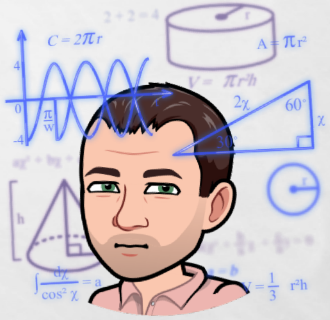
# THÈME : STATISTIQUES ET PROBABILITÉS

NOTRE DAME DU VOEU
LYCÉE

## SÉQUENCE 11 : ÉPREUVE ET LOI DE BERNOULLI



| | |
|---|---|
| Je dois connaître | • Probabilité associée à la répétition d'épreuves aléatoires identiques et indépendantes de Bernoulli.<br>• Loi de Bernoulli (0,1) de paramètre $p$, espérance. |
| Je dois Savoir-faire | • Reconnaître une situation aléatoire modélisée par une loi de Bernoulli.<br>• Représenter par un arbre de probabilités la répétition de $n$ épreuves aléatoires identiques et indépendantes de Bernoulli avec $n \leq 4$ afin de calculer des probabilités.<br>• Simuler *N* échantillons de taille *n* d'une loi de Bernoulli et représenter les fréquences observées des *1* par un histogramme ou un nuage de points.<br>• Interpréter sur des exemples la distance à $p$ de la fréquence observée des 1 dans un échantillon de taille $n$ d'une loi de Bernoulli de paramètre $p$. |

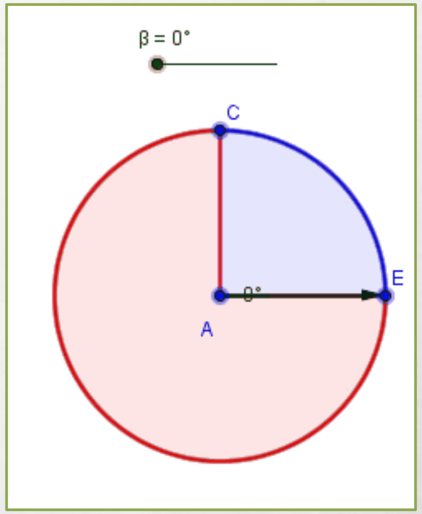
1

MEVEL CHRISTOPHE

## 1°) Épreuve de Bernoulli

**Un exemple pour introduire:**

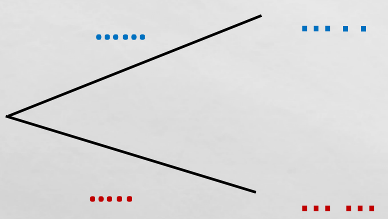

On fait tourner la roue de loterie présentée ci-contre:
On obtient la couleur **« rouge »** avec la probabilité de ……
et la couleur **« bleu »** avec la probabilité de ……




Le joueur est **gagnant** lorsque la flèche s'arrête sur **la zone bleue** comme sur la figure ci-contre.
On décide de noter **S (comme succès)** cette éventualité et
de noter **E (comme échec)** l'éventualité contraire c'est-à-dire « la flèche tombe sur **la zone rouge** ».

Par conséquent, $P(S) = \ldots\ldots.\ et\ P(E) = \ldots\ldots..$

**Arbre pondéré associé au jeu :**




**Généralisation:**
**Une expérience à deux issues, …………….. ou ………………, est appelée une ……………………………………………………………………………………**

- **Le paramètre $p$ est ………………………………………………….**
- **La probabilité de l'échec E est noté … … ..**

**Autrement dit: P(S) = p et P(E) = q = 1- p.**

## 2°) Loi de Bernoulli

**Définition:**
**On appelle ………………………………………………………………………………………………………………….**
**$p$ est le paramètre de cette loi de Bernoulli.**

