Chapitre M6

Statistique et Probabilités 4

FLUCTUATION DE FREQUENCE SELON LES ECHANTILLONS, PROBABILITES

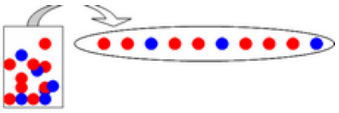
Capacités	Connaissances
Expérimenter, à l'aide d'une simulation informatique, la prise d'échantillons aléatoires de taille n fixée, extraits d'une population où la fréquence p relative à un caractère est connue	Distribution d'échantillonnage d'une fréquence
Calculer la moyenne de la série des fréquences f_i des échantillons aléatoires de même taille n prélevés. Comparer la fréquence p de la population et la moyenne de la série des fréquences f_i des échantillons aléatoires de même taille n prélevés, lorsque p est connu.	Moyenne de la distribution d'échantillonnage d'une fréquence.
Calculer le pourcentage des échantillons de taille n simulés, pour lesquels la fréquence relative au caractère étudié appartient à l'intervalle donné $[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}]$ et comparer à une probabilité de 0,95. Exercer un regard critique sur des données statistiques en s'appuyant sur la probabilité précédente.	Intervalle de fluctuation.

Contenu du dossier :

- ☐ **Cours**
- ☐ **Exercices (Livre Chapitre 2 pages 23-32)**
- ☐ **Corrigé des exos**
- ☐ **Evaluation EM6**
- ☐ **Correction évaluation EM6**

PBP M6

Lancer un dé, lancer une pièce (pile ou face) avec une pièce ou encore tirer une boule dans une urne sont ***des expériences aléatoires***



Vocabulaire

Une expérience aléatoire est une expérience dont le résultat est lié au hasard.

L'issue est le résultat possible d'une expérience aléatoire. (exemple :lancer de pièces : 2 issues Pile ou Face)

Un évènement est un ensemble constitué d'issues. (exemple « obtenir des nombres pairs en lançant un dé : 3 issues possibles 2, 4, 6). Il peut être **élémentaire**, si l'évènement est constitué que d'une seule issue.

L'échantillon d'une expérience aléatoire est l'ensemble des résultats obtenus par répétition de l'expérience. Le nombre de résultats est **la taille de l'échantillon.**

Lorsque l'on cherche à évaluer les chances de réalisation d'un événement, on dit qu'on évalue **sa probabilité.** C'est la somme des probabilités des issues qui le composent. Lorsque toutes les issues possibles ont **la même** probabilité, on dit qu'elles sont **équiprobables.**

I. Expérience aléatoire

Activité 1

On lance un dé à jouer à 6 faces et souhaite savoir s'il est bien équilibré, c'est-à-dire si la probabilité d'apparition d'une face donnée, est la même, quelle que soit la face. Pour cela nous allons faire 2 expériences :

- *1ère : on lance dix fois le dé*
- *2ème : Faites la même expérience mais avec 150 lancers.*

***Noter** les résultats dans les tableaux ci-dessous. **Observer** les résultats et **conclure** si oui ou non le dé est bien équilibré.*

Faces	**1**	**2**	**3**	**4**	**5**	**6**	
Effectifs							**10**
Fréquences							**100**
Effectifs							**150**
Fréquences							**100**

.....

.....

.....

.....

Quelles sont les **issues** possibles à cette expérience ?

.....

Vous jouez aux petits chevaux, quel doit être **l'évènement** pour que vous puissiez commencer le jeu ?.....

Quelle est la taille de **l'échantillon dans cette expérience** ?.....

.....

Quelle est la probabilité d'obtenir

- un 2 ?.....
- un 4 ?.....
- un 6 ?.....

Quelle est la probabilité d'obtenir un nombre pair ?

.....

On peut simuler l'expérience précédente à l'aide d'une calculatrice ou d'un tableur (voir page 10).

Activité 2 : Simulation de 150 lancers de dé à la calculatrice.

