

Fluctuation d'échantillonnage

1er Gestion - Administration

1. Notion de fluctuation d'échantillonnage

Ce premier paragraphe constitue un rappel du cours de seconde.

Lorsque l'on répète n fois une expérience aléatoire, l'ensemble des résultats collectés constitue

Qu'est ce qu'une expérience **aléatoire** ?

.....

Exemple d'expérience aléatoire :

Exemple :

- On lance 50 fois un dé cubique bien équilibré.
 - Quelles sont les éventualités (ou issues) possibles?
 - Quelle est la taille de l'échantillon?
 - On réalise un premier échantillon pour lequel on obtient :

Face	1	2	3	4	5	6
Effectif	8	5	9	9	8	11

- Quelle est la fréquence de sortie de la face 2?
- Quelle est la fréquence de sortie de la face 5?

- On lance recommence 3 fois la prise d'échantillon de taille 50. On obtient :

Face	1	2	3	4	5	6
Fréquence échantillon A	0.16	0.1	0.18	0.18	0.16	0.22
Fréquence échantillon B	0.20	0.10	0.16	0.12	0.22	0.20
Fréquence échantillon C	0.12	0.22	0.14	0.16	0.18	0.18

Obtient-on les mêmes résultats? Qu'observe-t-on?

.....
.....
.....

- On simule par informatique 5000 lancers de dé.
- Quelle est la taille de l'échantillon?
- On réalise la prise de 3 échantillons de on obtient les résultats suivant :

Face	1	2	3	4	5	6
Fréquence échantillon A	0.165	0.170	0.162	0.161	0.168	0.174
Fréquence échantillon B	0.162	0.161	0.167	0.178	0.167	0.165
Fréquence échantillon C	0.174	0.166	0.163	0.162	0.168	0.167

- Qu'observe-t-on?
.....
.....
- Quelle est l'influence de la taille des échantillons?
.....
.....
- Vers quelle valeur tend la fréquence de sortie de chaque face?
.....
.....
- A quoi correspond cette valeur?
.....
.....

À retenir

- La fréquence d'un même caractère sur plusieurs échantillons de même taille prend des valeurs différentes selon les échantillons.
- Ce phénomène s'appelle
- Plus la taille de l'échantillon est grande
- Pour un grand échantillon, la fréquence est

2. Distribution d'échantillonnage d'une fréquence

On considère une population où la fréquence d'un caractère est p et dans laquelle on prélève au hasard N échantillons de taille n .
 La liste des fréquences $f_1; f_2; f_3; \dots; f_N$ du caractère obtenue sur les N échantillons constitue une

Exemple :

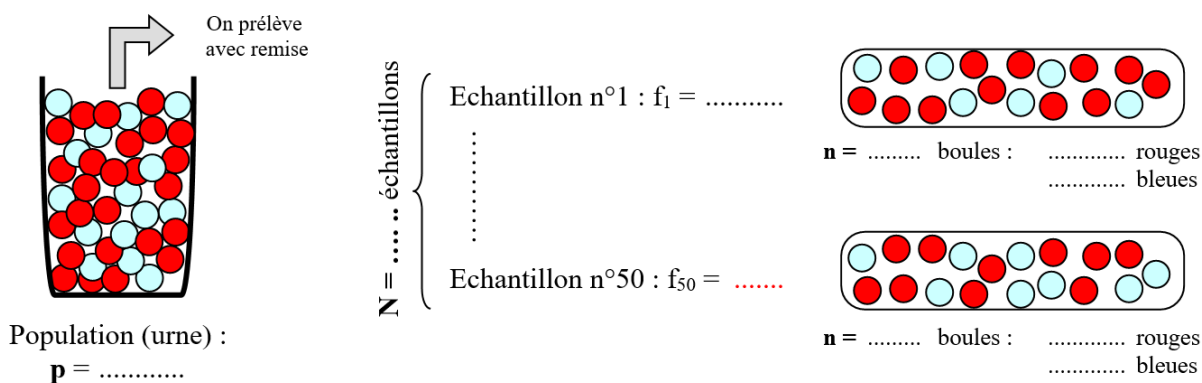
Une urne contient 60 % de boules rouges et 40 % de boules bleues. On prélève avec remise $N = 50$ échantillons aléatoires de même taille $n = 100$.