**Autrement dit, on a définit une variable aléatoire $X$ tel que $P(X=1)=P(S)=p$ $et$ $P(X=0)=P(E)=1-p$.**

**Propriété:**
**La loi de Bernoulli est la loi de probabilité ci-contre:**

| $x_i$ | 0 | 1 |
|---|---|---|
| $P(X=x_i)$ | … … . | … … |

**Théorème:**
**L’espérance d’une loi de Bernoulli est … … … … … … … . ..**

**Exemple:**
La loi de Bernoulli de l’exemple page 2 est :

Son espérance $E(X)=$ …, résultat que l’on retrouve aussi grâce à sa loi de probabilité.

| $x_i$ | …. | …. |
|---|---|---|
| $P(X=x_i)$ | ….. | …. |

## 3°) Répétition d'épreuves aléatoires de Bernoulli

**Lorsque l'on répète de manière identique et indépendante des épreuves de Bernoulli, on peut modéliser la situation par un arbre de probabilités.**

**Exemple :**

À partir du jeu de loterie exposé à la page 2 du cours, nous pourrions vouloir représenter la situation lorsque l'on joue trois fois de suite.

Arbre pondéré modélisant la situation

(S,S,S)

(……….)

(……….)

(……….)

(……….)

(……….)

(……….)

(……….)

Les issues

**Propriétés: (Principe multiplicatif)**

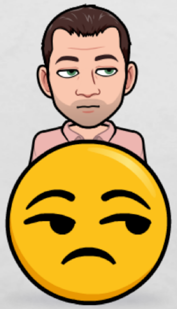
- **Pour calculer la probabilité d'un chemin, on effectue ……………………………………………………………………………………………………**
- **Pour calculer la probabilité d'obtenir $n$ succès ($n \leq 4$), ……………………………………………………………………………………………………**
  **……………………………………………………………………………………………………**

Dans cet exemple, on peut obtenir exactement 2 succès avec les issues $(SSE), (SES)$ et $(ESS)$. Soit **…** chemins possibles.

De plus, $\boldsymbol{P(SSE)} = \; \ldots \times \; \ldots \times \; \ldots = \; \ldots$ .

On en conclut que la probabilité d'avoir deux succès dans le jeu de loterie à trois lancers est donné par le calcul : … $\times$ … $=$ …

On peut donc définir la variable aléatoire $X$ qui donne le nombre de côté « pile » obtenues à l'issue des trois lancers.
Dans ce cas, $\boldsymbol{X} = \{\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots\}$.



La loi de probabilité associée à ce jeu est :

| $x_i$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $P(X = x_i)$ | …. | …. | …. | …. |

Son espérance $\boldsymbol{E(X)} = \; \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots . .$

## 4°) Simulations et échantillons

### a) Définition et simulation en langage Python

**Lorsque l'on répète $n$ fois de façon indépendante, une expérience aléatoire à deux issues, on obtient une série de $n$ résultats que l'on appelle ……………………………………………………………………………………………………………**

Cette fonction piece() permet de simuler le lancer d'une pièce avec comme sortie **1 pour PILE** et **0 pour FACE**.

```
In [1]: from random import *

        def piece():
          x=randint(0,1)        #randint(a,b) retourne aléatoirement un entier entre a et b
          return x

        print (piece())

0
```

Cette fonction jeu2() permet de déterminer le nombre de 1 qui est apparu lorsque l'on a lancé 10 fois la pièce.
Ce nombre est stocké dans la **variable S**

```
In [286]: from random import *

          def jeu2():
            S=0
            for i in range (10):
              x=randint (0,1)
              S=S+x
            return (S/10)

          print(jeu2())

0.4
```

Cette fonction nous retourne en sortie **la fréquence observée de 1** dans notre échantillon de 10 lancers de pièce.

Cette fonction echantillon1() simule 10 lancers de pièce et retourne la **liste des résultats**.