Type casio (GRAPH 25+PRO)

Pour simuler les lancés on va utiliser le mode **STAT.**

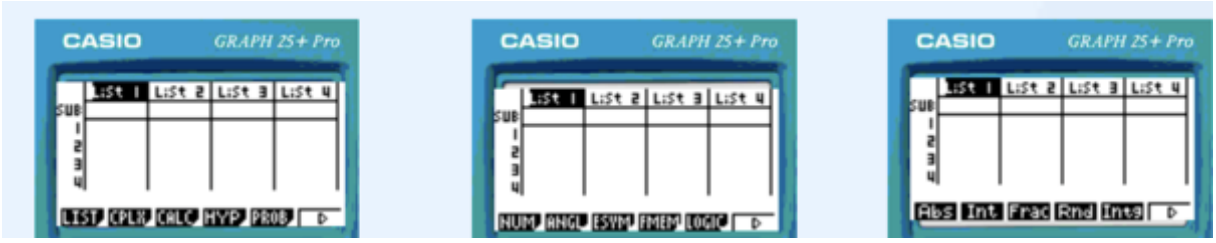
Nous allons créer une séquence de 150 nombres aléatoires compris entre 1 et 6 dans la **List 1.**

Avec le pavé directionnel **placer** le curseur sur **List 1** puis **sélectionner** alors

puis

(▶) et (NUM) puis

(Int) pour afficher **Int** en bas de l'écran**.**



Sélectionner

taper

puis

(PROB) et

(RAND) et

(LIST)

Taper alors 1 5 0) x 6) + 1

Et

Alors apparaît dans la colonne de **List1** une liste de 150 nombres aléatoires compris entre 1 et 6.
(*Formule de calcul pour générer une liste :* **Int(RanList#(150)×6)+1**)

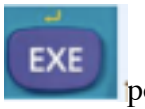
Nous allons trier la **List1** par ordre croissant.

Faire d'abord

2 fois puis **sélectionner** alors

(TOOL) et

(SRTA) puis **saisir**

et 

pour présiser le nombre de listes à trier.

Saisir

pour classer la première liste dans l'ordre croissant.

A l'aide de
descendre List 1 jusqu'à trouver un nombre 2 . **Noter** le rang du dernier 1. **Continuer** jusqu'à trouver un nombre 3. **Noter** le rang du dernier 2 et ainsi de suite. Il vous suffit de calculer l'effectif de chaque face et de compléter le tableau ci-après.

Faces	1	2	3	4	5	6	
Effectifs							150
Fréquences							100

II. Moyenne de fréquences sur une série d'échantillons

II.1. Observer, sur des séries d'échantillons de même taille, la moyenne des fréquences d'une valeur d'un caractère

Une fabrique produit en grande quantité des jouets en bois, de couleurs différentes, dont 60 % sont rouges. On simule sur tableur la prise de 6 échantillons aléatoires de 100 jouets dans la production. Voici les fréquences de jouets rouges obtenues :

Echantillon numéro	1	2	3	4	5	6
Fréquence de jouets rouges	0,68	0,58	0,7	0,72	0,7	0,52

La moyenne des fréquences est le rapport $\frac{Somme\ des\ fréquences}{Nombre\ d'échantillons}$

Activité 3

Cocher la case correspondant à la bonne réponse, et **compléter** s'il y a lieu.

1. **a)** La fréquence *p* de jouets rouges dans la production est égale à :

☐0,4 ☐0,6

b) La fréquence *f5* de jouets rouges obtenue avec l'échantillon numéro 5 est égale à :

☐0,4 ☐0,6 ☐Autre : .....

2. a) La moyenne des fréquences des 6 échantillons est égale à *p* :

☐Vrai ☐Faux car .....

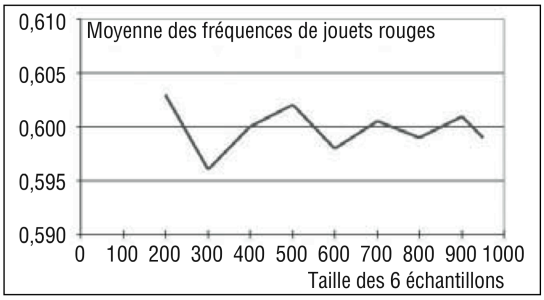
b) Avec de nouvelles séries de 6 échantillons aléatoires de taille 100, on constaterait que la moyenne des fréquences de jouets rouges fluctue :

☐Vrai ☐Faux

II.2. Visualiser la stabilisation de la moyenne des fréquences, lorsque la taille des échantillons augmente

On simule sur tableur la prise de 6 échantillons aléatoires de jouets, cette fois de taille 200, puis 6 échantillons aléatoires de taille 300, etc.

