

DM 5

THÈME : DEUX CLASSIQUES EN PROBABILITÉ

Sujet A : faire exercice 1.

Sujet * : exercices 1 et 2.

Exercice 1

Inégalité de concentration : Bienaymé-Tchebychev ou Chernoff?

Pour tout n de $\mathbb{N}^*$, S_n désigne une variable aléatoire définie sur $(\Omega, \mathcal{A}, \mathbf{P})$ de loi binomiale $\mathcal{B}(n, p)$, avec $0 < p < 1$, et x un réel strictement positif fixé. On pose $q = 1 - p$.

1. À l'aide de l'inégalité de Bienaymé-Tchebychev, majorer la quantité $\mathbf{P}(|S_n - np| \geq x\sqrt{n})$ indépendamment de n et p .

L'objectif de la suite est d'établir une deuxième majoration.

2. Montrer que pour tout $\lambda \in \mathbb{R}_*^+$

$$\mathbf{P}(S_n - np \geq nx) \leq \frac{\mathbf{E}(e^{\lambda(S_n - np)})}{e^{n\lambda x}}.$$

3. Établir la formule $\mathbf{E}(e^{\lambda(S_n - np)}) = e^{-n\lambda p} (p \cdot e^\lambda + q)^n$.

4. Pour tout $\lambda \geq 0$, on pose : $\psi(\lambda) = \ln(p \cdot e^\lambda + q) - \lambda p$.

- a) Montrer que : $\sup_{\lambda \geq 0} \psi''(\lambda) \leq \frac{1}{4}$.

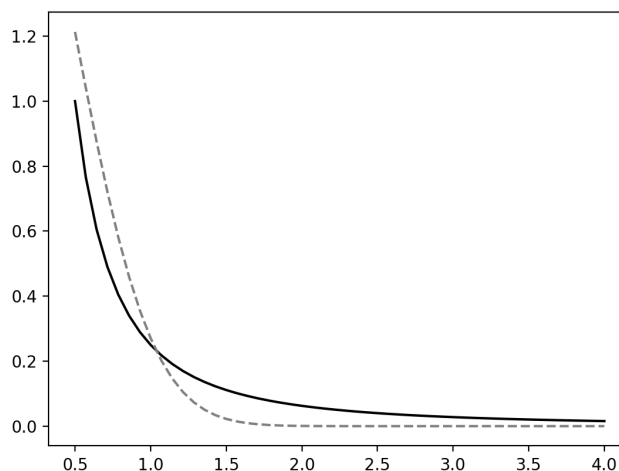
- b) En déduire que : $\sup_{\lambda \geq 0} \psi(\lambda) \leq \frac{\lambda^2}{8}$.

5. Montrer les inégalités :

$$\mathbf{P}(S_n - np \geq nx) \leq e^{-2nx^2} \quad \text{et} \quad \mathbf{P}(|S_n - np| \geq x\sqrt{n}) \leq 2e^{-2x^2}.$$

6. Commenter le résultat suivant.

```
1 import numpy as np
2 import matplotlib.pyplot as plt
3
4 x=np.linspace(0.5,4,50)
5 y=1/(4*x**2)
6 z=2*np.exp(-2*x**2)
7 plt.plot(x,y,color='black')
8 plt.plot(x,z,'--',color='gray')
9 plt.show()
```



Exercice 2

Problème du collectionneur

On suppose qu'un enfant collectionne des images de joueurs de football qu'il trouve en cadeau dans des tablettes de chocolat. Une tablette contient une seule image et il y a $N \geq 2$ images de joueurs différentes numérotées de 1 à N , que l'entreprise de chocolat a réparties avec la même fréquence $1/N$. Notons X_n le numéro de l'image trouvée dans la n -ième tablette ouverte par l'enfant. Alors, les hypothèses sur la répartition des images impliquent que $(X_n)_{n \in \mathbb{N}^*}$ est une suite de variables uniformes sur $\llbracket 1, N \rrbracket$.

Pour tout $k \in \llbracket 1, N \rrbracket$, on note

$$\tau_k^N = \inf \left\{ n \in \mathbb{N}^* \mid \text{Card}(\{X_1, \dots, X_n\}) = k \right\}.$$

Pour tout $N \in \mathbb{N}^*$, on définit aussi

$$H_N = \sum_{k=1}^{N-1} \frac{1}{k} \quad \text{et} \quad S_N = \sum_{k=1}^{N-1} \frac{1}{k^2}.$$

1. Simulation

- a) Écrire une fonction python `testcoeff()` qui prend en argument une matrice ligne et compte le nombre de termes non nuls de la matrice.
- b) Compléter le programme suivant qui permet de simuler τ_k^N .

```
1 import numpy as rd
2 import numpy.random as rd
3
4 def tnk(N,k):
5     C=0
6     A=np.zeros(N)
7     while ...
8         i=rd.randint(0,N)
9         # tire au hasard un entier entre 0 et N-1
10        A[i]+=1
11        C= ...
12    return C
```

- c) Que renvoie `tnk(5,7)` ?

2. a) Calculer $\mathbf{P}(\tau_2^N - \tau_1^N = m)$ pour tout $m \in \mathbb{N}^*$. Plus généralement, calculer pour tout $m \in \mathbb{N}^*$,

$$\mathbf{P}(\tau_{k+1}^N - \tau_k^N = m).$$

- b) Calculer également $\mathbf{E}(\tau_{k+1}^N - \tau_k^N)$.

3. Posons $T_N = \tau_N^N$. Calculer $H_N^{-1} \cdot \mathbf{E}(T_N)$.
4. Exprimer la variance de T_N en fonction de H_N et S_N .
5. Déterminer une constante C tel que pour tout $\varepsilon \in \mathbb{R}_*^+$, tout $N \in \mathbb{N}$,

$$\mathbf{P}(|H_N^{-1} T_N - N| \geq N\varepsilon) \leq \frac{C}{\ln(N)^2 \varepsilon^2}.$$