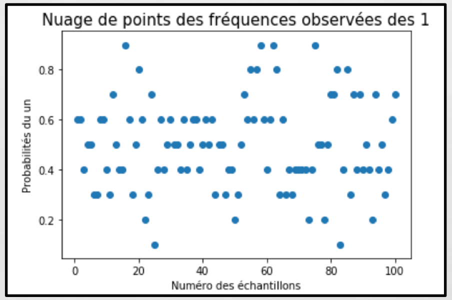
```
In [4]: from random import *

        def echantillon1():
          L=[]                  #Création d'une liste vide
          for i in range (10):
            L.append(piece())       #Remplir la liste L
          return (L)

        print (echantillon1())

[1, 0, 0, 0, 0, 0, 1, 0, 0, 0]
```

## b) Étude des échantillons

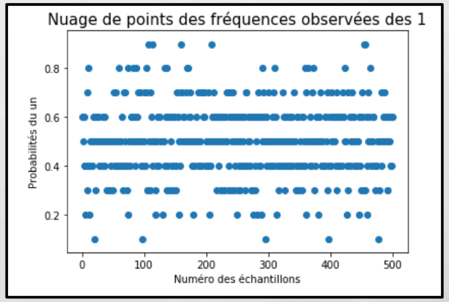
**Si on réalise plusieurs échantillons de même taille, …………………………………………………………………………………………
Cela s’appelle …………………………………………………………………**

**Exemple :**

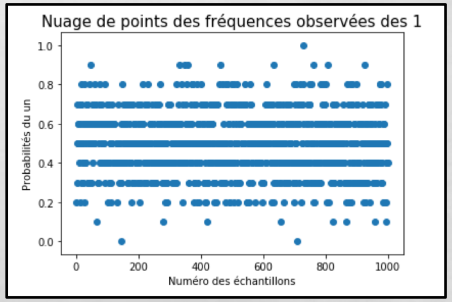
Ci-contre les nuages de points représentant la fréquence observée de 1 (succès) avec des simulations permettant d’obtenir 100, 500 puis 1000 échantillons de 10 lancers de pièces avec comme succès l’apparition du 1 modélisant le côté PILE.



**100 échantillons**



**500 échantillons**

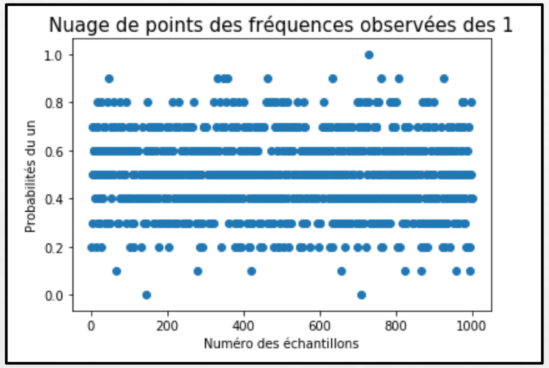


**1000 échantillons**

**On voit dans les trois cas que la fréquence observée varie entre 0,1 et 0,9 alors que la proportion théorique de Pile est 0.5.**

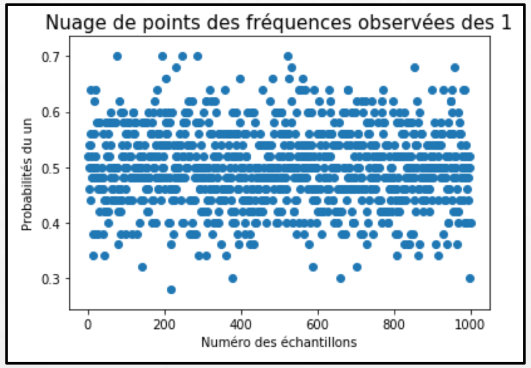
Jouons maintenant plutôt sur la taille des 1000 échantillons.