Le graphique indique l'évolution de la moyenne des fréquences de jouets rouges des 6 échantillons, selon leur taille.



Activité 4

Rayer l'encadré inexact**.**

a) La moyenne des fréquences de jouets rouges fluctue / ne fluctue pas selon la taille des 6 échantillons.

b) Lorsque la taille des 6 échantillons augmente, la fluctuation de la moyenne des 6 fréquences de jouets rouges est plus / moins grande.

c) Lorsque la taille des 6 échantillons devient très grande, cette moyenne se stabilise / ne se stabilise pas vers la fréquence p.

d) Le nombre p est / n'est pas la probabilité d'obtenir un jouet rouge lors du prélèvement aléatoire d'un jouet dans la production de l'entreprise.

III. Intervalle de fluctuation

III.1. Interpréter des fréquences calculées sur une population et des échantillons de cette population

En 2007, la population française comptait 63 578 000 personnes, dont 15 901 940 avaient moins de 20 ans. On simule sur tableur un prélèvement de 50 échantillons aléatoires de 200 personnes de la population française et on calcule pour chacun la fréquence des personnes de moins de 20 ans :

0,28	0,25	0,22	0,24	0,23	0,15	0,24	0,22	0,23	0,2
0,24	0,23	0,25	0,19	0,25	0,29	0,31	0,24	0,28	0,29
0,22	0,27	0,24	0,25	0,27	0,24	0,27	0,31	0,23	0,28
0,27	0,22	0,24	0,23	0,25	0,21	0,25	0,24	0,26	0,27
0,28	0,22	0,24	0,23	0,25	0,33	0,25	0,24	0,25	0,27

Activité 5

1. Cocher la case correspondant à la bonne réponse et **compléter.**

a) La fréquence p des moins de 20 ans dans la population française, arrondie à 0,01,

est 0,25:

☐Vrai ☐Faux car .....

b) On en déduit qu'en choisissant au hasard 100 personnes dans la population, exactement p parmi elles ont moins de 20 ans :

☐Vrai ☐Faux

2. **Compléter** le tableau**.**

Fréquence des moins de 20 ans	0,15	0,19	0,2	0,21	0,22	0,23	0,24
Nombre d'échantillons			1		5	6	10

Fréquence des moins de 20 ans	0,25	0,26	0,27	0,28	0,29	0,31	0,33
Nombre d'échantillons	9	1		4	2		

III.2. Déterminer l'intervalle de fluctuation des fréquences pour des échantillons de taille n et comprendre sa signification

La fréquence d'une valeur du caractère étudié dans une population est p. Lorsque $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$: la probabilité que la fréquence de cette valeur pour un échantillon aléatoire de taille n pris dans la population soit dans l'intervalle de fluctuation $\left[p - \frac{1}{\sqrt{n}};\ p + \frac{1}{\sqrt{n}}\right]$ est supérieure à 0,95.

Activité 6

On s'intéresse à la population française de l'activité 1.

1.Compléter. On a $p = 0{,}25$ et $n =$ .....**.**

Ainsi, $\left[p - \frac{1}{\sqrt{n}};\ p + \frac{1}{\sqrt{n}}\right] = \left[\ldots\ldots\ldots - \frac{1}{\sqrt{200}}\ ;\ \ldots\ldots\ldots + \frac{1}{\sqrt{200}}\right] \approx [\ldots\ldots ;\ \ldots\ldots]$

2. Rayer les encadrés inexacts et **compléter.**

Il y a échantillons 1 / 2 / 3 parmi les 50 pour lesquels la fréquence n'est pas dans l'intervalle de fluctuation. Pour ces 50 échantillons, le pourcentage des fréquences qui appartiennent à cet intervalle est / n'est pas supérieur à 95 %, car.....

III.3. Comment utiliser un intervalle de fluctuation ?

Méthode 1

Une valeur V du caractère étudié dans une population a pour fréquence p. On s'intéresse à un échantillon de taille n prélevé dans cette population.

Étape 1 Repérer ou calculer la fréquence p, puis la fréquence f de la valeur V du caractère, dans l'échantillon.

