Feuille 4 - Approche fréquentiste et inégalités de concentration. Chaînes de Markov.

M1 Enseignement Mathématiques - Probabilités et Statistiques

Un des objectifs de cette feuille est de justifier l'approche fréquentiste travaillée dans les feuilles précédentes à l'aide des inégalités de concentration.

Un autre objectif est de travailler le formalisme relatif aux chaînes de Markov introduit en Terminale. Pour réviser ce dernier thème, on pourra également reprendre le travail effectué en *Initiation à la recherche* au premier semestre.

`Exercice 1.` *Retour sur le problème - Convergence de la moyenne vers l'espérance.*

1. On reprend l'algorithme (dans sa version corrigée !) de la première partie du problème étudié, concernant le déplacement de la puce. On considère X une variable aléatoire dont la valeur correspond à la position de la puce après les 50 sauts. Calculer l'espérance et la variance de X.

2. On souhaite répèter l'expérience 1000 fois, et renvoyer la moyenne des positions de la puce. (C'est-à-dire que 1000 fois, la puce va partir de la position 0 et effectuer 50 sauts.) Implémenter l'algorithme : quelle valeur moyenne obtenez-vous ?

3. Démontrer (ou redémontrer !) la loi faible des grands nombre en supposant connue l'inégalité de Bienaymé-Tchebychev.

4. Montrer qu'on peut appliquer la loi faible des grands nombres pour interpréter le résultat obtenu à la question **1.** . On pourra introduire des variables aléatoires $X_1, X_2, \cdots$ correspondant aux expériences successives : on les supposera indépendantes, et suivant chacune la même loi que X.

5. A partir de l'inégalité de Bienaymé-Tchebychev, donner pour tout entier $n > 0$ et tout réel $\epsilon > 0$ une majoration de

$$p\left(\left|\frac{X_1 + \dots X_n}{n} - E(X)\right| \geq \epsilon\right)$$

en fonction de n.

6. En déduire une condition suffisante sur le nombre d'expériences à simuler pour obtenir une estimation de l'espérance à 10^{-1} près avec une probabilité supérieure à $0,99$, autrement dit pour que

$$p\left(\left|\frac{X_1 + \dots X_n}{n} - E(X)\right| \geq 0,1\right) \leq \frac{1}{100}.$$

`Exercice 2.` *Retour sur le problème du Duc de Toscane - Convergence de la fréquence vers la probabilité.*

1. On reprend l'algorithme de l'exercice 1 de la Feuille 2. En répétant l'expérience 1000 fois, donner une valeur approchée de la probabilité d'obtenir 9 et de la probabilité d'obtenir 10. Comparer avec les vraies valeurs. Même chose en répétant l'expérience 10000 fois.

2. On souhaite justifier cette façon de procéder : par exemple on va essayer de justifier pourquoi la fréquence d'obtention du 9 converge vers la probabilité. Dans ce but, on introduit une suite de variables aléatoires de Bernoulli $(X_n)_{n\in\mathbb{N}^*}$ supposées indépendantes, telle pour tout entier n, la variable X_n prenne la valeur 1 si et seulement si on obtient 9 au n-ième tirage. Appliquer la loi des grands nombres pour justifier la méthode utilisée en **1** .

3. En utilisant une inégalité déduite de l'inégalité de Bienaymé-Tchebychev, donner une condition suffisante sur le nombre d'expériences à simuler pour obtenir une estimation de la probabilité d'obtenir 9 à 10^{-2} près avec une probabilité supérieure à $0,95$.

`Exercice 3.` *Chaînes de Markov.*
1. Donner les principaux points des programmes relatifs aux chaînes de Markov. Donner pour chacun un énoncé de cours possible en vous inspirant des manuels.
2. Démontrer que la matrice de transition d'un processus de Markov est une matrice stochastique.
3. Prouver que les matrices stochastiques ont une valeur propre commune.
4. Donner un théorème mettant en jeu la matrice de transition et le passage de l'état n à l'état $n+1$. Le démontrer.
5. Donner un énoncé possible relatif aux *probabilités invariantes*, et le démontrer.
6. Donner (sans démonstration !) un énoncé du théorème de Perron-Frobenius (on peut chercher dans la littérature, par exemple dans l'article `https://www.apmep.fr/IMG/pdf/07-Bonneval_C.pdf` à l'APMEP.

`Exercice 4.` *Inégalités de concentration.*
1. Quelles inégalités de concentrations apparaissent au programme de Terminale, et comment s'articulent-elles ?
2. Proposer pour chacune un exemple d'exercice et sa correction.

`Exercice 5.` *Méthode de Monte-Carlo dans un cas particulier.* On va justifier ici le fait suivant, utilisé sans démonstration dans certaines formulation de la méthode de Monte-Carlo pour le calcul de π.

> **Théorème :** Soit X et Y deux variables aléatoires indépendantes et suivant la loi uniforme sur $[0,1]$, et soit D le quart de disque unité défini par
>
> $$D = \left\{(x,y)\ :\ x \geq 0,\ y \geq 0,\ x^2+y^2 \leq 1\right\},$$
>
> alors on a $\quad P\left((X,Y) \in D\right) = \dfrac{\pi}{4}.$

1. Soit $a, b, c \in [0,1]$ avec $a < b$, démontrer que

$$P\left((X,Y) \in [a,b] \times [0,c]\right) = c(b-a).$$

2.a. Soit f la fonction définie sur $[0,1]$ par $f(x) = \sqrt{1-x^2}$. Justifier sans calculer d'intégrale les égalités suivantes :

$$\int_0^1 f(x)dx = \frac{\pi}{4}.$$

2.b. Montrer qu'un changement de variable peut également permettre d'obtenir l'égalité de la question précédente. (Cela peut être utile dans un contexte où on a défini l'intégrale à partir de sommes de Riemann.)
3. Démontrer que f est décroissante sur $[0,1]$.
4. On fixe $n \geq 2$. Démontrer et illustrer sur un dessin les inclusions suivantes :

$$\bigcup_{k=0}^{n-1} \left[\frac{k}{n}, \frac{k+1}{n}\right] \times \left[0, f\left(\frac{k+1}{n}\right)\right] \ \subset\ D\ \subset\ \bigcup_{k=0}^{n-1} \left[\frac{k}{n}, \frac{k+1}{n}\right] \times \left[0, f\left(\frac{k}{n}\right)\right]$$

5. En déduire

$$\sum_{k=0}^{n-1} \frac{1}{n} f\left(\frac{k+1}{n}\right) \ \leq \ P\left((X,Y) \in D\right) \ \leq \ \sum_{k=0}^{n-1} \frac{1}{n} f\left(\frac{k}{n}\right)$$

6. Conclure de ce qui précède que $\quad P\left((X,Y) \in D\right) = \dfrac{\pi}{4}$.

7. Ecrire un algorithme s'appuyant sur la méthode décrite et renvoyant une valeur approchée de π.