 ; 0,549]$

Donc on REJETTE l'hypothèse selon laquelle $p = 0,5$ (pièce bien équilibré) avec un niveau de confiance de 95% 2 (ou un risque d'erreur de 5%)

28 Un groupe de citoyens demande à la municipalité d'une ville la modification d'un carrefour en affirmant que 40 % des automobilistes tournent en utilisant une mauvaise file.
Un officier de police constate que, sur 500 voitures prises au hasard, 190 prennent une mauvaise file.

1. Avec l'hypothèse faite par le groupe de citoyens, déterminer un intervalle de fluctuation asymptotique de la fréquence des automobilistes fautifs au seuil de confiance de 95 %.

2. D'après l'échantillon, peut-on considérer, au seuil de confiance 95 %, comme exacte l'affirmation du groupe de citoyens ?

1) D'après l'échantillon, on a : $n = 500$ et $p = \frac{40}{100} = 0{,}40$

Conditions : $n = 500 \geq 30$, $n \times p = 200 \geq 5$ et $n \times (1 - p) = 300 \geq 5$

L'intervalle de fluctuation asymptotique de la fréquence des automobilistes fautifs au seuil de confiance de 95% est :

$$\left[0{,}40 - 1{,}96 \times \frac{\sqrt{0{,}40 \times 0{,}60}}{\sqrt{500}} ; 0{,}40 + 1{,}96 \times \frac{\sqrt{0{,}40 \times 0{,}60}}{\sqrt{500}}\right] \approx [0{,}357; 0{,}443]$$

2) Comme $f = \frac{190}{500} = 0{,}38$

Or, $0{,}38 \in [0{,}357; 0{,}443]$

Donc on ACCEPTE comme exacte l'affirmation du groupe de citoyens avec un niveau de confiance de 95% (ou un risque d'erreur de 5%).

29

Selon la loi de Mendel, en croisant deux hétérozygotes Aa et Aa, où A est dominant, on doit obtenir les phénotypes A et a dans les proportions $\frac{3}{4}$ et $\frac{1}{4}$.

Pour un caractère donné, on a obtenu les résultats suivants :

phénotype A : 130 ; phénotype a : 70.

Au seuil de risque 5 %, peut-on dire que la loi de Mendel est vérifiée par ce caractère ?

<u>Méthode 1</u> : En raisonnant sur le phénotype « A »

D'après l'échantillon, on a : $n = 130 + 70 = 200$ et $p = \frac{3}{4} = 0{,}75$

Conditions : $n = 200 \geq 30$, $n \times p = 150 \geq 5$ et $n \times (1 - p) = 50 \geq 5$

L'intervalle de fluctuation asymptotique de la fréquence des phénotypes A au seuil de confiance de 95% est :

$$\left[0{,}75 - 1{,}96 \times \frac{\sqrt{0{,}75 \times 0{,}25}}{\sqrt{200}} ; 0{,}75 + 1{,}96 \times \frac{\sqrt{0{,}75 \times 0{,}25}}{\sqrt{200}}\right] \approx [0{,}689; 0{,}811]$$

Comme $f = \frac{130}{200} = 0{,}65$

Or, $0{,}65 \notin [0{,}689; 0{,}811]$

Donc on REJETTE l'hypothèse selon laquelle la loi Mendel serait vérifiée par ce caractère avec un niveau de confiance de 95% (ou un risque d'erreur de 5%).

<u>Méthode 2</u> : En raisonnant sur le phénotype « a »

D'après l'échantillon, on a : $n = 130 + 70 = 200$ et $p = \frac{1}{4} = 0{,}25$

Conditions : $n = 200 \geq 30$, $n \times p = 50 \geq 5$ et $n \times (1 - p) = 150 \geq 5$

L'intervalle de fluctuation asymptotique de la fréquence des phénotype A au seuil de confiance de 95% est :

$$\left[0{,}25 - 1{,}96 \times \frac{\sqrt{0{,}25 \times 0{,}75}}{\sqrt{200}} ; 0{,}75 + 1{,}96 \times \frac{\sqrt{0{,}25 \times 0{,}75}}{\sqrt{200}}\right] \approx [0{,}189; 0{,}311]$$

Comme $f = \frac{70}{200} = 0{,}35$

Or, $0{,}35 \notin [0{,}189; 0{,}311]$

Donc on REJETTE l'hypothèse selon laquelle la loi Mendel serait vérifiée par ce caractère avec un niveau de confiance de 95% (ou un risque d'erreur de 5%).