RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

TP Simulations variables aléatoires réelles

NB : Le code rassemble tous les exercices du TD.
Pour lancer le code : se rendre à partir du terminal à l'emplacement où se trouve le fichier téléchargé, taper dans le terminal « ./main.sh ».
Pour voir le graphique après avoir fini de lancer le code : se rendre dans result. Le graphique rendu par le programme correspond au fichier graphe.jpeg.

Exercice 1

1. Loi de Bernoulli de paramètre p

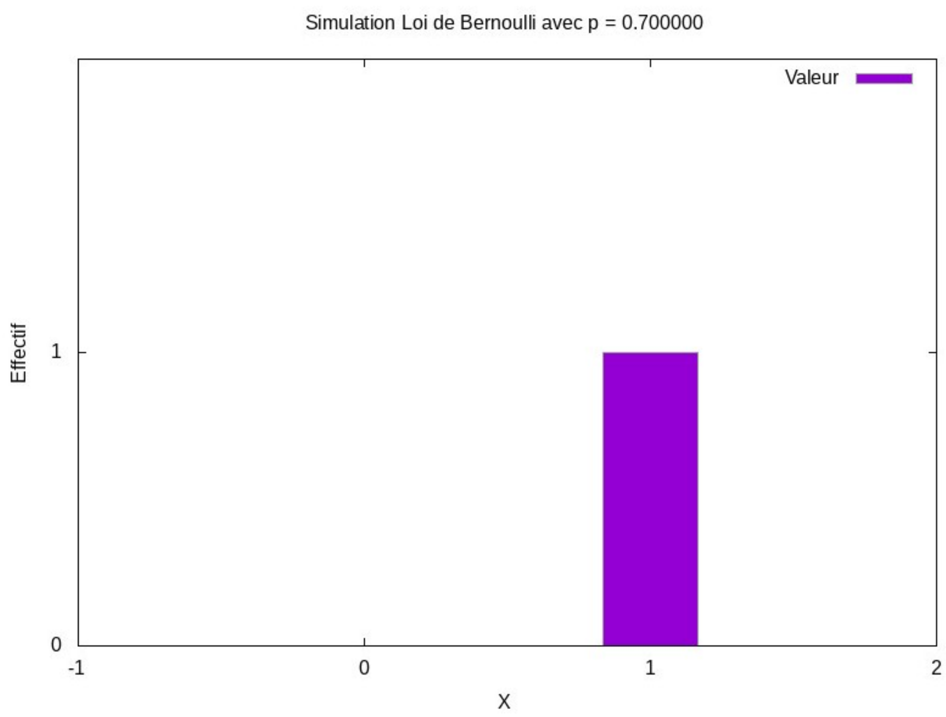
Approche théorique
On sait que $E_{th}(X) = p$
et $Var_{th}(X)=p(1-p)$

On choisit de prendre p=0.7
on a donc $E_{th}(X)=0.7$ et $Var_{th}(X)=0.21$

Approche expérimentale
D'un point de vue expérimental, on approximera toujours l'espérance par la moyenne expérimentale et la variance par

$$s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2$$

A partir du code, on effectue un tirage aléatoire avec une probabilité p=0.7.
On obtient le graphique suivant :



RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

Etant donné que p=0.7, on a de fortes chandes de tomber sur un succès (c'est-à-dire X=1). Effectivement, on obtient X=1 avec la simulation.

Dans le code, on répète plusieurs fois Bernoulli (on effectue donc une loi Binomiale codée à la question 2 pour n=1) pour voir l'espérance et la variance moyenne obtenue. On choisit de répéter l'expérience 1000 fois (n=1000). On récupère les résultats de nos tirages.

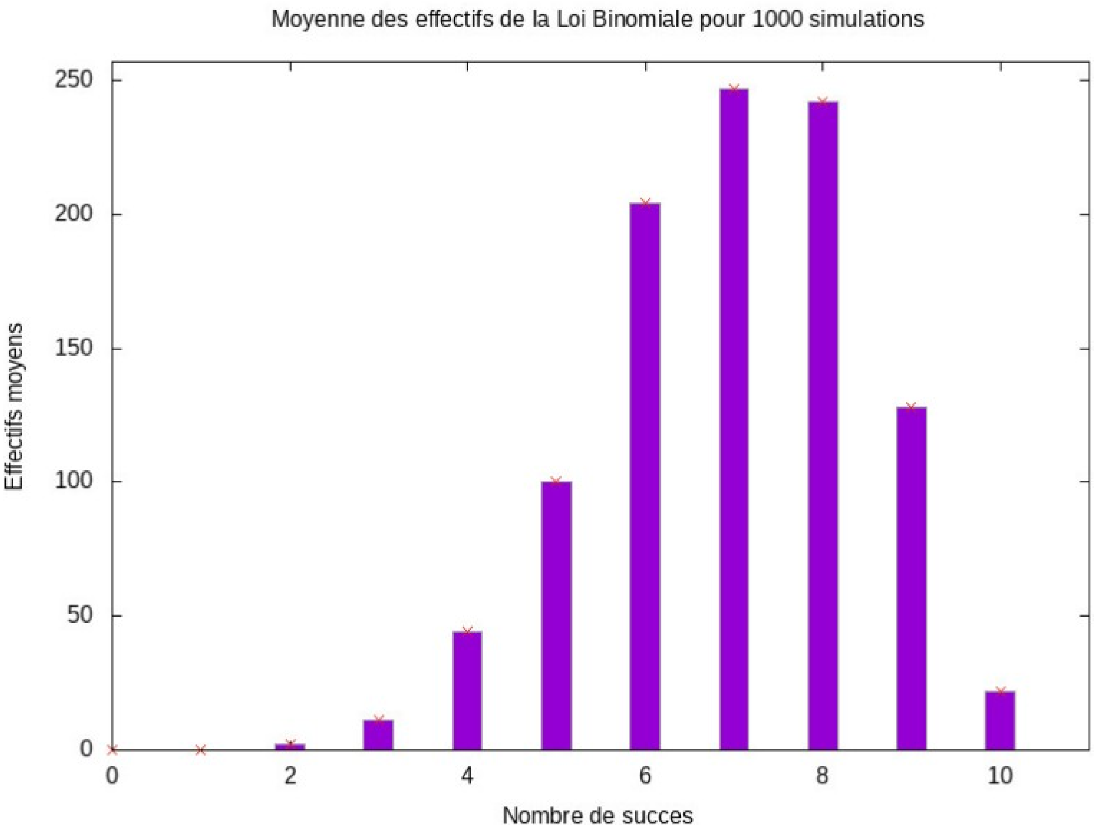
On obtient $E_{exp}(X) = 0*0.31+ 1* 0.69$ soit une espérance environ égale à la valeur théorique.
$Var_{exp}(X)= (1/999)*(690*(1-0.69)^2+310*(-0.69)^2) = 0.2141$ soit environ la valeur théorique.

2. Loi de Binomiale de paramètres (n, p)

<u>Approche théorique</u>
On sait que $E_{th}(X) = np$
et $Var_{th}(X) = np(1-p)$

On choisit de prendre p=0.7 et n=10
On va répéter cette expérience 1000 fois.
On a donc E(X)=7 et $Var_{th}(X)=2.1$

<u>Approche expérimentale</u>
On choisit ici de tracer les moyennes de effectifs de la loi binomiale pour 1000 simulations (pour chaque simulation, n=10).



On obtient les valeurs suivantes :

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

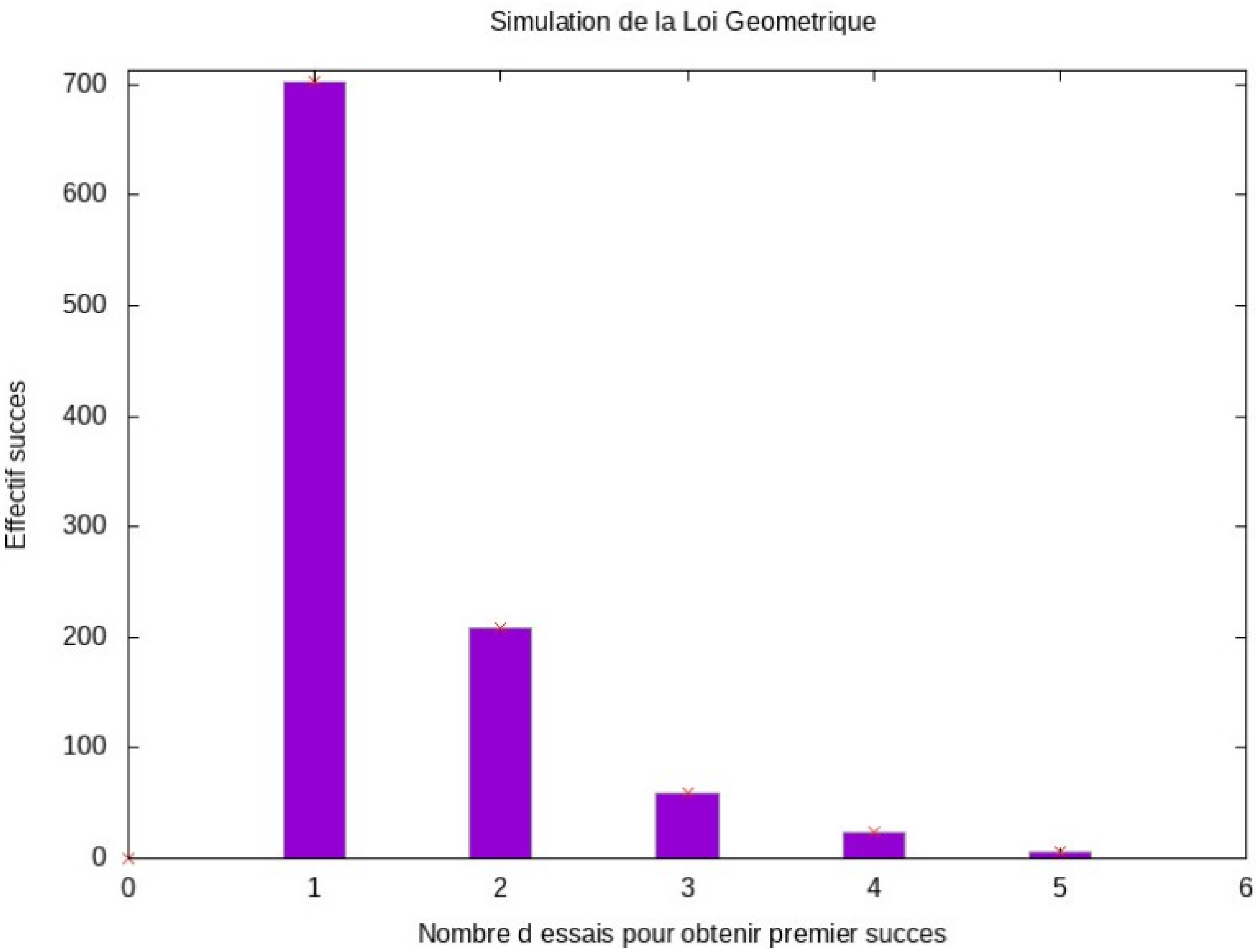
X =	0	1	2	3	4	5	6	7	8	9	10
effectif	0	0	2	11	44	100	204	247	242	128	22

