Probabilités
Agrégation interne, Année 18/19

1 Lois normales

Exercice 1.1. [Quelques calculs avec la loi centrée réduite]

La densité d'une loi normale centrée réduite est donnée par $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. On considère une variable aléatoire Z de loi $\mathcal{N}(0,1)$.

1. Calculer l'espérance et la variance de Z.
   Calculer pour tout $n \geq 0$, $E[Z^n]$.
   Calculer pour $t \in \mathbb{R}$, $E[e^{tZ}]$. Retrouvez ainsi le résultat précédent.
2. Soit F la fonction de répartition de Z.
   Montrer que pour tout x, $F(-x) = 1 - F(x)$.
   Montrer que pour tout $x \geq 0$, $\mathbb{P}(|Z| \leq x) = 2F(x) - 1$.
   Montrer que quand x tend vers $+\infty$, $1 - F(x) \sim \frac{1}{x} f(x)$.
   Valeurs de F à connaître : $F(1,96) \simeq 0,975, \quad F(2,58) \simeq 0,995$.
3. Déterminer la loi de Z^2.

Exercice 1.2. [Passage de la loi normale centrée réduite à une loi normale quelconque]
Soit Z une variable aléatoire de loi $\mathcal{N}(0,1)$. Soit σ, m des réels, montrer que $\sigma Z + m$ suit la loi $\mathcal{N}(m, \sigma^2)$.
Soit X une variable aléatoire de loi $\mathcal{N}(m, \sigma^2)$, montrer que $\frac{1}{\sigma}(X - m)$ suit la loi $\mathcal{N}(0,1)$.

Exercice 1.3 (Méthode de Box et Müller)**.**

1. Soit X et Y deux vas indépendantes de loi $\mathcal{N}(0,1)$. Déterminer la loi de $X^2 + Y^2$.
2. Soit U et V deux variables aléatoires indépendantes de loi uniforme sur $[0,1]$.
   Posons $S = -2\ln U$ et $\Theta = 2\pi V$. Montrer que S et Θ sont deux variables aléatoires indépendantes de lois respectives la loi exponentielle de paramètre $\frac{1}{2}$ et la loi uniforme sur $[0, 2\pi]$.
   Montrer que $X = \sqrt{S}\cos\Theta$ et $Y = \sqrt{S}\sin\Theta$ sont indépendantes de loi $\mathcal{N}(0,1)$.

Exercice 1.4. Sur un grand nombre de personnes on a constaté que la répartition du taux de cholestérol suit une loi normale avec les résultats suivants :
- 56% ont un taux inférieur à 165 cg ;
- 34% ont un taux compris entre 165 cg et 180 cg ;
- 10% ont un taux supérieur à 180 cg.

Quelle est le nombre de personnes qu'il faut prévoir de soigner dans une population de 10 000 personnes, si le taux maximum toléré sans traitement est de 182 cg ?

2 Théorème central limite, intervalle de confiance

Exercice 2.1. Utiliser l'inégalité de Chebyshev pour trouver un nombre n de jets tel que la probabilité d'obtenir une proportion de piles entre 49% et 51% soit au moins égale à 96%. (On suppose la pièce équilibrée).

Exercice 2.2. On lance une pièce équilibrée et on souhaite obtenir une proportion de ≪ piles ≫ entre 49% et 51% avec une probabilité au moins égale à 96%. Déterminer le nombre de jets nécessaire en utilisant l'approximation par une loi normale. Comparer avec l'exercice précédent.

Exercice 2.3. On veut estimer la probabilité p d'obtenir un 6 lorqu'on lance un dé.
Pour ceci on lance le dé 12000 fois.
On obtient 1890 fois le 6.
Déterminer un intervalle de confiance de niveau 95%.

