# I Statistiques avec R-cran

## 1 intervalle de fluctuation $\quad I=\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$

### 1.1 Énoncé

- on souhaite à l'aide d'un logiciel, simuler le lancer de 400 pièces de monnaie équilibrées, puis calculer la fréquence d'apparition de «pile». On veut ensuite répéter 2000 fois cette expérience.
- Au résultat «pile», on associe la valeur 1 et à «face» le résultat 0.
  On génère donc 400 nombres aléatoires 0 ou 1 et on note la fréquence des 1.
  On répète 2000 fois cette expérience. On obtiens une série statistique de fréquence d'apparition du nombre 1.
- On cherche alors la proportion des valeurs situées à l'intérieur de l'intervalle $I=\left[p-\dfrac{1}{\sqrt{n}};p+\dfrac{1}{\sqrt{n}}\right]$ où $p$ est la probabilité théorique et $n$ la taille de l'échantillon.

Voici un algorithme pour simuler cette situation.

**Entrées** :
**début**
  **pour** j allant de 1 jusqu'à 2000 faire **faire**
    Créer une liste de 400 nombres au hasard de 0 ou 1.
    Calculer la fréquence d'apparition de 1.
    Noter cette fréquence dans une liste.
  Calculer les bornes $a=p-\dfrac{1}{\sqrt{n}}$ et $b=p+\dfrac{1}{\sqrt{n}}$.
  Calculer le nombre de valeurs dans l'intervalle $[a;b]$ puis la fréquence.
**Sorties** : La liste de 2000 fréquences d'apparition de 1

*Remarque : Les listes commencent à 1 avec R-cran c'est pourquoi $j$ va de 1 à 2000.*

### 1.2 Expérimentation avec R-cran

On aura donc ici $n=400$, $p=\dfrac{1}{2}$ et $I=\left[\dfrac{1}{2}-\dfrac{1}{\sqrt{400}};\dfrac{1}{2}+\dfrac{1}{\sqrt{400}}\right]=[0,45;0,55]$

#### Réalisation d'un échantillon de taille 400

Ouvrez R-cran (R250xp au lycée)

```
x=sample(0:1,400,re=T)      # On génère les 400 nombres 0 ou 1 dans la variable x.
x                           # Pour visualiser les 400 nombres.
table(x)                    # Pour visualiser le tableau des effectifs.
sum(x==1)                   # pour compter le nombre de 1
sum(x==1)/400               # Pour avoir la fréquence des 1.
table(x)/400                # Pour visualiser le tableau des fréquences.
```

#### Simulation

Nous allons donc réaliser les 2000 simulations.

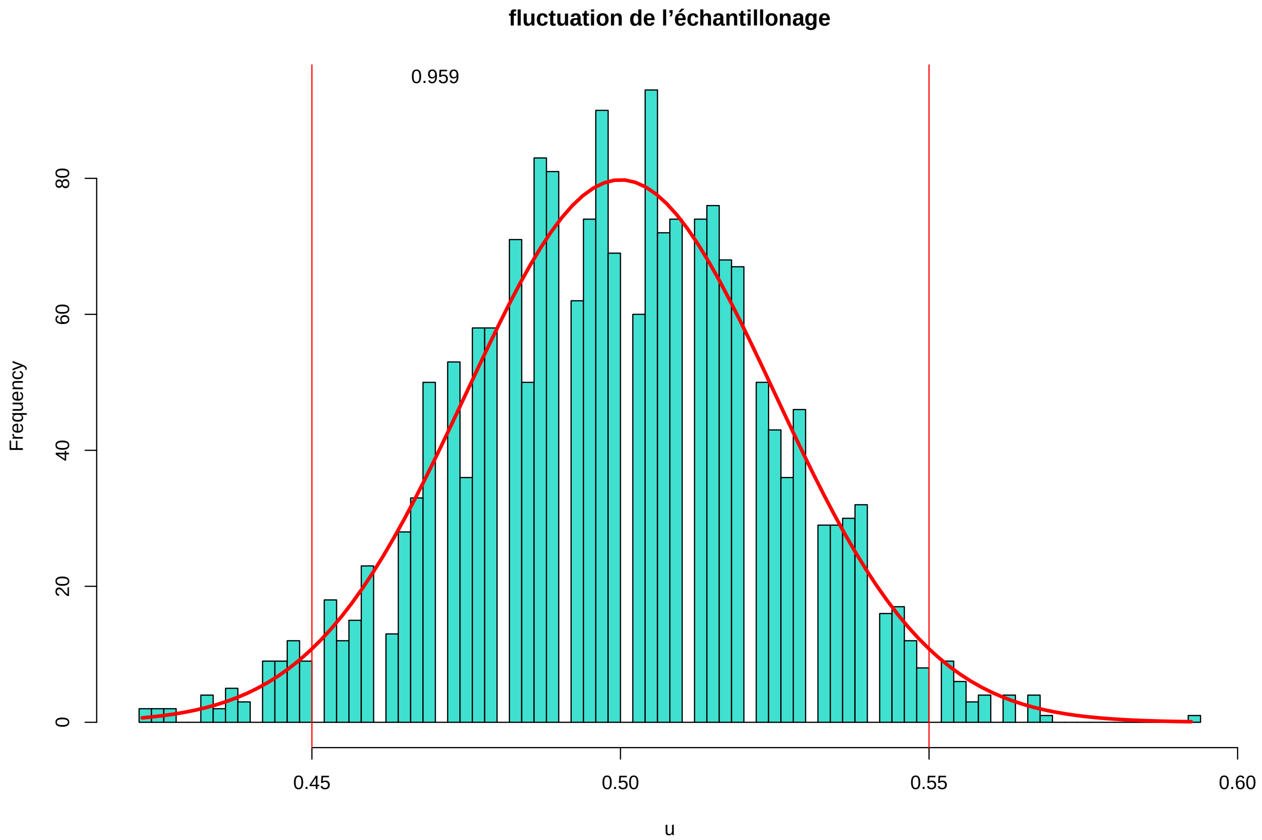
```
u=c()                        # Pour initialiser la série u.
for(i in 1:2000){a=ceiling(2*runif(400))-1; u[i]=sum(a==1)/400}
```

On a donc une liste $u$ des 2 000 valeurs de fréquences d'apparition du nombre 1.