- Quelle est la fréquence p du caractère "boule rouge" dans l'urne ?

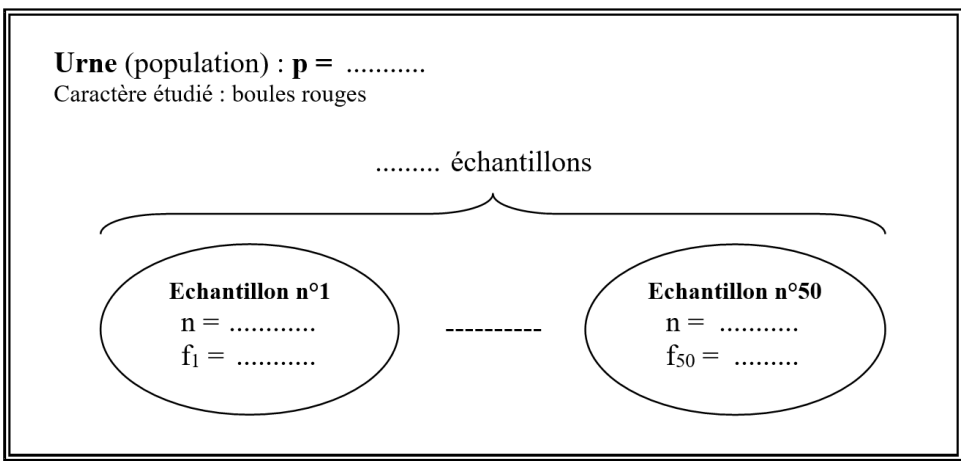
- Obtient-on la même fréquence f des boules rouges dans chaque échantillon ?

- On obtient la distribution d'échantillonnage de la fréquence des boules rouges suivantes :

Numéro échantillon	1	2	3	4	5	...	47	48	49	50
Fréquence f_i des rouges	0.61	0.60	0.64	0.53	0.6	...	0.55	0.62	0.6	0.51



Un schéma utile pour bien faire la différence entre p et les différentes valeurs de f_i :



3. Fréquence moyenne d'une distribution d'échantillonnage

La fréquence moyenne \bar{f} d'une distribution d'échantillonnage est donnée par la formule :

$$\bar{f} = \frac{\dots\dots\dots}{\dots\dots\dots}$$

avec : le nombre d'échantillons
 et : la fréquence de l'échantillon i

Exemple :

Le tableau suivant donne la fréquence d'apparition de la face 6 dans 5 échantillons de 100 lancers d'un dé cubique :

Échantillon : i	1	2	3	4	5
Fréquence de l'échantillon f_i	0.15	0.16	0.19	0.17	0.13

Calculer la fréquence moyenne de la distribution d'échantillonnage :

