TP 10: Estimation ponctuelle.

28 février 2025

Introduction :Exercice 1

On lance 10 fois de suite une pièce honnête. On cherche à estimer la probabilité de l'événement A : obtenir au moins une fois une succession de 3 résultats identiques.

1. On rappelle que si a et b sont deux entiers naturels, l'instruction `rd.randint(a,b)` renvoie un entier aléatoire de $[\![a, b-1]\!]$. Compléter la fonction `simul()` suivante afin qu'elle simule cette expérience et renvoie la valeur 1 lorsque cet événement est réalisé et la valeur 0 lorsqu'il ne l'est pas.

```python
import numpy.random as rd
def simul():
    U = rd.randint(1,3,10)
    x = 0
    for i in range(2,10):
        if .................... :
            x = 1
    return x
```

2. On renouvelle n fois cette expérience. On note X_k la variable de Bernoulli indicatrice de la réalisation de l'événement A lors de la k-ème tentative. On pose $\bar{X}_n = \dfrac{1}{n} \sum_{k=1}^{n} X_k$

   Selon la loi faible des grands nombres $\lim\limits_{n \to +\infty} \mathbb{P}(|\bar{X}_n - \mathbb{P}(A)| \geqslant \varepsilon) = 0$. On peut donc **estimer** que la valeur de $\mathbb{P}(A)$ est environ égale à $\bar{X}_n$ pour n assez grand.
   Ecrire un programme qui permet de calculer une valeur approchée de $\mathbb{P}(A)$.

3. Le programme suivant permet de visualiser la manière dont $\bar{X}_n$ se rapproche de $\mathbb{P}(A)$ lorsque n augmente. Expliquer ce programme et le résultat affiché.

```python
import matplotlib.pyplot as plt
import numpy as np
plt.close()
N = 1000
essais = [50,100,1000]
for k in range(3):
    M = np.zeros(N)
    n = essais[k]
    for i in range(N):
        X = np.zeros(n)
```

```
            for j in range(n):
                X[j] = simul()
            M[i] = np.mean(X)
        plt.subplot(1,3,k+1)
        plt.hist(M, range=(0,1), bins=50)
plt.show()
```

Exercice 2 :Estimation du paramètre d'une loi de Poisson.

On considère n variables aléatoires $X_1, \cdots, X_n$ indépendantes et identiquement distribuées de même loi de Poisson de paramètre λ inconnu.

1. Montrer que $Y_n = \dfrac{1}{n}\sum_{k=1}^{n} X_k$ est un estimateur de λ. Calculer son espérance.
2. Posons $V_n = \dfrac{1}{n}\sum_{i=1}^{n}(X_i - Y_n)^2$, la variance empirique des X_i.
   (a) Montrer que pour tout entier n non nul $\mathbb{E}(V_n) = \dfrac{n-1}{n}\lambda$.
   (b) En déduire un estimateur Z_n de λ tel que $\mathbb{E}(Z_n) = \lambda$. Z_n est appelée variance empirique débiaisée des X_i.
3. Ecrire un programme, utilisant `rd.poisson()` qui simule N fois une loi de Poisson de paramètre $\lambda = 3$ et qui affiche l'estimation que l'on peut faire de λ selon les deux méthodes. Comparer l'erreur commise avec différentes valeurs de λ et de N. Quelle est le meilleur estimateur ?
4. En Python la variance empirique se calcule avec la fonction `np.var()`. Compléter le programme suivant afin qu'il simule $N = 2500$ fois le programme de la question 3 avec $n = 100$ et qu'il affiche les histogrammes empiriques des variables aléatoires Y_n et Z_n.

```
N = 2500
n = 100
X = rd.poisson(l, size=[n,N])
Y = ...
Z = ...
m = min(min(Y),min(Z))
M = max(max(Y),max(Z))
plt.close()
plt.subplot(1,2,1)
plt.hist(Y, range=(m,M), bins=50, color="yellow", edgecolor="red")
plt.subplot(1,2,2)
plt.hist(Z, range=(m,M), bins=50, color="yellow", edgecolor="red")
plt.show()
```

Exercice 3 : Estimation du paramètre d'une loi uniforme à densité.

On veut estimer le paramètre a d'une variable aléatoire X suivant une loi uniforme sur $[0, a]$. On considère n variables aléatoires $X_1, \cdots, X_n$ indépendantes de même loi que X. On pose

$$\bar{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k, M_n = \max(X_1, \cdots, X_n) \text{ et } V_n = 2X_n.$$

1. Calculer $\mathbb{E}(\bar{X_n})$. En déduire un premier estimateur Y_n d'espérance a.
2. (a) Montrer que la fonction de répartition de M_n est donnée par :

   $$\forall x \in \mathbb{R}, \quad F_{M_n}(x) = [F(x)]^n$$

   où F est la fonction de répartition de X.

   (b) En déduire que M_n est une variable aléatoire à densité et qu'une densité de M_n est donnée par :

   $$\forall x \in [0, a], f_{M_n}(x) = \frac{n}{a^n}x^{n-1} \text{ et sinon } f_{M_n}(x) = 0$$

   (c) Calculer $\mathbb{E}(M_n)$. En déduire un second estimateur Z_n d'espérance a.
3. Montrer que Y_n est un estimateur convergent de a.
4. Calculer l'espérance et la variance de Z_n et de V_n. Ces estimateurs sont-ils sans biais ? Sont -ils convergent ?
5. Quel estimateur de a privilégierez-vous ?
6. Ecrire un script en python qui simule 1000 fois une loi uniforme sur $[0, a]$ et qui affiche les trois estimations. Tester sur différentes valeurs de a. Faire un commentaire.
7. Reprendre les instructions de l'exercice précédent et les modifier afin d'afficher les histogrammes empiriques des variables Y_n, Z_n et V_n pour $a = 3$ avec $n = 100$ en répétant l'expérience $N = 2500$ fois.

Exercice 4 : Variance d'une loi normale - Edhec 2020 E

On considère une variable aléatoire X suivant une loi normale $\mathcal{N}(0, \sigma^2)$ où $\sigma > 0$ est inconnu. On pose $Y = |X|$. On admet que Y est une variable aléatoire à densité.

1. (a) Montrer que la fonction de répartition de Y est donnée par :

   $$F_Y(x) = 2F_X(x) - 1 \text{ si } x \geq 0 \text{ et } F_Y(x) = 0 \text{ sinon}$$

   (b) En déduire une densité de Y.

   (c) Montrer que Y admet une espérance et que $\mathbb{E}(Y) = \sigma\sqrt{\frac{2}{\pi}}$.
2. On considère un échantillon $(Y_1, \cdots, Y_n)$ composé de variables aléatoires indépendantes suivant la même loi que X. On pose $S_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$.

   (a) Montrer que S_n est un estimateur de σ. Donner son espérance. Proposer un estimateur T_n d'espérance σ.

(b) Calculer $\mathbb{E}(X^2)$, $\mathbb{E}(Y^2)$, $\mathbb{V}(Y)$ et $\mathbb{V}(S_n)$.

(c) En déduire la variance de T_n.

3. Ecrire une fonction en python qui prend en entrée un entier naturel n et un réel σ positif et qui renvoie une simulation de S_n et une simulation de T_n.