### Utilisation de la simulation

```
a=1/2-1/sqrt(400)    # On calcule les deux bornes
b=1/2+1/sqrt(400)
hist(u,101)          # Pour un histogramme de u avec 100 classes
#On peut aussi mettre de la couler et un titre
hist(u,101,col="turquoise",main="fluctuation de l'echantillonage")
abline(v=a,col="red") # tracer les deux droites verticales
abline(v=b,col="red")
sum(u>=a & u<=b)     # On compte le nombre de valeurs dans l'intervalle I=[a;b]
sum(u>=a & u<=b)/2000 # Pour obtenir la fréquence
```

On trouve ... ... % des valeurs entre 0,45 et 0,55 donc il y a bien au moins 95% des valeurs situées à l'intérieur l'intervalle $\left[\frac{1}{2}-\frac{1}{\sqrt{400}};\frac{1}{2}+\frac{1}{\sqrt{400}}\right]$



# II Statistiques avec Xcas

## 1 intervalle de fluctuation $I=\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$

### 1.1 Énoncé

A l'aide d'un logiciel , on souhaite simuler le lancer de 400 pièces de monnaie équilibrées, puis calculer la fréquence d'apparition de «pile». On veut ensuite répéter 2000 fois cette expérience.

Au résultat «pile», on associe la valeur 1 et à «face» le résultat 0.

On génère donc 400 nombres aléatoires 0 ou 1 et on note la fréquence des 1.

On répète 2000 fois cette expérience. On obtiens une série statistique de fréquence d'apparition du nombre 1.

On cherche alors la proportion des valeurs situées à l'intérieur de l'intervalle $I=\left[p-\frac{1}{\sqrt{n}};p+\frac{1}{\sqrt{n}}\right]$ où $p$ est la probabilité théorique et $n$ la taille de l'échantillon.

Voici un algorithme pour simuler cette situation.

**Entrées** :
**début**
    **pour** j allant de 0 jusqu'à 1999 faire **faire**
        Créer une liste de 400 nombres au hasard de 0 ou 1.
        Calculer la fréquence d'apparition de 1.
        Noter cette fréquence dans une liste.
    Calculer les bornes $a=p-\frac{1}{\sqrt{n}}$ et $b=p+\frac{1}{\sqrt{n}}$.
    Calculer le nombre de valeurs dans l'intervalle $[a;b]$ puis la fréquence.
**Sorties** : La liste de 2000 fréquences d'apparition de 1

*Remarque : Les listes commencent à 0, c'est pourquoi j va de 0 à 1999 et non de 1 à 2000.*

### 1.2 Expérimentation avec Xcas

On aura donc ici $n=400$, $p=\frac{1}{2}$ et $I=\left[\frac{1}{2}-\frac{1}{\sqrt{400}};\frac{1}{2}+\frac{1}{\sqrt{400}}\right]=[0,45;0,55]$

#### Réalisation d'un échantillon de taille 400

Ouvrez Xcas

```
x:=randvector(400,'rand(2)');      // On génère les 400 nombres dans la variable x.
count_eq(1,x);                     // pour compter le nombre de 1
count_eq(1,x)/400;                 // Pour avoir la fréquence des 1.
count_eq(1,x)/400.0;               // Pour avoir le résultat sous forme décimal.
```

#### Simulation

Nous allons donc réaliser le programme.

Ouvrez un environnement programme (alt+p)

```
u:=[];                        // Pour initialiser la série u.
pour j de 0 jusque 1999 faire // une liste commence à 0
x:=randvector(400,'rand(2)'); // On génère 400 nombres aléatoirement 0 ou 1
u[j]:=count_eq(1,x)/400.0;    // On cherche la fréquence d'apparition du nombre 1
fpour;
```

Compiler le programme (F9)

**Utilisation de la simulation**

Ouvrez un environnement de graphique (alt+g)

```
f:=max(u)-min(u);            // On calcule l'étendue
a:=1/2-1/sqrt(400)           // On calcule les deux bornes
b:=1/2+1/sqrt(400)
classes(u,0,f/30)            // donne la série statistique regroupée en 30 classes
histogram(classes(u,0,f/30)); //Pour un histogramme de u avec 30 classes
purge(x)                     // On a déjà utilisé la variable x, c'est pour
droite(x=a);droite(x=b)      // tracer les deux droites verticales
2000-count_inf(a,u)-count_sup(b,u) // On compte le nombre de valeurs
                                  // dans l'intervalle I=[a;b]
(2000-count_inf(a,u)-count_sup(b,u))/2000 // Pour obtenir la fréquence
```

On trouve . . . . . . % des valeurs entre 0,45 et 0,55 donc il y a bien au moins 95% des valeurs situées à l'intérieur l'intervalle $\left[\frac{1}{2}-\frac{1}{\sqrt{400}};\frac{1}{2}+\frac{1}{\sqrt{400}}\right]$