Donc $E_{exp}(X)=6.932$ ce qui est très proche de la valeur théorique $E_{th}(X)$. On dira que $E_{exp}(X)=6.9$
$Var_{exp}(X) = (1/999)*(2*(2-6.9)^2+11*(3-6.9)^2+44*(4-6.9)^2+100*(5-6.9)^2+204*(6-6.9)^2+247*(7-6.9)^2+242*(8-6.9)^2+128*(9-6.9)^2+22*(10-6.9)^2) = 2.18$
On retrouve quasiment la valeur de la variance théorique.

3. Loi géométrique de paramètre p

On prendra p=0.7
<u>Approche théorique</u>
$E_{th}(X) = 1/p = 1.4$
$V_{th}(X) = (1-p)/p^2 = 0.6$

<u>Approche expérimentale</u>
On choisit de tracer les moyennes de effectifs de la loi géométrique pour 1000 simulations.



On obtient les résultats suivants :

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

X =	1	2	3	4	5
effectif	703	208	59	24	6

D’où l’espérance expérimentale :
$E_{exp}(X) = (1/1000)*(703+2*208+3*59+4*24+5*6) = 1.422$ soit environ 1.4

On obtient aussi une variance très proche de la valeur théorique :
$Var_{exp}(X) = (1/999)*(703*(1-1.4)^2+208*(2-1.4)^2+59*(3-1.4)^2+24*(4-1.4)^2+6*(5-1.4)^2) = 0.579$ soit environ 0.58

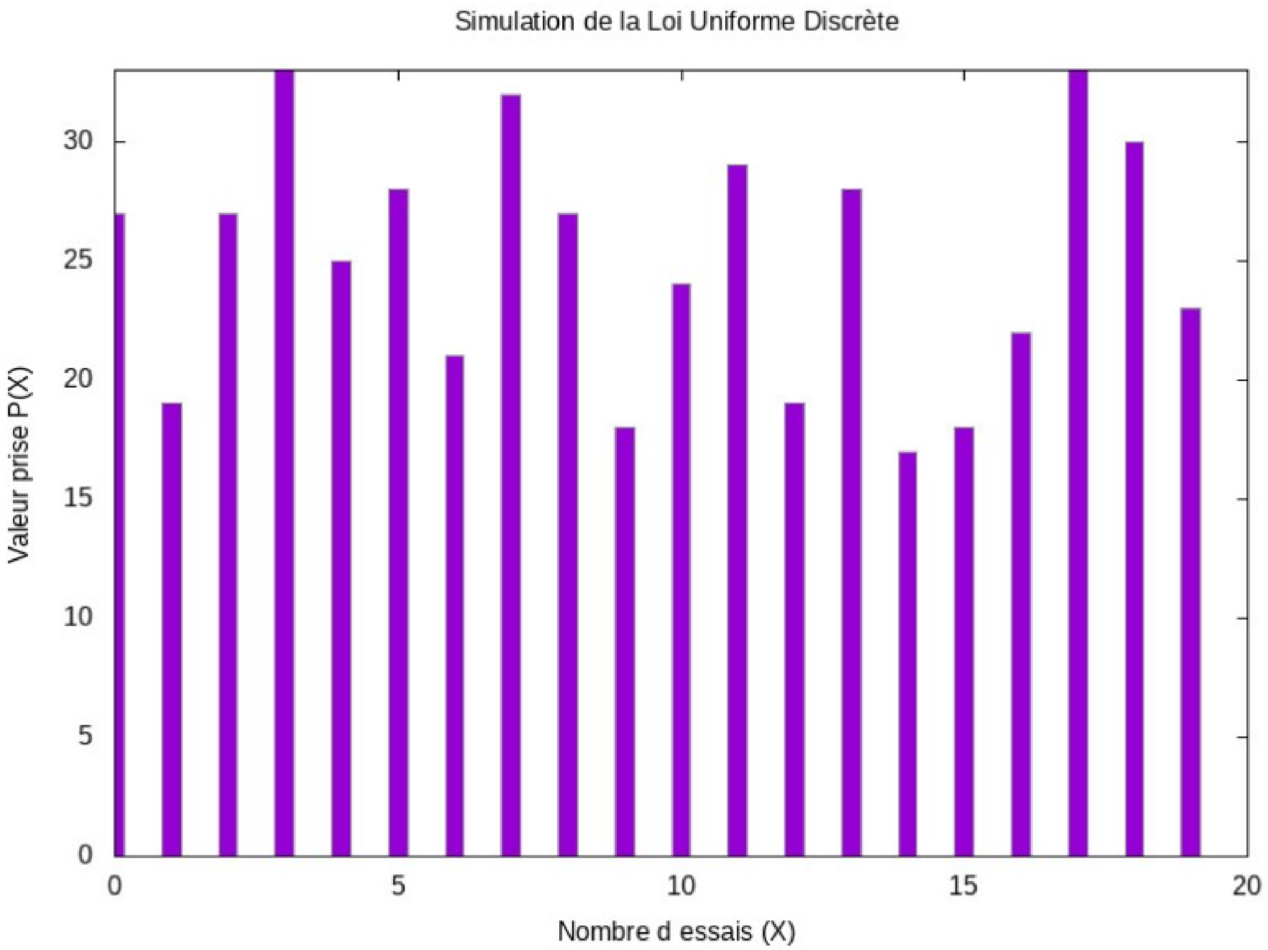
4. Loi uniforme sur les entiers de 1 à k avec k>=2.
On prendra k=20

<u>Approche théorique</u>

$$E_{th}(X) = (1/20)* \sum_{i=1}^{20} i = 10.5$$

$$Var_{th}(X) = (1/20)* \sum_{i=1}^{20} i^2 - (E_{th}(X))^2 = 33.25$$

<u>Approche expérimentale</u>
On répète la loi uniforme 500 fois.



RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

On obtient les résultats suivants :

X =	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
effectif	27	19	27	33	25	28	21	32	27	18	24	29	19	28	17	18	22	33	30	23

$E_{exp}(X) = 10.378$
On considerera que $E_{exp}(X) = 10.4$
$Var_{exp}(X) = 34.09$

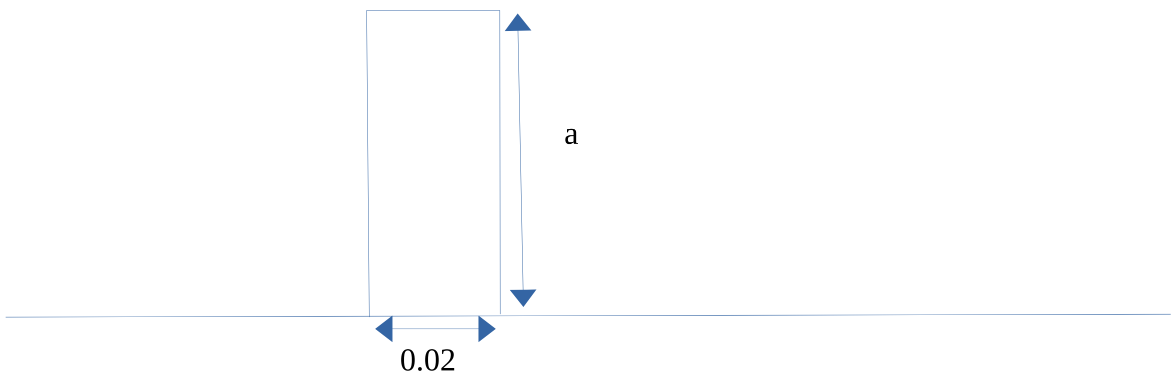
Les valeurs expérimentales correspondent aux valeurs théoriques.

