MATHEMATIQUES 2017-2018 LOIC CHAPELLIER lp2i, Jaunay- Marigny	**Thème 3 : Probabilités** SECONDE Plan de travail : Comment prendre une décision à partir d'observations ? ***Comment prendre une décision à partir d'observations ?***

Introduction – Sondage

Le lien

Parité hommes/femmes :

les disparités ne s'effacent pas dans

les plus grandes entreprises

Votre mission :

Rédiger en quelques lignes le critère que doit respecter une entreprise de 500 employés pour respecter la parité

Lire le cours :

Document : COprobabilités 2017- 2018

I – Comment prendre une décision à partir d'observations - Intervalles de fluctuation

Annexe technique :

Faire les exercices A1 à A6

Se contrôler avec la correction

Simulations:

Effectuer la simulation 1 : entreprise A

Effectuer la simulation 2 : entreprise B

A l'aide de ces simulations, quelle est l'entreprise , qui respecte le mieux la parité ?

Lire la synthèse et faire l'exercice 1

Prise en main de Python :

Vidéo 1 : effectuer une simulation

Vidéo 2 : se servir de mathplotlib

Quand vous êtes prêts !

Demander à votre professeur:

Socrative : intervalle de fluctuation

Simulation 1 : entreprise A

On considère une entreprise A composée de 40 employées dont 17 femmes.

On peut considérer que l'ensemble de la population active set composée de 50% de femmes et 50% d'hommes.

(voir le document : http://www.insee.fr/fr/statistiques/1288331?sommaire=1288404)

A partir d'un sac composé de types de billes noires et blanches en nombre égal, effectuer plusieurs tirages de 40 billes avec remise, noter les résultats dans le tableau suivant.

tirage	1	2	3	4	5	6	7	8	9
Nombres de billes noires									

Si on fait l'hypothèse que l'entreprise A a respecté la parité, la proportion de femmes dans l'entreprise est-elle surprenante ou bien semble t-elle conforme à ce qui était attendu ?

Simulation 2 : entreprise B

On considère une entreprise B composée de 2500 employés dont 1150 femmes.

Comme il serait fastidieux de refaire pour l'entreprise B la même expérience, on peut utiliser un programme qui va simuler très rapidement les 2500 tirages.

Partie A :

Voici par exemple un **algorithme** permettant de simuler *n* tirages d'une urne contenant autant de boules noires que de boules blanches.

```
from math import *
import matplotlib.pyplot as plt
import random

print(' ')
print(' ------ Simulation tirage ------ ')
print(' ')

n=int(input("nombre de tirages"))

compteur = 0# compte le nombre de boules noires
for i in range(n+1): #répète n fois le tirage
    a=random.randint(0,1)
    compteur = compteur+a
print(compteur)
```

a) Etudier l'algorithme précédent puis recopier le dans thonny pour effectuer des simulations de 2500 tirages.

b) Avec plusieurs simulations, pouvez-vous obtenir un nombre de boules noires inférieur à 1190

c) ***Si on fait l'hypothèse que l'entreprise B a respecté la parité, la proportion de femmes dans l'entreprise est-elle surprenante ou bien semble t-elle conforme à ce qui était attendu ?***

Partie B :
Une simulation consiste à reproduire, l'expérience faite en Partie A afin d'avoir un grand nombre de données. Bien sûr, il faut connaître la vraie proportion içi, 0,5 !
1) Expliquez pourquoi la fonction suivante permet de simuler n tirages dans le sac en question et de calculer la fréquence.

```
def frequence(n):
    compteur=0 #nombre de boules rouges
    for i in range(n+1):
        a=random.randint(0,1)
        compteur=compteur+a
    return compteur/n #la fréquence de boules rouges...
```

2) Dans la question suivante, on cherchera à construire une représentation graphique de la situation. On utilisera pour cela la commande : `import matplotlib.pyplot as plt`

Le morceau de programme suivant utilise des « listes » pour afficher sur un graphique les fréquences calculées pour des échantillons de taille 10,20,30,...,100.