**Nuage de points**



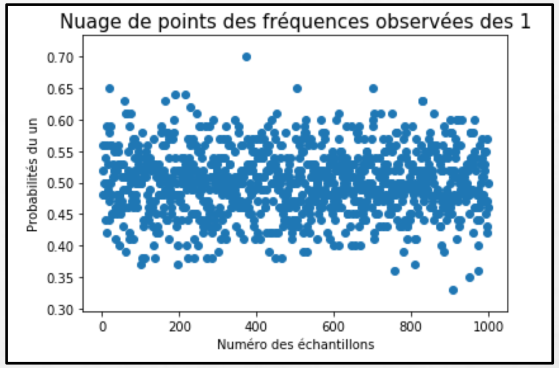

Fluctuation

**10 lancers**




Fluctuation

**50 lancers**

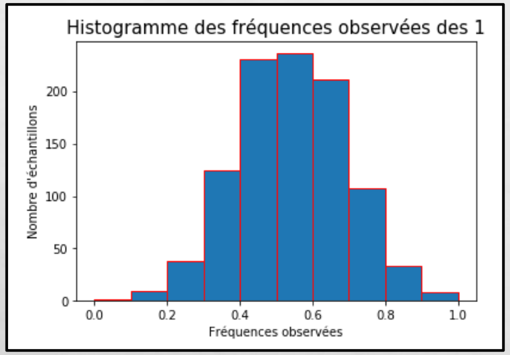



Fluctuation

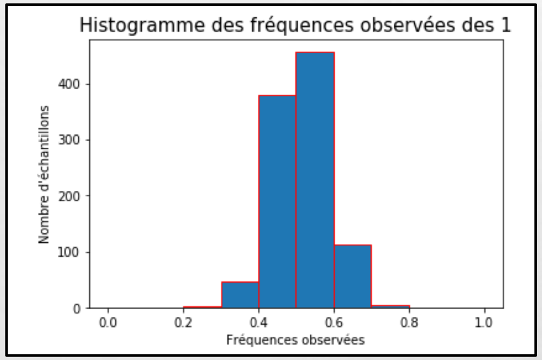
**100 lancers**

PAS POSSIBLE

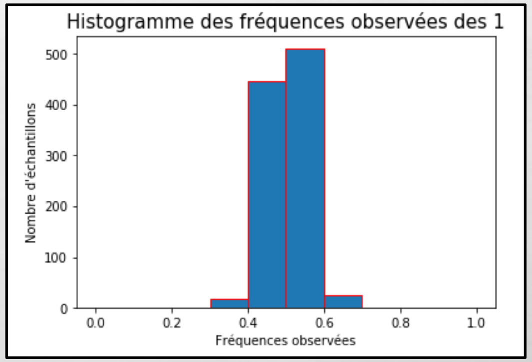
**Histogramme**




Fluctuation




Fluctuation




Fluctuation

**Conséquence:**
**Plus la taille des échantillons est …………, plus le phénomène de fluctuations …………………: les fréquences se rapprochent alors de la proportion théorique. (Ici, 0.5)**

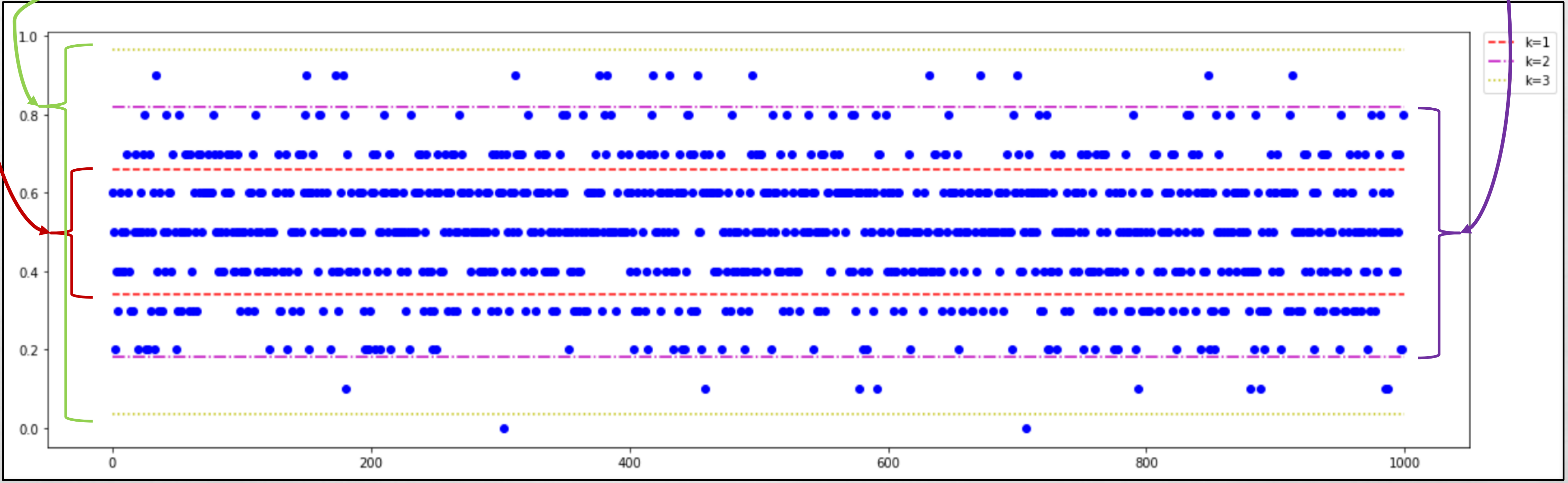
**Interprétation des résultats:**
**Pour étudier une simulation de $N$ échantillons de taille $n$ et on note $s$ l'écart-type de la série des fréquences obtenues et $p$ la proportion théorique.**