4. Loi uniforme sur [-1, 1].

<u>Approche théorique</u>
$E_{th}(X) = (-1+1)/2 = 0$
$V_{th}(X) = 2^2/12 = 1/3$

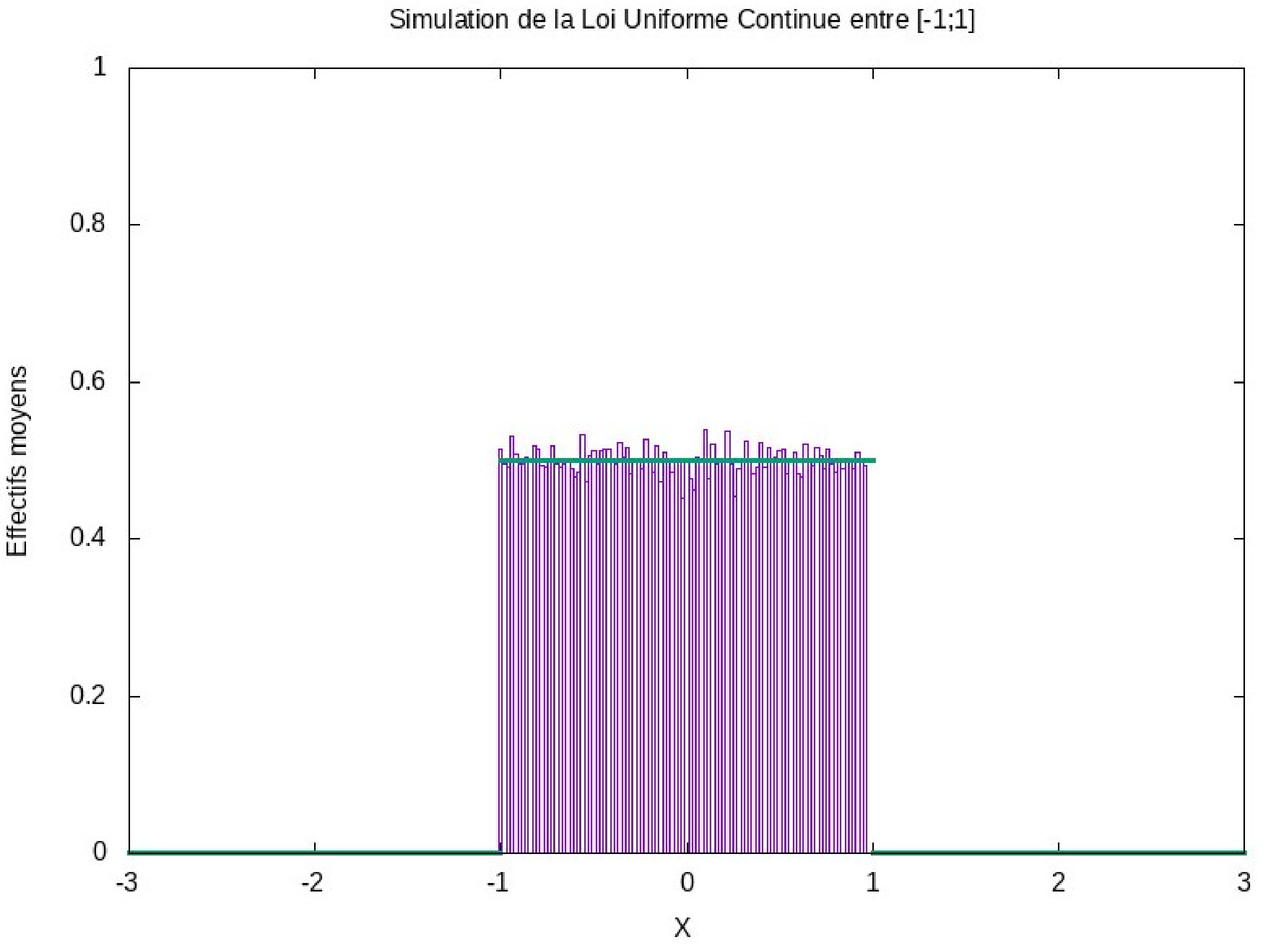
<u>Approche expérimentale</u>
Pour la représentation graphique, on doit faire un histogramme associé à la densité : pour cela, on simule un certain nombre de fois la loi uniforme sur [-1,1] (on le fait nbExp=100 000 fois dans le programme). Ensuite, on sépare l'intervalle [-1, 1] en plusieurs sous intervalles espacés de 0,02. (Par exemple, on a [-1, -0.98], [-0.98, -0.96] etc)
On calcule le nombre de VAR suivant la loi uniforme qui entre dans chacun des sous-intervalles. De cette façon on obtient des effectifs moyens. (effectif de l'intervalle / nbExp)

Pour obtenir une cohérence sur le graphique entre la densité théorique et les effectifs moyens, on choisit de faire en sorte que les rectangles de l'histogramme possèdent une aire égale à l'effectif moyen.
Sachant que la taille des intervalles est de 0.02 on fait le calcul suivant :
0.02 * a = effectif moyen



On obtient ainsi le graphique suivant :

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess



En vert, on a représenté la densité théorique de la loi uniforme sui [-1, 1] et en violet l'histogramme Etant donné que nbExp est très grand, on calcule l'espérance et la variance directement dans le code (voir l'affichage du code quand on lance la loi uniforme sur [-1, 1]).

On ajoute dans le code le tableau resultUni qui récupère les simulations de la loi uniforme sur [-1, 1] et on l'utilise pour calculer l'espérance et la variance :

```
//calculs de l'esperance et de la variance
for (int i=0 ; i<nbExp ; i++){
    esperance += resultUni[i];

}
esperance /= nbExp;

for (int i=0;i<nbExp;i++){
    variance += pow((resultUni[i]-esperance), 2);

}
variance /= (nbExp-1);
printf("esperance : %lf\n", esperance);
printf("variance : %lf\n", variance);
```

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

Pour le graphique obtenu ci-dessus, on obtient des résultats très proches des valeurs théoriques :
esperance : -0.001609
variance : 0.333840

Exercice 2

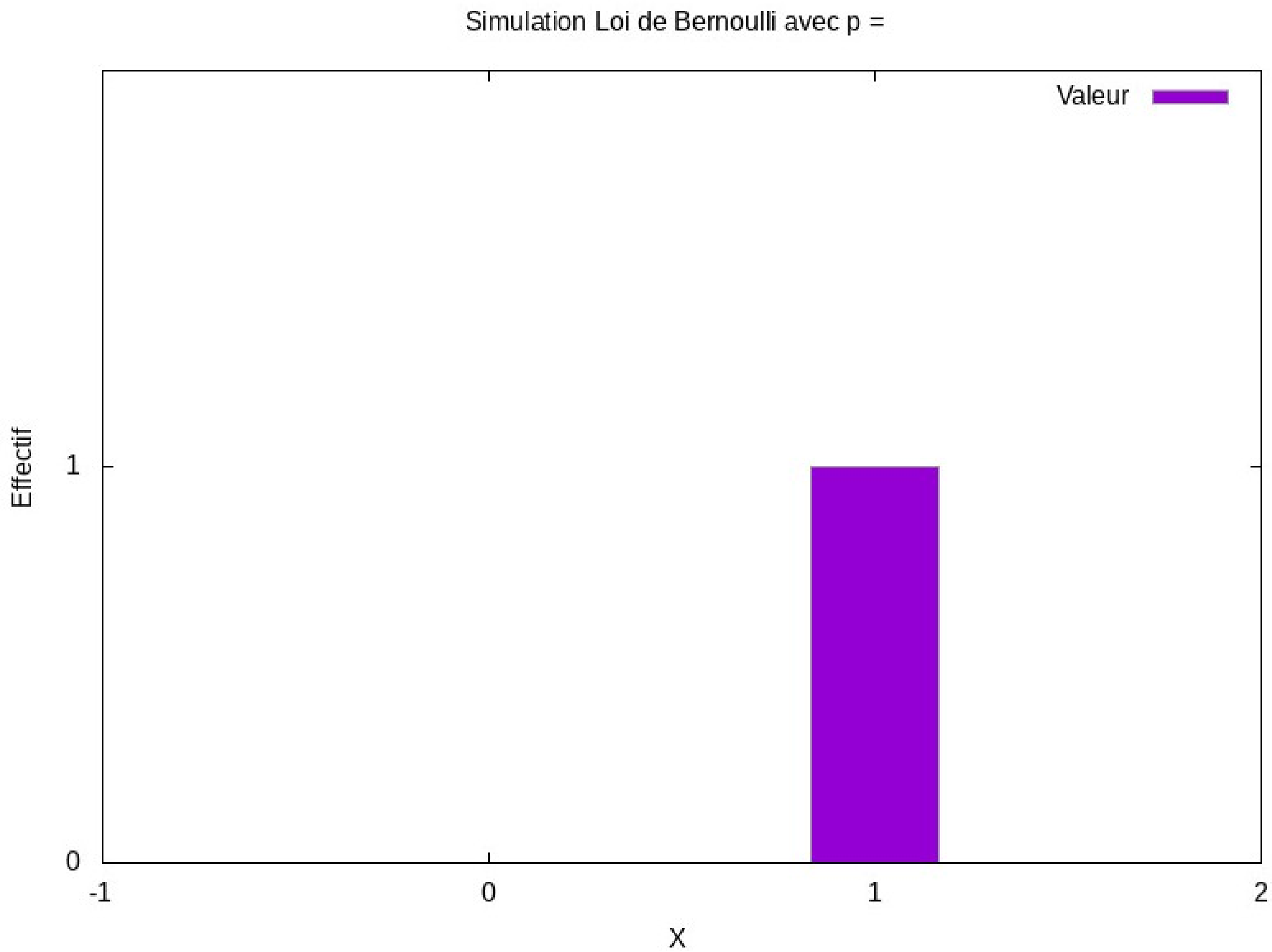
1. Loi de Bernoulli de paramètre p

Approche théorique
On sait que $E_{th}(X) = p$
et $Var_{th}(X)=p(1-p)$

On choisit de prendre p=0.7
on a donc $E_{th}(X)=0.7$ et $Var_{th}(X)=0.21$

Approche expérimentale

A partir du code, on effectue un tirage aléatoire avec une probabilité p=0.7.
On obtient le graphique suivant :



RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

Etant donné que p=0.7, on a de fortes chandes de tomber sur un succès (c’est-à-dire X=1). Effectivement, on obtient X=1 avec la simulation.