Étape 2 Déterminer l'intervalle de fluctuation $I = \left[p - \frac{1}{\sqrt{n}};\ p + \frac{1}{\sqrt{n}}\right]$

Étape 3 Conclure. Si la fréquence f, dans l'échantillon :

- Appartient à I, elle n'est pas « significativement différente » de p , au niveau 95 %
- N'appartient pas à I, elle est « significativement différente » de p, au niveau 95 %.

Exemple: Le taux de chômage dans la population active française fin 2008 était 7,7 %. A cette même période, pour deux villes françaises, on relevait :

Le taux de chômage de la ville 1 est-il significativement différent de celui de la population française, au niveau 95 % ? Et celui de la ville 2 ?

Solution

• *On s'intéresse à la ville 1.*

Étape 1 Les fréquences sont données dans l'énoncé :la fréquence p du nombre de chômeurs en France est $p = 0{,}077$; la fréquence $f1$ du nombre de chômeurs dans la ville 1 est égale à $f1 =$ ***0,089.***

	Ville 1	**Ville 2**
Effectif de la population active	5 530	98 100
Taux de chômage	8,9 %	8,9 %

Étape 2 On détermine l'intervalle de fluctuation, avec $p = 0{,}077$ et $n = 5\ 530$:

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

Étape 3 On conclut :

f_1 appartient à cet intervalle, donc le taux de chômage de la ville 1 est / n'est pas significativement différent de celui de la population française, au niveau 95 %.

• *On s'intéresse à la ville 2.*

Étape 1 Les fréquences sont données dans l'énoncé :

la fréquence p du nombre de chômeurs en France est $p =$ ..... ;

la fréquence f_2 du nombre de chômeurs dans la ville 2 est égale à $f_2 =$ ..... .

Étape 2 On détermine l'intervalle de fluctuation, avec $p =$ ..... et $n =$ ..... :

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

Étape 3 On conclut :

f_2 n'appartient pas à cet intervalle, donc le taux de chômage de la ville 2 est / n'est pas significativement différent de celui de la population française, au niveau 95 %.

(On peut donc être amené à s'interroger, et remarquer d'abord que la population active d'une ville ne constitue pas un échantillon aléatoire de la population active française.)

Exercices: ☐ **26 page 27** ☐ **27 page 27** ☐ **28 page 27** ☐ **29 page 28** ☐ **30 page 28**
☐ **31 page 28** ☐ **32 page 28** ☐ **33 page 28** ☐ **36 page 29** ☐**37 page 29**
☐ **42 page31** ☐ **43 page 32** ☐ **44 page 32** ☐ **46 page 32** ☐ **47 page 32**
☐ **48 page 32** ☐ **49 page 32**

Simulation d'un lancer de dé

(Utilisation du tableur EXCEL)

PARTIE A : Simulation de 10 lancers par 3 personnes

1. Simulation de 10 lancers

1.1.Entrer en **A1** la formule
=ALEA.ENTRE.BORNES(1;6)

1.2.Cliquer sur la cellule **A1**, puis **recopier** la formule jusqu'à la cellule **A10** à l'aide de la poignée de recopie (se placer en bas à droite de la cellule et **obtenir** une croix noire).

1.3. Quelle est la taille de l'échantillon ainsi obtenu ?

.....

.....

A1 =ALEA.ENTRE.BORNES(1;6)

	A	B	C	D	E
1	3				
2	3				
3	1				
4	2				
5	5				
6	6				
7	4				
8	5				
9	6				
10	1				
11					
12					

2. Comparaison de 3 échantillons de même taille.

2.1.Recopier la formule dans la plage de cellules allant de B1 à C10 par le même procédé.

2.2.Préparer le tableau ci-dessous dans les colonnes E,F,G,……K de la manière suivante:

A12

	A	B	C	D	E	F	G	H	I	J	K	L
1	1	5	4									
2	1	1	3		FACE	1	2	3	4	5	6	
3	1	1	4		Fréquence(A)							
4	4	6	4		Fréquence(B)							
5	4	2	5		Fréquence(C)							
6	1	6	6									
7	1	6	5		Etendue							
8	2	1	2									
9	1	3	1									
10	2	4	3									
11												
12												
13												

2.3. En F3 entrer la formule **=NB.SI($A$1:$A$10;F2)/10**

(Cette fonction permet de compter le nombre de fois qu'apparaît le contenu de la cellule F2 (c'est à dire 1) dans la plage de données correspondant à la colonne A).