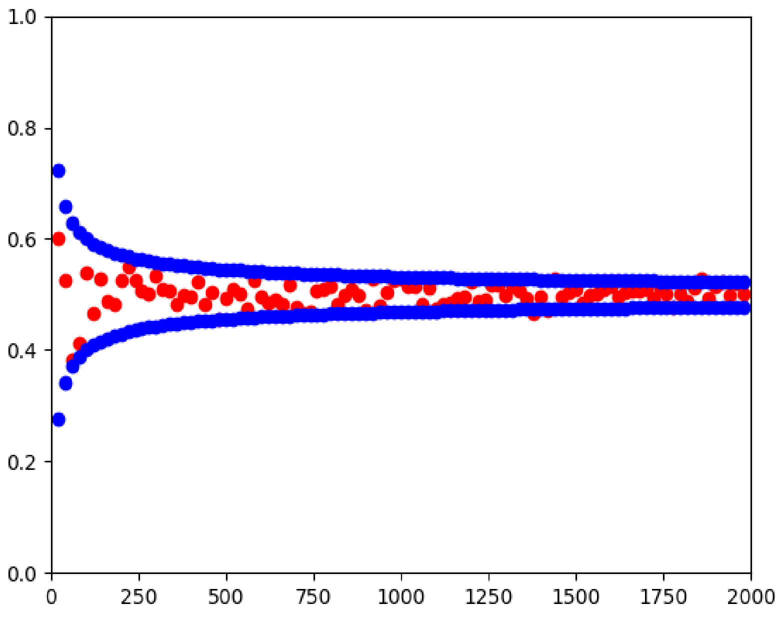
```
L=[10*k for k in range(1,10)]
F=[]
for i in L:
    F.append(frequence(i))
plt.plot(L,F,'ro')
plt.xlim(0,100)
plt.ylim(0,1)
plt.show()
```

Modifiez le programme afin qu'il affiche les fréquences calculées pour n = 20, 40, 60, 80,..., 2000.

3) Recopier le programme à la suite de la fonction fréquence et observer la répartition, que constatez-vous ?

Synthèse des simulations

Ouvrir le programme trompette et lancer la simulation, on obtient :



Conclusion de l'expérience : Plus précisément, les mathématiciens ont établi la propriété suivante :

Propriété :

Si $n \geqslant 25$ ***et*** $0{,}2 \leqslant p \leqslant 0{,}8$, il y a 95 % de chance que la fréquence observée f soit comprise dans l'intervalle $\left[p-\frac{1}{\sqrt{n}}\,;p+\frac{1}{\sqrt{n}}\right]$. Cet intervalle est appelé « **intervalle de fluctuation** » **de *f* au seuil 0,95**.

Illustration de la propriété : http://www.lp2i-poitiers.fr/doc/math/tirage3.php?affiche_propriete

A l'aide de cette propriété, quelle entreprise respecte le mieux la parité homme-femme ? Justifiez votre réponse.

Application :
Dans une entreprise C, il y a 5000 employées dont 2440 femmes.
Pouvez-vous conclure si l'entreprise respecte ou non la parité ?

Exercice 1 :

Dans la population d'enfants de 11 à 14 ans d'un département français le pourcentage d'enfants ayant déjà eu une crise d'asthme dans leur vie est de 21%. Un médecin d'une ville de ce département est surpris du nombre important d'enfants le consultant ayant des crises d'asthme et en informe les services sanitaires. Ceux–ci décident d'entreprendre une étude et d'évaluer la proportion d'enfants de 11 à 14 ans ayant déjà eu des crises d'asthme. Ils sélectionnent de manière aléatoire 100 jeunes de 11 à 14 ans de la ville. Cette étude a dénombré 29 jeunes ayant déjà eu des crises d'asthme.

A votre avis, le médecin a-t-il raison de s'inquiéter ?

Annexe technique : Intervalles de fluctuation:

Exercice A1
Au casino Belle-Vue, sur 2 500 lancers d'un dé à 6 faces, 1312 ont donné un nombre pair.

1) On fait l'hypothèse que les dés ne sont pas truqués : il y a alors autant de chances d'obtenir un nombre pair qu'un nombre impair. On assimile alors la situation à un échantillon de taille $n=2500$, avec une probabilité $p=0{,}5$ d'obtenir un nombre pair. Déterminer l'intervalle de fluctuation au seuil de 95%.

2) Peut-on conclure, avec un niveau de confiance de 95 % que les dés sont truqués ?

Exercice A2
On a vendu à un grossiste 50 000 appareils électroniques en certifiant qu'au moins 80 % ne présentent aucun défaut de fonctionnement. En prélevant au hasard 400 appareils et en les testant, le grossiste s'aperçoit que seulement 70 % n'ont pas de défaut de fonctionnement.

1) On assimile la situation à un échantillon de taille $n=400$, avec une probabilité $p=0{,}8$ d'appareils ne présentant aucun défaut de fonctionnement. Déterminer l'intervalle de fluctuation au seuil de 95%.

2) Peut-on conclure, avec un niveau de confiance de 95 % que le grossiste a été trompé ?

Exercice A3
Dans une population, la proportion d'un certain caractère est 0,45.

Dans un premier groupe de 150 personnes de cette population, on en compte 60 qui possèdent le caractère.

1) On assimile la situation à un échantillon de taille $n=150$, avec une probabilité $p=0{,}45$.

Déterminer l'intervalle de fluctuation au seuil de 95%.

Peut-on penser que le caractère est mal représenté dans ce groupe ?

2) Dans un deuxième groupe de 500 personnes de cette population, on en compte 195 qui possèdent le caractère. Peut-on penser que le caractère est mal représenté dans ce groupe ?

Exercice A4
Pour un certain type de cancer, une variété de souris présente une taux parfaitement connu de 20% de cancers spontanés. On utilise un traitement sur un lot de 100 souris prises au hasard. A la fin du traitement on décèle 14 souris atteintes du cancer.

Peut-on alors considérer que ce traitement est efficace ?

Exercice A5
Dans une usine automobile, on contrôle les défauts de peinture de type « grains ponctuels sur le capot ». Lorsque le processus est bien contrôlé, on a 20 % de ce type de défauts.

Lors d'un contrôle de 50 véhicules, on observe que 13 véhicules présentent des défauts. Faut-il s'inquiéter ?

Exercice A6
Une machine ensache des œufs de Pâques: elle choisit au hasard des œufs dans une cuve qui contient, à part égale, uniquement des œufs au chocolat noir et des œufs au chocolat blanc. Il y a 50 000 œufs dans la cuve. Chaque sachet contient 100 œufs.

1. Expliquer pourquoi le sachet peut être assimilé à un échantillon.

2. Nicolas ouvre un paquet qui contient 42 œufs au chocolat noir et 58 au chocolat blanc. Peut-on considérer que la machine soit bien réglée ? Justifier.

Annexe technique : Intervalles de fluctuation:
Eléments de correction

Exercice A1
Au casino Belle-Vue, sur 2 500 lancers d'un dé à 6 faces, 1312 ont donné un nombre pair.

1) On fait l'hypothèse que les dés ne sont pas truqués : il y a alors autant de chances d'obtenir un nombre pair qu'un nombre impair. On assimile alors la situation à un échantillon de taille $n=2500$, avec une probabilité $p=0{,}5$ d'obtenir un nombre pair. Déterminer l'intervalle de fluctuation au seuil de 95%.

$n=2500\geqslant 25$ et $0{,}2\leqslant p=0{,}5\leqslant 0{,}8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0{,}5-\frac{1}{\sqrt{2500}};0{,}5+\frac{1}{\sqrt{2500}}\right[=]0{,}48;0{,}52[$

2) $f=\frac{1312}{2500}\approx 0{,}5248$ qui n'appartient pas à l'intervalle de fluctuation donc, on peut conclure, avec un niveau de confiance de 95 % que les dés sont truqués.

Exercice A2
On a vendu à un grossiste 50 000 appareils électroniques en certifiant qu'au moins 80 % ne présentent aucun défaut de fonctionnement. En prélevant au hasard 400 appareils et en les testant, le grossiste s'aperçoit que seulement 70 % n'ont pas de défaut de fonctionnement.

1) On assimile la situation à un échantillon de taille $n=400$, avec une probabilité $p=0{,}8$ d'appareils ne présentant aucun défaut de fonctionnement. Déterminer l'intervalle de fluctuation au seuil de 95%.

$n=400\geqslant 25$ et $0{,}2\leqslant p=0{,}8\leqslant 0{,}8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0{,}8-\frac{1}{\sqrt{400}};0{,}8+\frac{1}{\sqrt{400}}\right[=]0{,}75;0{,}85[$

2) $f=0{,}7$ n'appartient pas à l'intervalle de fluctuation donc, on peut conclure, avec un niveau de confiance de 95 % que le grossiste a été trompé.

Exercice A3
Dans une population, la proportion d'un certain caractère est 0,45.

Dans un premier groupe de 150 personnes de cette population, on en compte 60 qui possèdent le caractère.

1) On assimile la situation à un échantillon de taille $n=150$, avec une probabilité $p=0{,}45$.

$n=150\geqslant 25$ et $0{,}2\leqslant p=0{,}45\leqslant 0{,}8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0{,}45-\frac{1}{\sqrt{150}};0{,}45+\frac{1}{\sqrt{150}}\right[=]0{,}36835;0{,}53165[$

$f=\frac{60}{150}=0,4$ appartient à l'intervalle de fluctuation donc, on ne peut pas penser que le caractère est mal représenté dans ce groupe.

2) Dans un deuxième groupe de 500 personnes de cette population, on en compte 195 qui possèdent le caractère. Peut-on penser que le caractère est mal représenté dans ce groupe ?

$n=500\geqslant 25$ et $0,2\leqslant p=0,45\leqslant 0,8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0,45-\frac{1}{\sqrt{500}};0,45+\frac{1}{\sqrt{500}}\right[=]0,4053;0,4947[$

$f=\frac{195}{500}=0,39$ n'appartient pas à l'intervalle de fluctuation donc, on peut conclure, avec un niveau de confiance que le caractère est mal représenté dans ce groupe.

Exercice A4

Pour un certain type de cancer, une variété de souris présente une taux parfaitement connu de 20% de cancers spontanés. On utilise un traitement sur un lot de 100 souris prises au hasard. A la fin du traitement on décèle 14 souris atteintes du cancer.

$n=100\geqslant 25$ et $0,2\leqslant p=0,20\leqslant 0,8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0,20-\frac{1}{\sqrt{100}};0,20+\frac{1}{\sqrt{100}}\right[=]0,1;0,3[$

$f=\frac{14}{100}=0,14$ appartient à l'intervalle de fluctuation donc, on ne peut pas considérer que ce traitement est efficace.

Exercice A5

Dans une usine automobile, on contrôle les défauts de peinture de type « grains ponctuels sur le capot ». Lorsque le processus est bien contrôlé, on a 20 % de ce type de défauts.

Lors d'un contrôle de 50 véhicules, on observe que 13 véhicules présentent des défauts. Faut-il s'inquiéter ?

$n=50\geqslant 25$ et $0,2\leqslant p=0,2\leqslant 0,8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0,20-\frac{1}{\sqrt{50}};0,20+\frac{1}{\sqrt{50}}\right[=]0,0586;0,3414[$

$f=\frac{13}{50}=0,26$ appartient à l'intervalle de fluctuation donc, il n'y a pas lieu de s'inquiéter.

Exercice A6

Une machine ensache des œufs de Pâques: elle choisit au hasard des œufs dans une cuve qui contient, à part égale, uniquement des œufs au chocolat noir et des œufs au chocolat blanc. Il y a 50 000 œufs dans la cuve. Chaque sachet contient 100 œufs.

1. Expliquer pourquoi le sachet peut être assimilé à un échantillon.

On peut considérer que le sachet est un échantillon car les œufs sont choisis au hasard et il y a un grand nombre d'œufs dans la cuve.

2. Nicolas ouvre un paquet qui contient 42 œufs au chocolat noir et 58 au chocolat blanc. Peut-on considérer que la machine soit bien réglée ? Justifier.

$n=100\geqslant 25$ et $0,2\leqslant p=0,5\leqslant 0,8$

L'intervalle de fluctuation au seuil de 95% est donc, $\left]0,50-\frac{1}{\sqrt{100}};0,50+\frac{1}{\sqrt{100}}\right[=]0,4;0,6[$

$f=\frac{42}{100}=0,42$ appartient à l'intervalle de fluctuation donc, on ne peut pas considérer que la machine ne soit pas bien réglée.