Dans le code, on répète plusieurs fois Bernoulli. On choisit de répéter l’expérience 1000 fois. On récupère la variance et l’espérance directement dans le code :

```
void exo2EspVar(float p){
    double esperance =0;
    double variance =0;
    float *resultBer=malloc(1000 * sizeof(int));
    float u;
    for (int i=0 ; i<1000 ; i++){
        resultBer[i]=I_Bernoulli(p);
        esperance += resultBer[i];
    }
    esperance/=1000;
    for (int i=0 ; i<1000 ; i++){
        variance += pow((resultBer[i]-esperance),2);

    }
    variance/=(1000-1);

    printf("esperance : %lf\nvariance : %lf\n", esperance, variance);
}
```

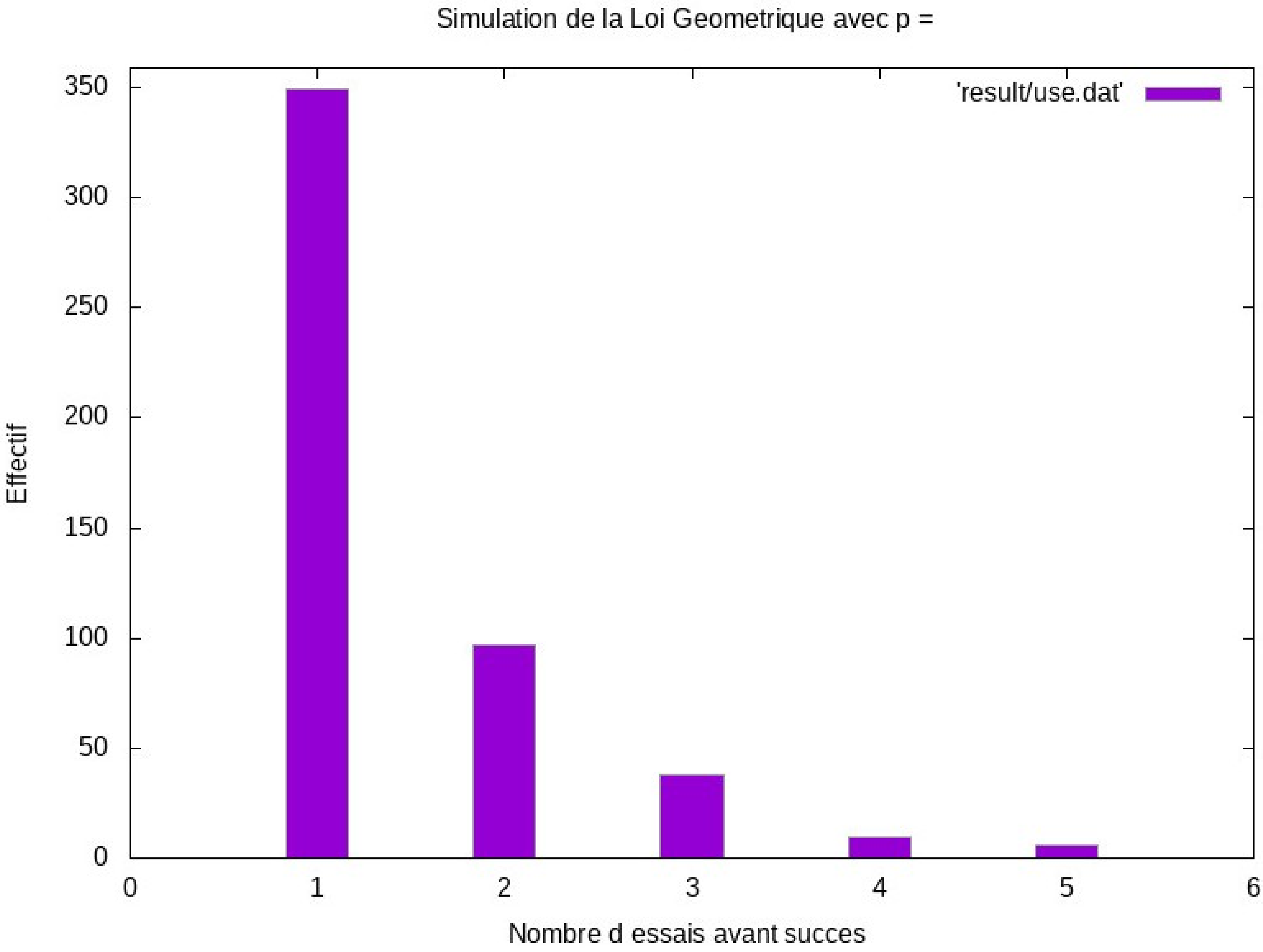
On obtient $E_{exp}(X) = 0.712000$ soit une espérance environ égale à la valeur théorique.
$Var_{exp}(X)= 0.205261$ soit environ la valeur théorique.

2. Loi géométrique de paramètre p

On prendra p=0.7 et on répètera l’expérience 500 fois.

Approche théorique
$E_{th}(X) = 1/p = 1.4$
$V_{th}(X) = (1-p)/p^2 = 0.6$

Approche expérimentale
On choisit de tracer les moyennes de effectifs de la loi géométrique pour 500 simulations.



X=	1	2	3	4	5
effectif	349	97	38	10	6

$E_{exp}(X) = 1.454$
$Var_{exp}(X) = 0.665$

On obtient des valeurs proches des valeurs théoriques.

3. Loi de poisson de paramètre λ.

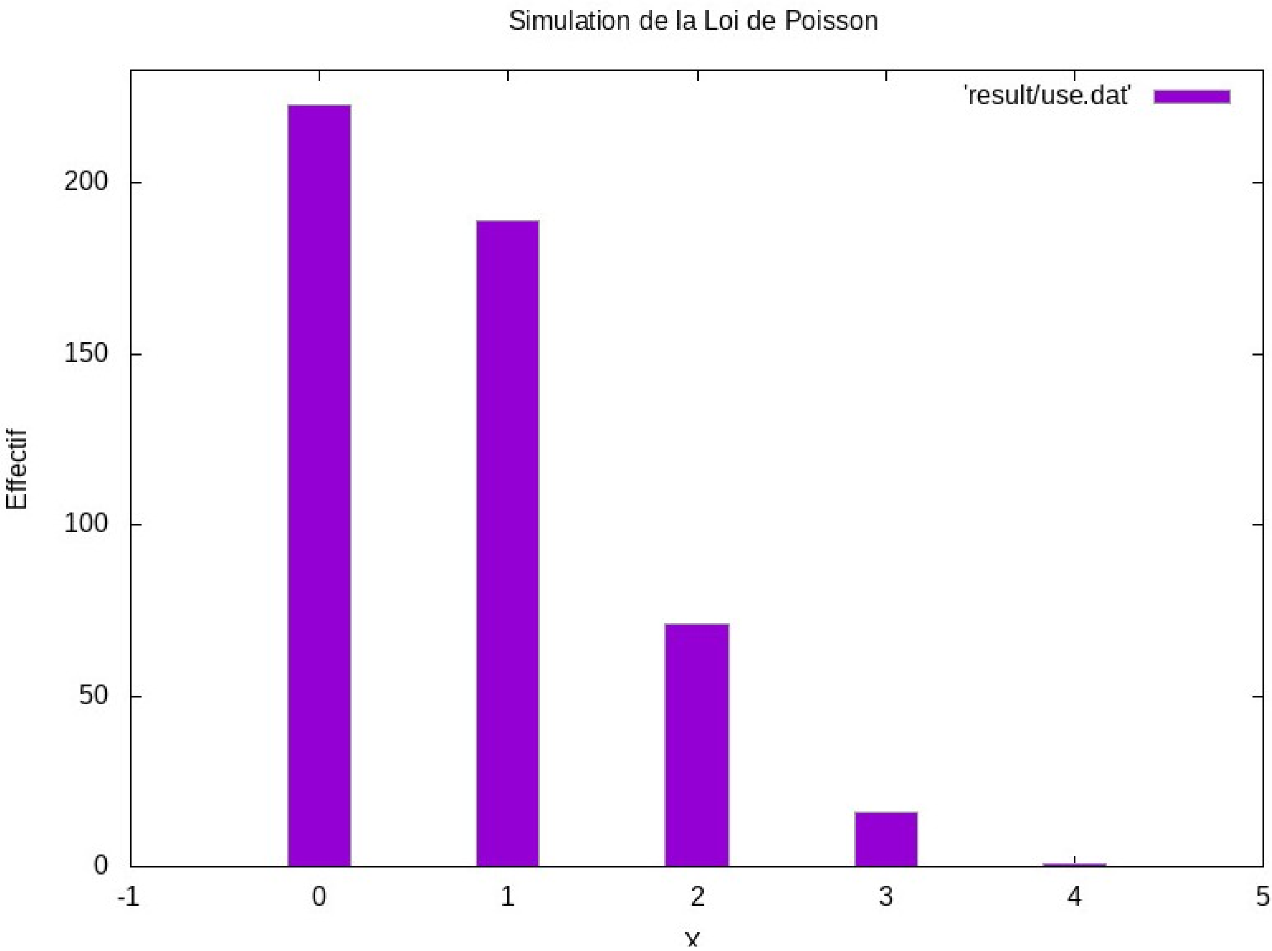
On prendra λ=0.8 et on répètera l’expérience 500 fois.