3 Méthode de Monte-Carlo

Exercice 3.1. [Calcul d'intégrale par la méthode de Monte-Carlo]
Soit f une application continue de $[0,1]^d$ dans $\mathbb{R}$.
On désire calculer $I = \int_{[0,1]^d} f(x)dx$.
On considère une suite $(X_n)_{n\geq 1}$ de variables aléatoires indépendantes de loi uniforme sur $[0,1]^d$.
On note $T_n = \frac{f(X_1)+\cdots+f(X_n)}{n}$.

1. Pourquoi la suite $(T_n)_{n\geq 1}$ converge t-elle presque sûrement vers I ?
2. On note $\sigma^2 = \int_{[0,1]^d}(f(x)-I)^2dx$.
   Donner la variance de T_n.
   Le théorème central limite dit que $\sqrt{n}\frac{T_n - I}{\sigma}$ converge en loi vers une loi normale centrée réduite.
3. Donner un intervalle de confiance asymptotique au niveau 95% pour I.
4. Le problème est que l'intervalle précédent dépend de σ que l'on ne connait pas.
   Soit $\widehat{\sigma}_n^2 = \frac{1}{n-1}\sum_{k=1}^{n}(f(X_k)-T_n)^2$.
   Montrer que $E[\widehat{\sigma}_n^2] = \sigma^2$ et que $\widehat{\sigma}_n^2$ converge p.s. vers σ^2.
   Donner alors un intervalle de confiance asymptotique au niveau 95% pour I ne faisant pas intervenir σ.
5. On suppose qu'il existe K tel que pour tout $x, y \in [0,1]^d$, $|f(x)-f(y)| \leq K\|x-y\|$.
   Soit $n \geq 1$ et $I_n = \frac{1}{n^d}\sum_{1\leq i_1,\cdots,i_d\leq n} f(\frac{i_1}{n},\cdots,\frac{i_d}{n})$.
   Montrer que $|I - I_n| \leq \frac{K\sqrt{d}}{n}$.

6. Comparer en termes de coût de calcul et de précision la méthode probabiliste de Monte-Carlo et la méthode déterministe pour calculer une valeur approchée de I.

Exercice 3.2. [Un calcul d'aire simple]
Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et de loi uniforme sur $C = [-1,1]\times[-1,1]$. Soit D le disque de rayon 1 et de centre $(0,0)$.

1. Calculer $p = P(X_n \in D)$.
2. Expliciter une méthode de Monte-Carlo pour estimer p. Donner la variance associée.
3. Trouver le nombre d'itérations pour que l'estimation de p soit précise à 0,1 près avec une probabilité supérieure ou égale à 0,95.

Exercice 3.3. [Comparaisons d'estimateurs]
Soit X une loi de Cauchy de densité $\frac{1}{\pi(1+x^2)}$.
On veut calculer $p = P(X > 2)$ par une méthode de Monte Carlo. Calculer la variance des variables aléatoires utilisés dans les exemples suivants :

1. On pose $g(x) = \mathbf{1}_{x>2}$ et on considère $(X_1,\dots,X_n)$ un n-échantillon de X.

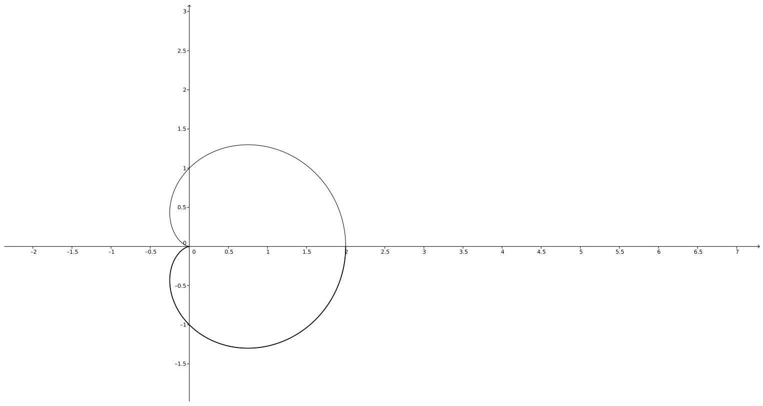
   On estime p par $\frac{1}{n}\sum_{i=1}^{n} g(X_i)$.
2. Montrer que $p = \frac{1}{2}P(|X| < 2)$.