2.4. Quelle formule doit-on alors entrer dans les cellules F4 et F5 pour obtenir la fréquence d'apparition du 1 :

- dans la colonne B ? .....
- dans la colonne C ? .....

2.5. En F7 entrer la formule = MAX(F3 :F5)-MIN(F3 :F5)(cette fonction permet de calculer l'étendue)

2.6.Compléter le tableau.

F3 : =NB.SI($A$1:$A$10;F2)/10

	A	B	C	D	E	F	G	H	I	J	K
1	5	3	4								
2	4	6	5		FACE	1	2	3	4	5	6
3	6	4	2		Fréquence(A)	0,20	0,10	0,10	0,30	0,10	0,20
4	4	6	2		Fréquence(B)	0,30	0,00	0,20	0,20	0,00	0,30
5	1	4	1		Fréquence(C)	0,20	0,20	0,00	0,20	0,30	0,10
6	3	1	1								
7	4	1	6		étendue	0,10	0,20	0,20	0,10	0,30	0,20
8	1	3	4								
9	6	6	5								
10	2	1	5								
11											

3. Représentation graphique

3.1.Sélectionner la plage de cellules allant de **F2 à K5**. Cliquer sur l'onglet **« Insertion »** puis **sélectionner** puis

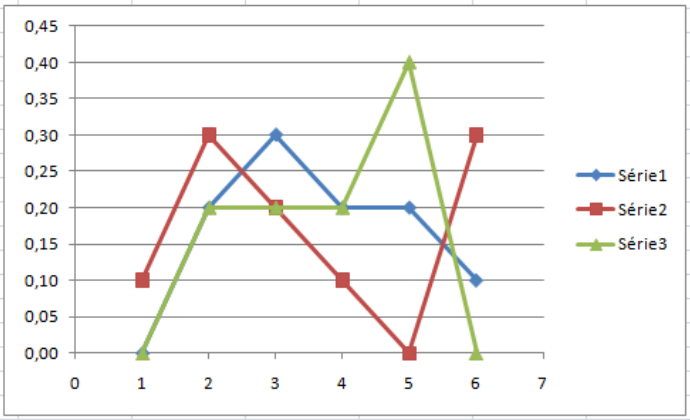
3.2. Une fois le graphique obtenu, **appuyer** plusieurs fois sur **F9** pour obtenir d'autres simulations et **observer** les résultats.

3.3. Quel est le plus grand écart de valeurs observé ? .....

.....

A12

	A	B	C	D	E	F	G	H	I	J	K
1	5	1	3								
2	4	3	4		FACE	1	2	3	4	5	6
3	2	6	5		Fréquence(A)	0,00	0,20	0,30	0,20	0,20	0,10
4	3	2	4		Fréquence(B)	0,10	0,30	0,20	0,10	0,00	0,30
5	3	4	5		Fréquence(C)	0,00	0,20	0,20	0,20	0,40	0,00
6	4	3	2								
7	6	6	3		étendue	0,10	0,10	0,10	0,10	0,40	0,30
8	2	2	2								
9	5	2	5								
10	3	6	5								
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											



PARTIE B :

1. Echantillons de taille 100

1.1. Sur la feuille 2 du classeur, **créer** trois échantillons de taille 100 dans les colonnes A,B,C, puis **reporter** les fréquences dans un tableau similaire à celui de la partie A.

N.B : Pour plus de rapidité, faire un copier-coller de la feuille 1 puis apporter les modifications nécessaires

1.2.Créer un graphique permettant de comparer les trois distributions de fréquence.

1.3.Simuler d'autres séries à l'aide de la touche F9.

1.4. Quel est le plus grand écart de valeurs observé ? .....

.....

2. **Echantillons de taille 1000**

2.1. Sur la feuille 3 du classeur, **reprendre** le même procédé pour des échantillons de taille 1000.

2.2. Quel est le plus grand écart de valeurs observé ?.....

.....

PARTIE C :

1. **Insérer** une 4ème feuille dans le classeur et y recopier les trois graphiques obtenus précédemment.

1.1. Dans quel cas la fluctuation d'échantillonnage est elle la plus importante ?

.....

.....

1.2. Que peut-on en conclure quant à la taille de l'échantillon ?

.....

.....

.....

1.3. Vers quelle valeur théorique se rapproche t-on lorsque la taille de l'échantillon est grande ?

.....

.....

.....

.....