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

<u>Analyse théorique</u>

$E_{th}(X) = \lambda = 0.8$
$Var_{th}(X) = 0.8$

<u>Analyse expérimentale</u>



X	0	1	2	3	4
effectif	223	189	71	16	1

$E_{exp}(X) = 0.766$
$Var_{exp}(X) = 0.680$

4. Loi exponentielle de paramètre λ.

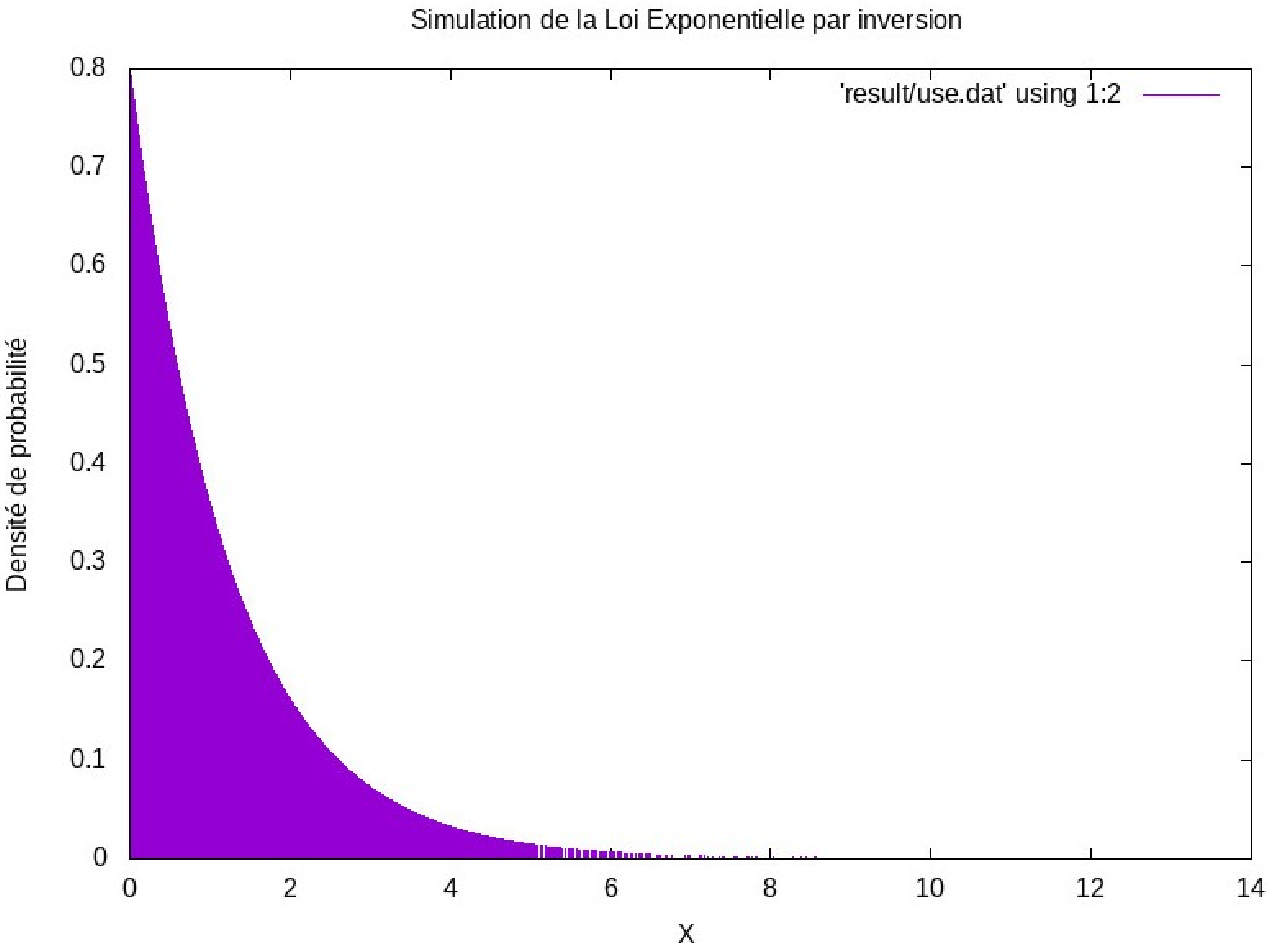
On prendra λ = 0.8 et on répètera l'expérience 10 000 fois.

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

Analyse théorique

$E_{th}(X) = 1/\lambda = 1.25$
$Var_{exp}(X) = 1/\lambda^2 = 1.5625$
densité de probabilité : $f(x) = \lambda e^{-\lambda x}\, 1_{[0, +inf[}$

Analyse expérimentale



Le graphique représente la densité de probabilité de la loi exponentielle tracée de façon expérimentale. La forme de la courbe correspond bien à l’expression théorique de la densité de probabilité de la loi exponentielle.

On calcule l’espérance et la variance directement dans le code (dans accueil.c) :

```c
double esperance=0;
double variance=0;
fprintf(inputFile, "%d %d %f\n",chosenExercise,chosenQuestion, lambda);
float *x=malloc(10000*sizeof(float));
for(int i=0; i<10000; i++){
    x[i] = Exponential(lambda);
    float y = law_Exponential(lambda, x[i]);
    esperance+= x[i];
    fprintf(inputFile,"%f %f\n", x[i], y);
}
esperance/=10000;
for(int i=0; i<10000; i++){
    variance += pow((x[i]-esperance),2);
}
variance/=9999;
printf("esperance : %lf\nvariance : %lf\n", esperance, variance);
```

On obtient des valeurs proches des valeurs théoriques :
esperance : 1.244843
variance : 1.540563

Exercice 3

1. On prendra N=1000

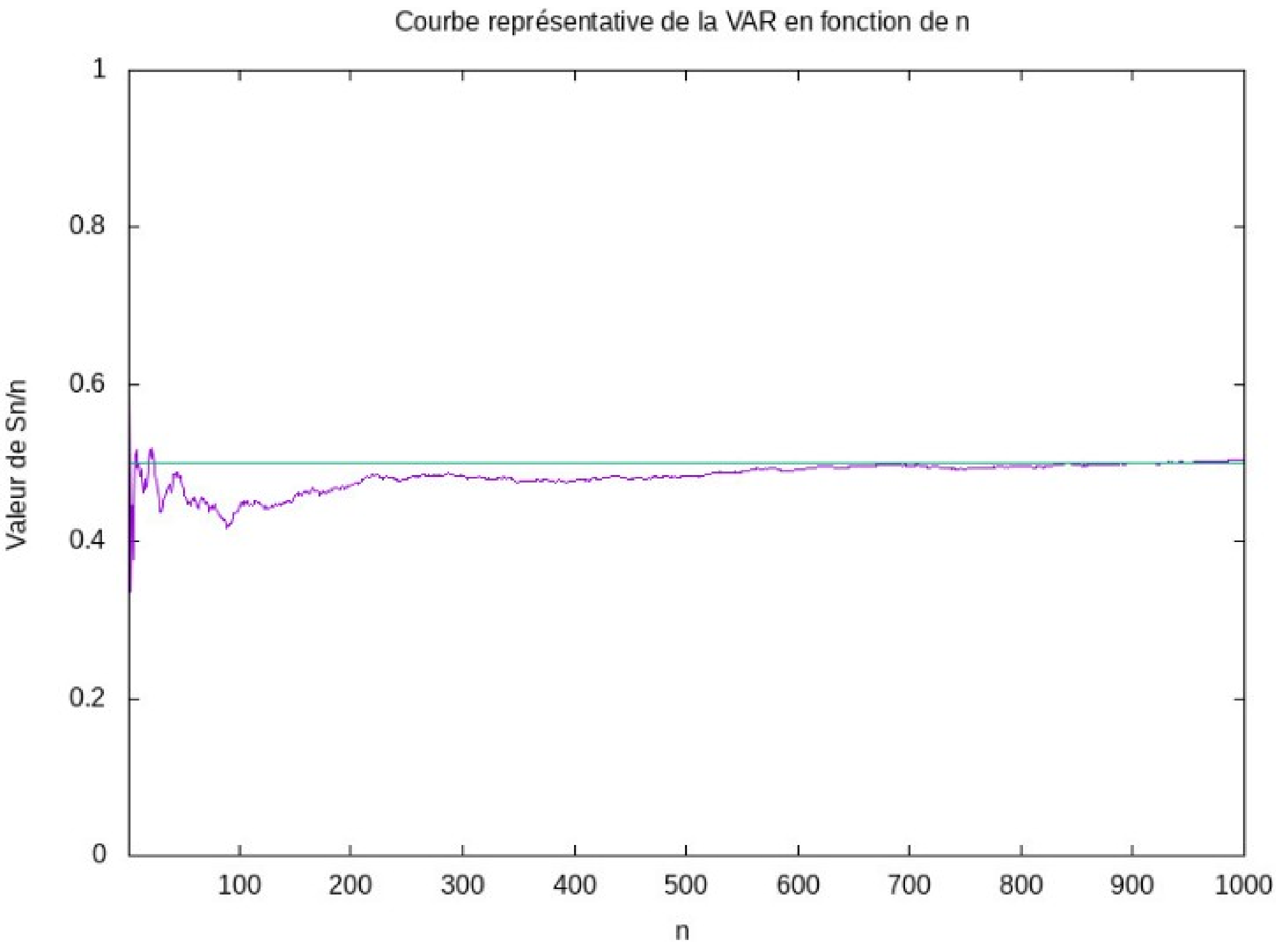
Sn = (1/n) * $\sum_{i=1}^{n} Xi$ avec les Xi qui suivent la loi uniforme sur [0,1]

Pour ce programme on représente Sn/n en fonction de n.
Pour n=1, on aura donc $S_1 = X_1$
$S_2/2 = (1/2) * (X_1 + X_2)$
$S_3/3 = (1/3) * (X_1 + X_2 + X_3)$ et ainsi de suite.