   On estime p par $\frac{1}{2n}\sum_{i=1}^{n} g(|X_i|)$.
3. Montrer que $p = \frac{1}{2} - \int_0^2 \frac{1}{\pi(1+x^2)}dx$.
   On considère $(X_1,\dots,X_n)$ un n-échantillon de la loi uniforme sur $[0,2]$.

   On estime p par $\frac{1}{2} - \frac{2}{n}\sum_{i=1}^{n} \dfrac{1}{\pi(1+X_i^2)}$
4. Montrer que $p = \int_0^{\frac{1}{2}} \frac{x^{-2}}{\pi(1+x^{-2})}dx$.
   Soit h définie par $h(x) = \mathbf{1}_{x>2}\frac{2}{x^2}$. Montrer que h est une densité de probabilité.
   Soit $(U_1,\dots,U_n)$ un n-échantillon de loi uniforme sur $[0,1]$. Soit $X_i = \frac{2}{U_i}$.
   Soit $\psi(x) = \mathbf{1}_{x>2}\frac{f(x)}{h(x)}$.

   Montrer qu'on peut estimer p par $\frac{1}{n}\sum_{i=1}^{n}\psi(X_i)$.

Exercice 3.4. [Un calcul de volume]
On considère la cardioïde définie par $(x^2+y^2-x)^2 = x^2+y^2$.



1. Montrer que son équation en polaire est donnée par $\rho = 1 + \cos\theta$.
2. On considère le volume de révolution obtenu en faisant tourner la cardioïde autour de l'axe des x.

   Proposer une méthode probabiliste pour estimer le volume du corps obtenu (un genre de pomme).

   Implanter la avec un logiciel de votre choix.
3. Faire le calcul exact du volume.

Exercice 3.5. [Exemple jouet pour calculer une série]

Soit pour $n \geq 1$, $p_n = \frac{1}{n(n+1)}$ et $q_n = \displaystyle\sum_{k=1}^{n} p_k$.

1. Calculer q_n en fonction de n. Que vaut $\displaystyle\sum_{n=1}^{+\infty} p_n$?
2. Soit U une variable aléatoire de loi uniforme sur $[0,1]$. On définit la variable aléatoire X comme étant la partie entière de $\frac{1}{U}$. Déterminer la loi de X.
3. Proposer une méthode probabiliste pour estimer $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^3}$. (On écrira $\displaystyle\sum_{n=1}^{+\infty} \frac{1}{n^3} = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \frac{n+1}{n^2}$)

4 Un peu de statistiques

Exercice 4.1. [Un exemple d'estimateur]

On suppose qu'une certaine donnée suit une loi uniforme sur $[0,\theta]$ et on cherche à estimer le paramètre θ.

Soit $X_1, \ldots, X_n$ un échantillon de loi uniforme sur $[0,\theta]$.

On définit la statistique T_n par $T_n = \max\{X_1, \ldots, X_n\}$.

1. Calculer $E[T_n]$. L'estimateur T_n est-il sans biais ? asymptotiquement sans biais ?

2. Montrer que T_n converge en loi vers θ. T_n est-il un estimateur convergent ? fortement convergent ?
3. Calculer le risque quadratique de l'estimateur.
4. Donner un intervalle de confiance non asymptotique pour θ au niveau 95%.

Exercice 4.2. [Repas]
Un restaurateur peut servir 75 repas uniquement sur réservation. En pratique, 20% des clients qui réservent ne viennent pas. Le restaurateur souhaite pouvoir servir les clients qui se présentent avec une probabilté supérieur à 90%. Déterminer le nombre maximal de clients que le restaurateur peut accepter.