$$\bar{f} = \dots\dots\dots$$

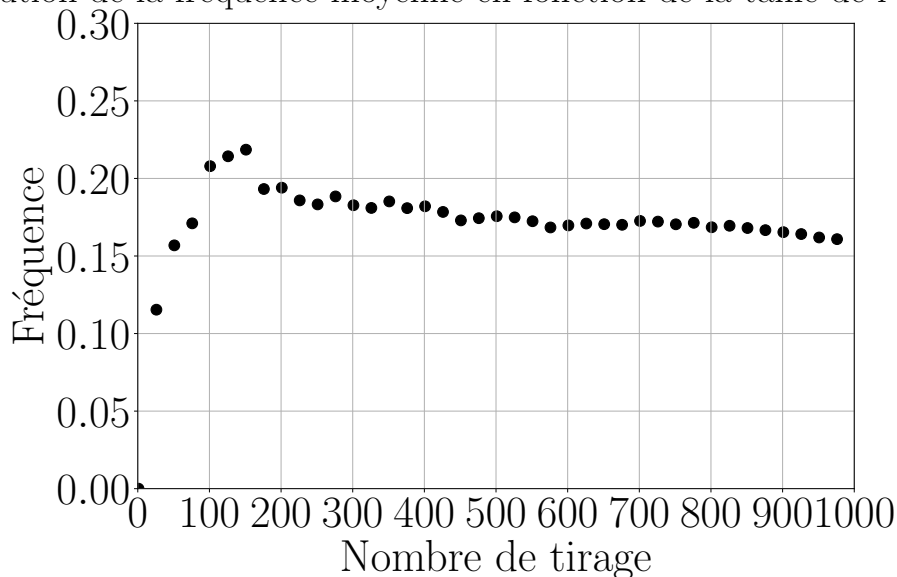
À retenir

Lorsque la taille n des échantillons ou le nombre N d'échantillons est assez important, ..

Exemple :

On s'intéresse à la fréquence de sortie de la face 2 dans le lancer d'un dé cubique. On réalise $N = 10$ échantillons. La taille n de chaque échantillon varie de 1 à 1 000. Le graphique ci-contre montre l'évolution de la fréquence moyenne en fonction de la taille de l'échantillon.

Evolution de la fréquence moyenne en fonction de la taille de l'échantillon



Qu'observe t-on ? Donner une explication

.....
.....
.....
.....

4. Intervalle de fluctuation à 95%

On s'intéresse à un caractère dont on connaît la fréquence p dans la population totale.

Sous certaines conditions (n assez grand et p ni très petit ni très grand), la probabilité qu'une fréquence du caractère étudié sur un échantillon de taille n appartient à l'intervalle $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$,

Autre formulation possible :

Sous certaines conditions (n assez grand et p ni très petit ni très grand), au moins 95% des échantillons de taille n ont une fréquence du caractère étudié comprise entre $p - \frac{1}{\sqrt{n}}$ et $p + \frac{1}{\sqrt{n}}$.

Pourquoi fait-on cela ?

Cela permet de pouvoir exercer un regard critique sur les données statistiques.

En effet, la connaissance de cette condition permet de vérifier si

.....
.....

• Dans le cas où la fréquence f de l'échantillon est comprise dans l'intervalle $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$, on pourra conclure que

.....

• Dans le cas où la fréquence f de l'échantillon n'est pas comprise dans l'intervalle $\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}} \right]$, on pourra conclure que

.....

Exemple :

Parmi les 120 salariés d'une entreprise A, seulement 51 sont des femmes. L'entreprise B emploie 2 600 salariés dont 1 222 sont des femmes.

Dans une entreprise, la parité hommes - femmes sous-entend que l'identité sexuelle des personnes n'intervient pas dans leur recrutement.

Le respect de la parité suppose donc théoriquement un même nombre d'hommes et de femmes salariés au sein de l'entreprise.

Laquelle de ces deux entreprises respecte le mieux la parité hommes – femmes ?

Méthode de résolution :

- Analyse de la situation et des paramètres mis en jeu :
 - Quel est le caractère étudié au sein d'une entreprise ?
.....
 - Considérant que le respect de la parité est effectif, quelle est la fréquence p de femmes salariées dans l'entreprise ?
.....
 - Calculer la fréquence de femmes dans l'entreprise A :
.....
 - Calculer la fréquence de femmes dans l'entreprise B :
.....
 - Que peut-on en conclure à ce moment là de la résolution ?
.....
 - Est-ce vraiment le cas ? Quel paramètre n'avons pas encore pris en compte ?
.....
- Il faut calculer pour chacune de ces deux entreprises leur intervalle de fluctuation à 95 % avant de pouvoir donner la bonne réponse.
 - Calculer l'intervalle de fluctuation à 95 % pour l'entreprise A (arrondir au millième) :
.....
.....

Cela signifie qu'au moins 95 % des entreprises de 120 salariés qui respectent la parité hommes – femmes ont une fréquence de femmes comprises entre

- Calculer l'intervalle de fluctuation à 95 % pour l'entreprise B (arrondir au millième) :

.....
.....

Cela signifie qu'au moins 95 % des entreprises de 2 600 salariés qui respectent la parité hommes – femmes ont une fréquence de femmes comprises entre

- Conclusion : Quelle entreprise respecte le mieux la parité hommes – femmes ?

.....
.....
.....
.....
.....