Comme les X_i suivent la loi uniforme, on a m=1/2

On obtient le graphique suivant :

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

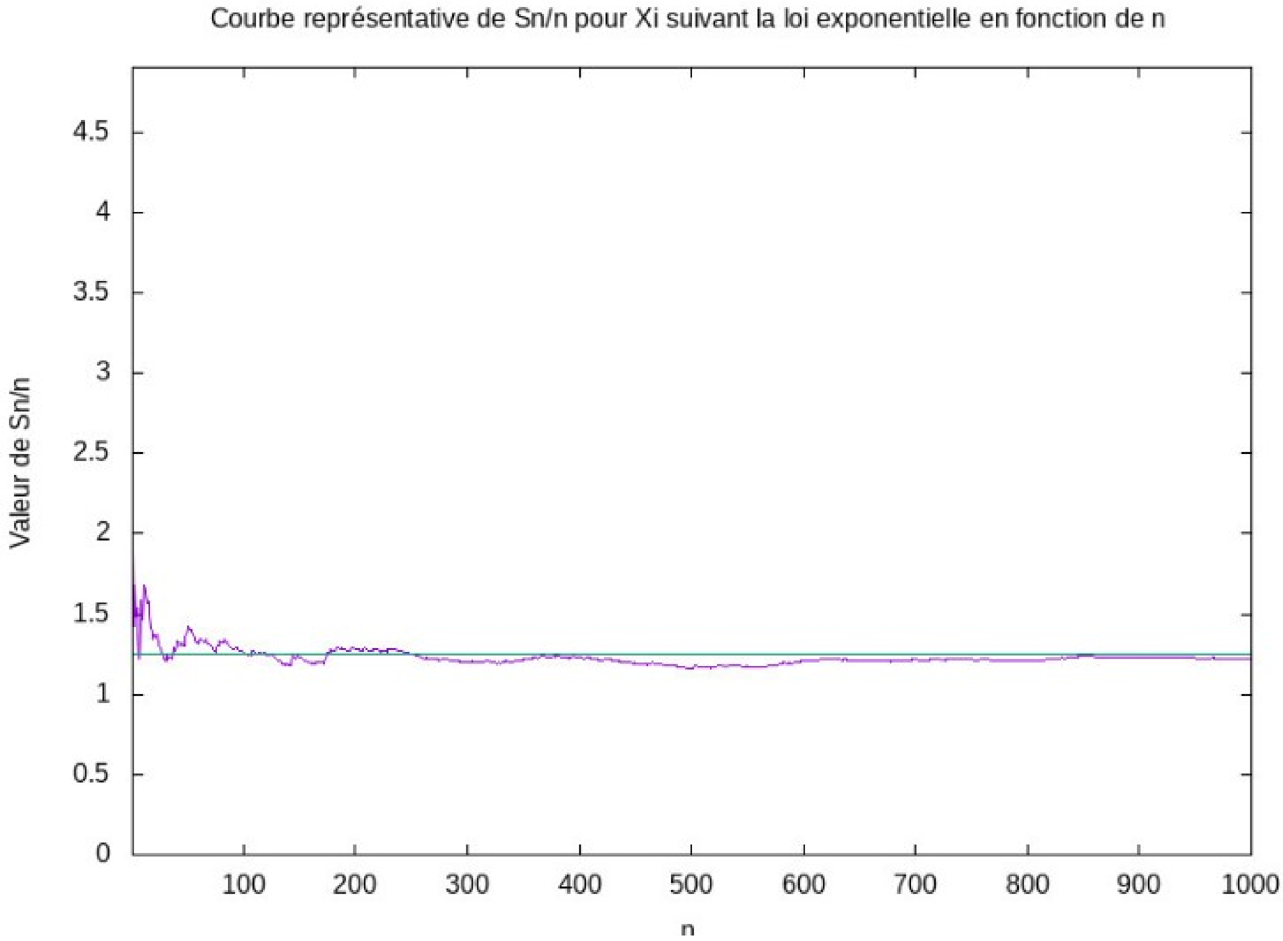


On observe que la moyenne des résultats pour un N grand tend vers l'espérance de la loi uniforme. On reconnaît alors la loi des grands nombres puisque la variable aléatoire Sn/n qui correspond à $\bar{X}$ converge presque sûrement vers $E(X_1)$.

2. On prendra λ = 0.8

Par approche théorique, comme les X_i suivent la loi exponentielle, pour tout i appartenant aux entiers de 1 à N, $E(X_i) = 1/0.8 = 1.25$
On trace la droite y = 1.25 et la courbe représentative de Sn/n :

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess



De nouveau, on observe que $\bar{X}$ converge presque sûrement vers $E(X_1)$ et donc on observe bien que la loi des grands nombres fonctionne.

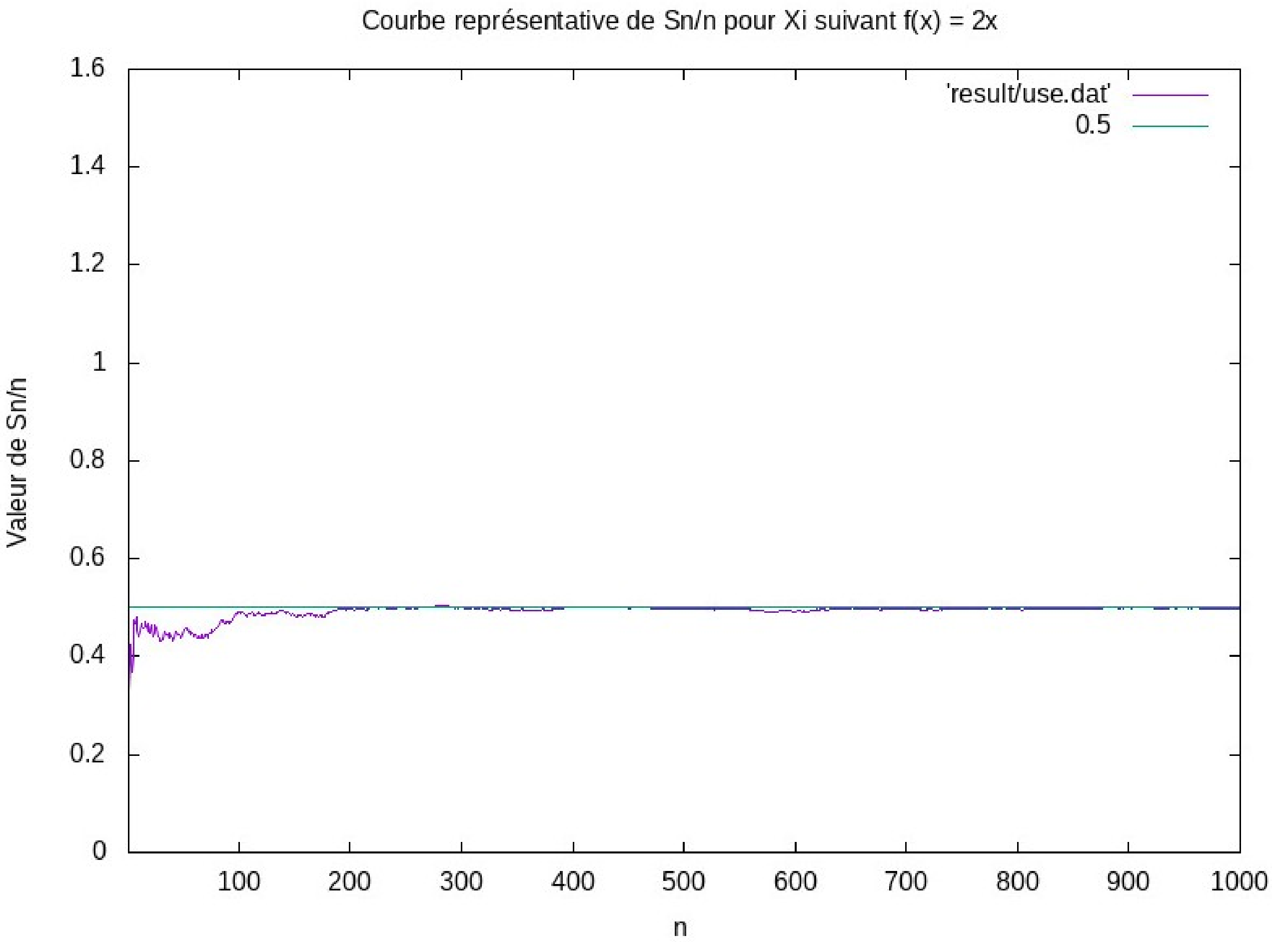
3. On a la fonction densité $f(x) = 2x * 1_{[0,1]}(x)$