**On s'intéresse aux pourcentages des fréquences contenues dans les intervalles**
- $[p - s\,;p + s]$
- $[p - 2s\,;p + 2s]$
- $[p - 3s\,;p + 3s]$.

**Exemple :** Reprenons la simulation de 1000 échantillons de taille 10 correspondant au nombre de lancers de pièces.
Python nous fournit l'écart-type $s \cong 0,156$ et $p = 0,5$.
Donc: $[p - s\,;p + s] = [0,5 - 0,156\,;0,5 + 0,156] = [\,0,344\,;0,656]$
$[p - 2s\,;p + 2s] = [0,5 - 2\times0,156\,;0,5 + 2\times0,156] = [\,0,188\,;0,812]$
$[p - 3s\,;p + 3s] = [0,5 - 3\times0,156\,;0,5 + 3\times0,156] = [\,0,032\,;0,968]$

La simulation avec Python nous fournit les données suivantes:

```
L'intervalle [p-1*s,p+1*s] est :  [0.34, 0.66]
La proportion de points entre [p-1*s,p+1*s] est  0.659
L'intervalle [p-2*s,p+2*s] est :  [0.172, 0.828]
La proportion de points entre [p-2*s,p+2*s] est  0.98
L'intervalle [p-3*s,p+3*s] est :  [0.025, 0.975]
La proportion de points entre [p-3*s,p+3*s] est  0.998
```



**Interprétation de ces résultats:**

On observe qu'environ $65{,}9\%$ des fréquences sont comprises entre $0{,}34$ $et$ $0{,}66$ : c'est donc ……………………………………………………….

A contrario, si on se place à une distance supérieure à $2s$ de $p$, on a seulement $2\%$ (proportion à $2s$ de 0,98) des fréquences : C'est …………………………………………………………….

**Propriété (admise) :**

**L'écart type $s$ de la série des fréquences est de l'ordre de $\frac{1}{2\sqrt{n}}$ où $n$ est la taille de l'échantillon.**

**Exemples:**

À des simulations illustrant la série des fréquences observées des 1 dans 1000 échantillons de taille n d'une loi de Bernoulli, on constate que .

1°) Pour $n = 10, s = 0{,}156$ et $\frac{1}{2\sqrt{n}} = \ldots\ldots\ldots \approx \ldots\ldots\ldots.$ sont très proches.

2°) Pour $n = 50, s = 0{,}074$ et $\frac{1}{2\sqrt{n}} = \ldots\ldots.. \approx \ldots\ldots\ldots..$ sont très proches.

3°) Pour $n = 100, s = 0{,}05$ et $\frac{1}{2\sqrt{n}} = \ldots\ldots\ldots \approx \ldots\ldots..$ sont égaux.