Par la méthode d'inversion, on calcule F :

Pour x appartenant à [0,1] :
si on note $F(x) = u$
on a $f(x)=2*x$ donc $F(x) = x^2$
d'ou $x^2 = u$
et donc $x= \sqrt{u}$ car x positif

D'où la fonction d'inversion :

```
double F_inv(){
    double u = rand_0_1();
    double x = sqrt(u);
    return x;
}
```



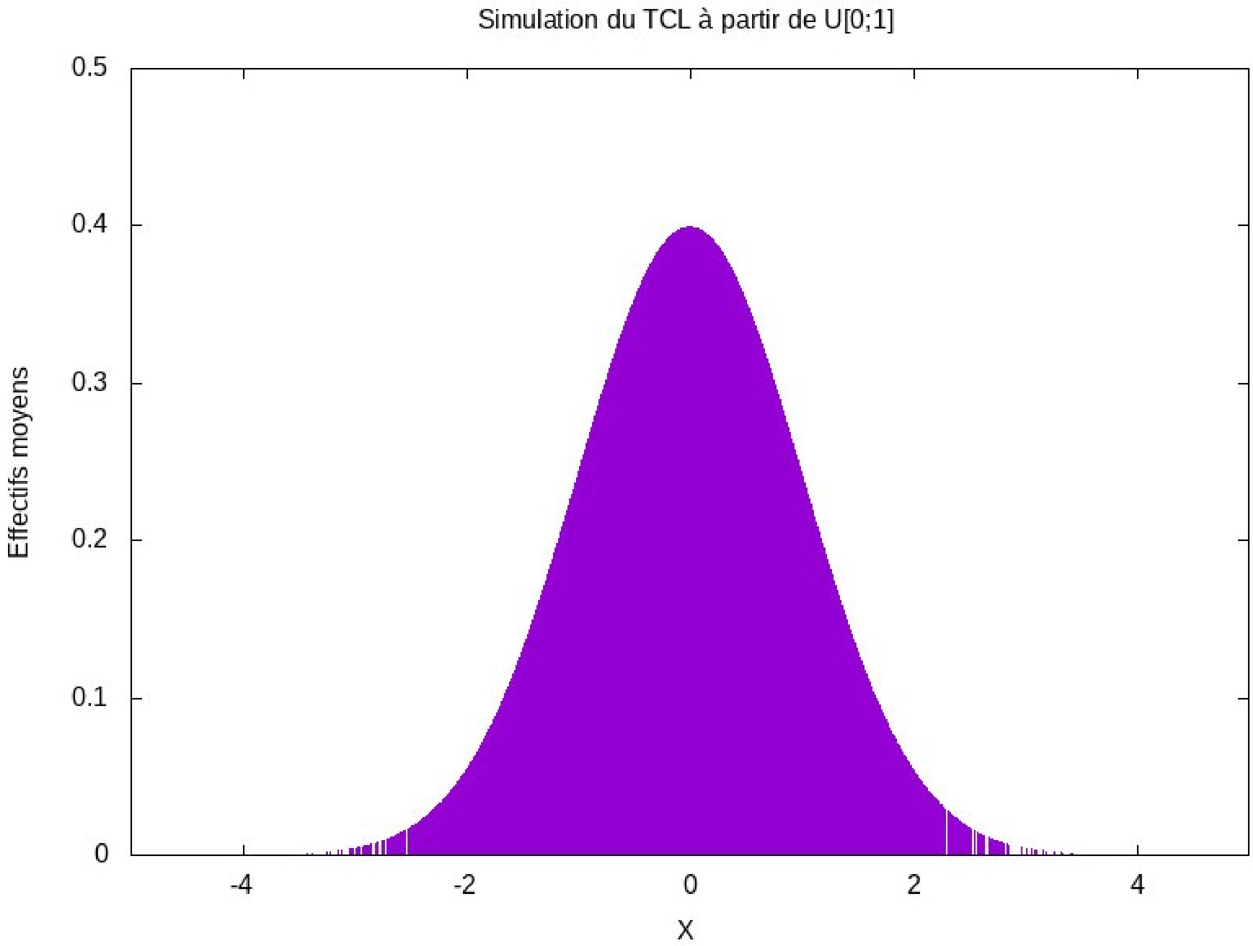
De nouveau, on observe que $\bar{X}$ converge presque sûrement vers $E(X_1)$ et donc on observe bien que la loi des grands nombres fonctionne.

Exercice 4

1.a On prendra n=12 simulations pour avoir une bonne représentation graphique de la convergence.

On obtient le graphique suivant :

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess



On observe que la simulation et le tracé de Zn donne une courbe représentative de la loi normale. On peut donc penser que, pour n assez grand, on peut approcher Zn par une loi normale (il s'agit du TCL).
Nous allons observer le TCL à la question 1.b

1.b On veut montrer la convergence en loi de Zn vers la loi gaussienne. On veut donc montrer la convergence simple de F_{Zn} vers F (fonction de répartition de la loi gaussienne).
On trace d'abord l'histogramme de F_{Zn} . En abscisse on mettra la valeur de x et en ordonnée le nombre de valeurs moyen pour lesquelles $X<x$.

Pour X_i suivant la loi uniforme, l'espérance est de 1/2 et la variance est de 1/12 d'où m=1/2 et σ = 1/sqrt(12) . On choisi de prendre n=12 c'est-à-dire que l'on étudie Z_{12}.

On considérera x appartenant à [-5, 5] car Z_{12} ne peut prendre que des valeurs entre -5 et 5.

On utilisera les résultats de 10 000 simulations de Z_{12}.

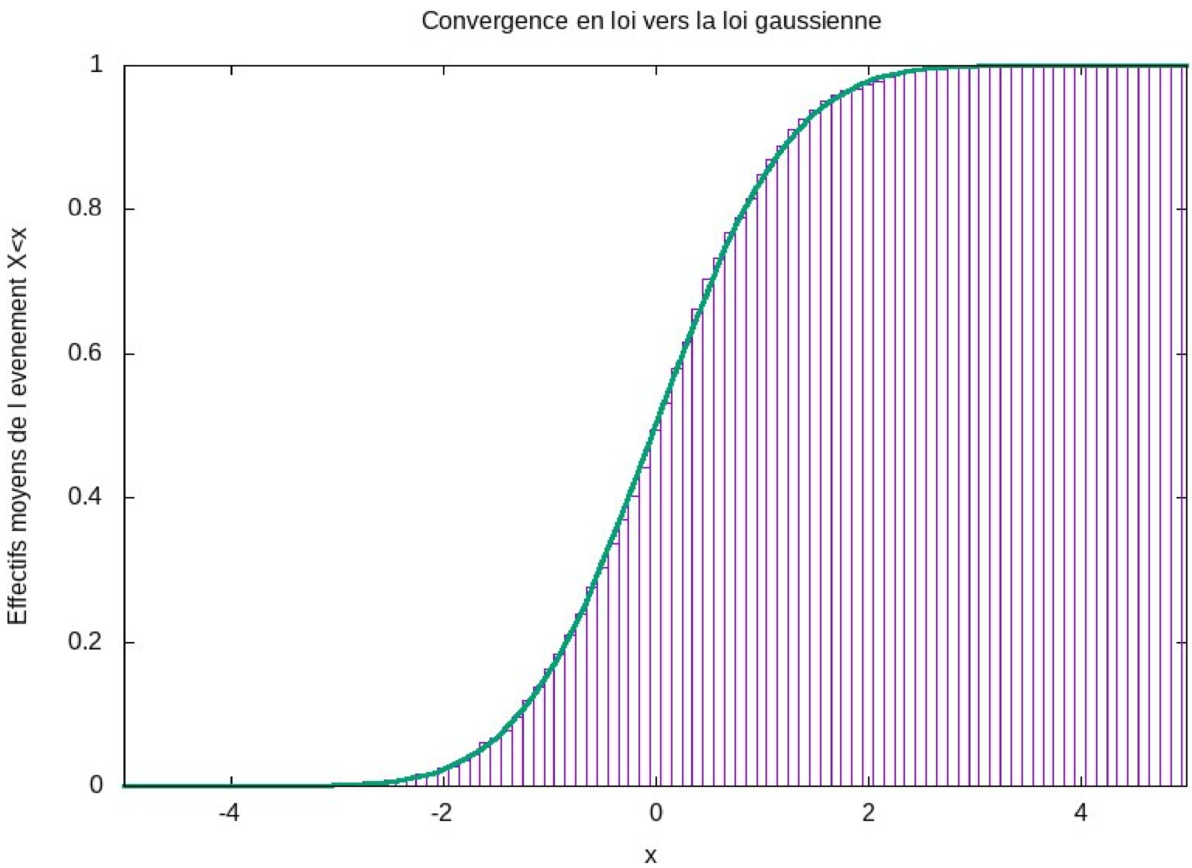
RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

Pour une simulation, on commence par simuler 12 fois la loi uniforme (ce qui donne les valeurs de Xi). On met les resultats dans la variable Sn (correspondant à S_{12}) puis on calcule Zn.
On répète la même chose 10 000 fois et on remplit interval[i].cpt :
inter[i].interval contient la valeur de x dans l'intervalle [-5,x] et inter[i].cpt contient le nombre de Zn dans [-5,x]

De cette façon on calcule les effectifs moyens et on trace le graphique.
On trace ensuite la fonction de répartition de la loi gaussienne.
Il n'est pas possible de l'exprimer à partir de fonctions usuelles mais on peut l'exprimer en fonction de la fonction d'erreur :

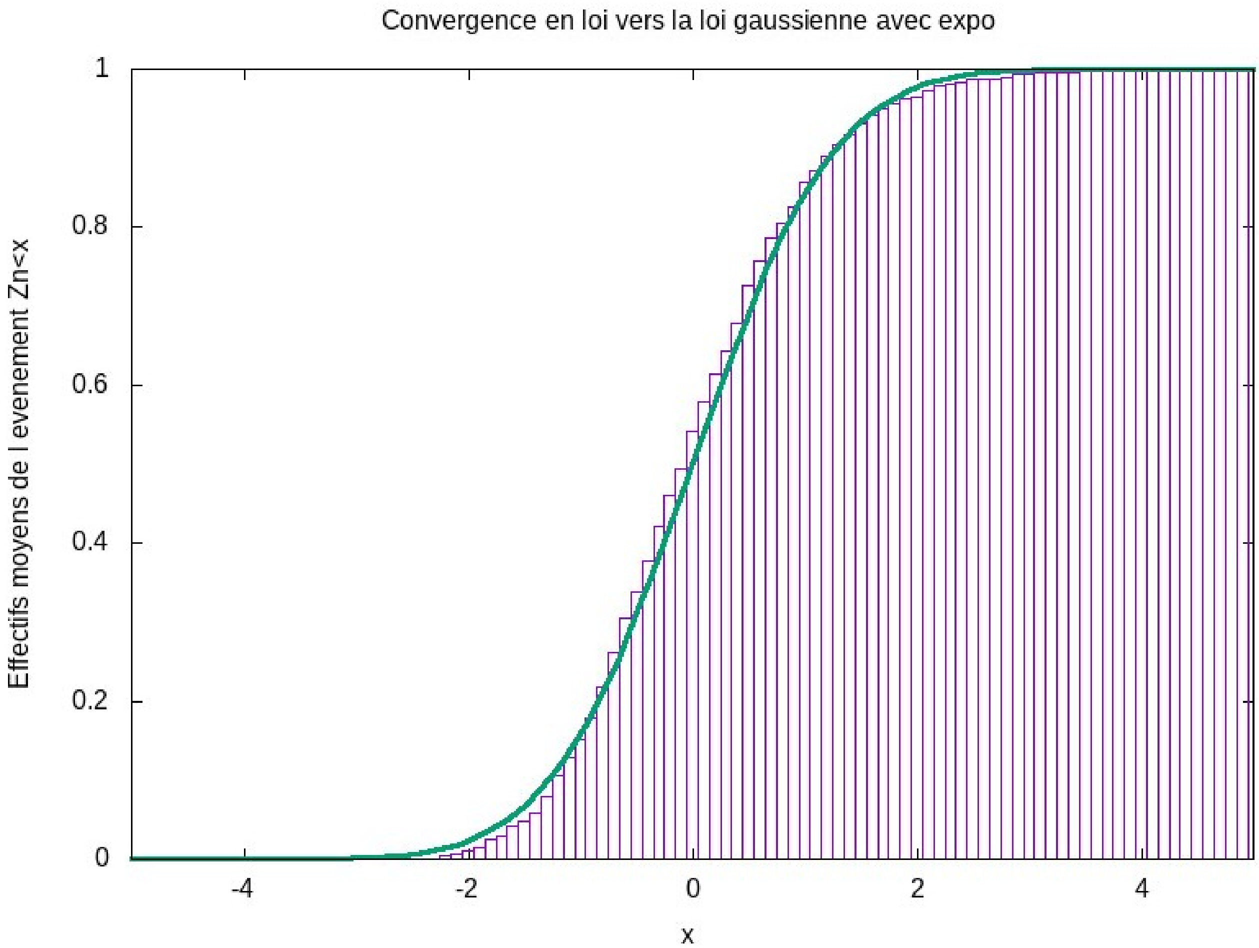
$$F(x) = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)$$

On finit par ajouter le tracé de F(x) sur [-5,5] sur notre graphique.



On peut ainsi observer la convergence en loi de F_{Zn} vers F.

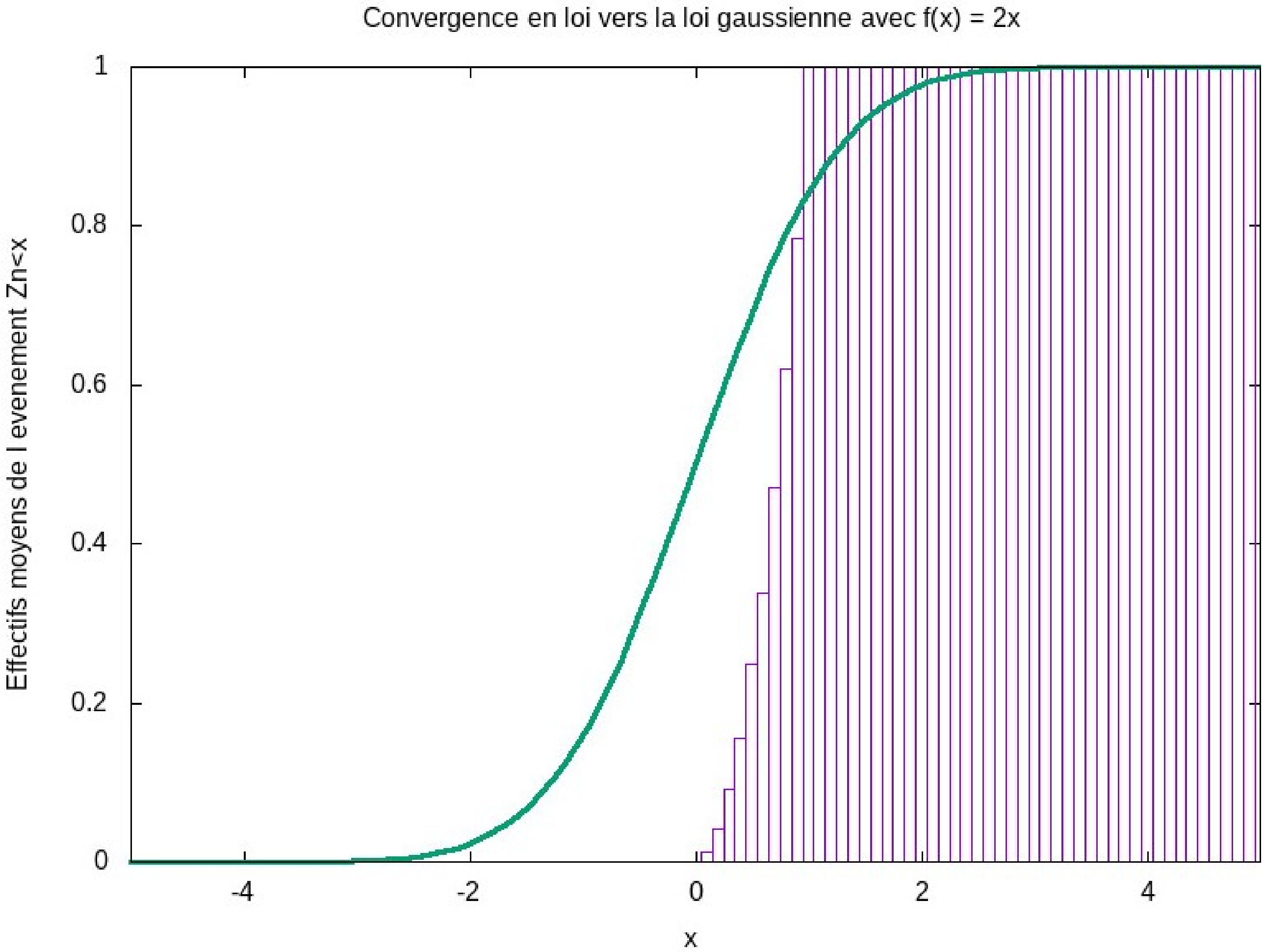
2. On reproduit le même principe pour des variables suivant la loi exponentielle. On prendra λ=2 et le même nombre de simulations que précédemment.



On observe de nouveau la convergence en loi de F_{Zn} vers F.

3. On utilisera la même fonction d’inversion F_inv que celle utilisée dans l’exercice 3.

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess



Exercice 5

RAHMANI Kevin LIANG Yixin CZAPLINSKI Tess

La méthode de Box-Muller consiste à générer deux variables aléatoires X et Y indépendantes et identiquements distribuées sur [0;1] de la forme :

$$X = \sqrt{-2\ln(U)}\cos(2\pi V) \quad et \quad Y = \sqrt{-2\ln(U)}\sin(2\pi V)$$

avec X, Y suivant une loi normale centrée réduite. En résumé, cet algorithme est utilisé pour générer des nombres aléatoires suivant une distribution normale à partir de nombre aléatoires uniformes en utilisant la transformation de Box-Muller. Les avantages par rapport aux méthodes de l'exercice précèdent sont que Box-Muller offre une plus grande simplicité de mise en œuvre, est beaucoup plus rapide à l'éxécution, offre un couple de variables aléatoires en deux dimensions (x,y), et que pour certaines distributions il y a une plus grande précision et stabilité des données.
On commence donc à générer deux nombres aléatoires suivant une loi uniforme sur [0;1] appelée U et V et on applique tout simplement la formule de la transformation de Box-Muller ci-dessus qu'on répète pour 10000 simulations.
Voici le graphique obtenu représentant la densité: