TD15 - Estimation

1 Estimation ponctuelle

Exercice 1 ⋆

Soit X une variable suivant une loi de Bernoulli de paramètre p inconnu. Pour n entier naturel non nul, on désigne par $(X_1, \dots, X_n)$ un n-échantillon de variables aléatoires indépendantes de même loi que X. On note $\overline{X_n} = \frac{X_1 + \cdots + X_n}{n}$.

1. Montrer que $\overline{X_n}$ est un estimateur de p.
2. Montrer que l'estimateur $\overline{X_n}$ a une espérance égale à p.

Exercice 2 ⋆⋆

On considère la variable aléatoire X dont la loi est donnée par : $P(X = -1) = p$, $P(X = 0) = 1 - 2p$ et $P(X = 1) = p$.
Pour un certain paramètre $p \in \left]0, \frac{1}{2}\right[$. On dispose d'un n-échantillon $(X_1, \dots, X_n)$ de X, et on cherche à déterminer le paramètre p.

1. Calculer $E(\overline{X_n})$.
2. Peut-on trouver les réels a et b tels que $aX_n + b$ ait pour espérance p ?
3. On note $T_n = \frac{1}{2n} \sum_{i=1}^{n} X_i^2$. Montrer que T_n est un estimateur d'espérance p. Calculer sa variance et montrer qu'il est convergent.

Exercice 3 ⋆⋆

Soit X une variable aléatoire admettant une variance σ^2. On désigne par $(X_1, \dots, X_n)$ un n-échantillon de variables aléatoires indépendantes de la loi de X et on pose $\overline{X_n} = \frac{1}{n} \sum\limits_{i=1}^{n} X_i$ et$T_n = \frac{1}{n-1} \sum\limits_{i=1}^{n} \left(X_i - \overline{X_n}\right)^2$.
Montrer que $E(T_n) = \sigma^2$.

Exercice 4 - Max de vraisemblance ⋆⋆

Soit $X_1, \dots, X_n$ des variables aléatoires mutuellement indépendantes suivant une loi de Poisson de paramètre $\lambda > 0$.
On pose $f_{x_1,\dots,x_n}(\lambda) = P(X_1 = x_1, \dots, X_n = x_n)$

1. Donner une expression de $f_{x_1,\dots,x_n}(\lambda)$.
2. Pour quelle valeur de λ la fonction $f_{x_1,\dots,x_n}$ est-elle maximale ?

Exercice 5 - Estimation de la variance ⋆⋆

Soit $(X_1, \dots, X_n)$ un n-échantillon d'une v.a. X admettant pour espérance m et pour variance σ^2.

1. On suppose que m est connu. Montrer que T_n est un estimateur de σ^2, où $T_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - m)^2$ et calculer son espérance.
2. On suppose que m n'est pas connu. On note $\overline{S_n}^2$ la variance empirique, où $\overline{S_n}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X_n})^2$
   (a) Montrer que, pour tout $i \in [\![1, n]\!]$, $E((X_i - \overline{X_n})^2) = V(X_i - \overline{X_n})$
   (b) D'autre part, montrer que, pour tout $i \in [\![1, n]\!]$, $E((X_i - \overline{X_n})^2) = V(X_i) - 2E\big((X_i - m)(\overline{X_n} - m)\big) + V(\overline{X_n})$ (on pourra ajouter et retrancher m dans le carré)
   (c) En déduire l'espérance de la variance empirique. En déduire un estimateur sans biais de σ^2.

Exercice 6 - Convergence ⋆⋆⋆

Soit a un réel strictement positif. On note f la fonction définie sur $\mathbb{R}$ par $f(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{3a^2}{x^4} & \text{si } x \geqslant a \end{cases}$.

1. Montrer que f est une densité de probabilité d'une var T et que T admet une espérance et une variance que l'on calculera.
2. (a) Déterminer la fonction de répartition de T.
   (b) Calculer les probabilités $P(T > 2a)$ et $P_{[T>2a]}(T > 6a)$.
3. Montrer que la variable $Z_n = \frac{2}{3n} \sum_{k=1}^{n} T_k$ est un estimateur de a dont on calculera l'espérance.
4. Montrer que pour tout $\epsilon > 0$, on a :
$$\lim_{n \to +\infty} P\left(|Z_n - a| \geqslant \epsilon\right) = 0.$$
   On dit alors que Z_n est un estimateur de a convergent.

2 Estimation par intervalle de confiance

Exercice 7 - Bienaymé-Tchebychev ⋆⋆

On suppose que le paramètre p, qui exprime la proportion de votants pour un candidat au second tour d'une élection, est inconnu, et on cherche à l'estimer. Pour tout entier n supérieur ou égal à 1, on considère un

n-échantillon $(X_1, \dots, X_n)$ de v.a.i.i.d de même loi de Bernoulli p. On pose $\overline{X_n}$ la moyenne empirique.
Calculer l'espérance et l'écart-type de $\overline{X_n}$ et à l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que :

$$\left[\overline{X_n} - \sqrt{\frac{5}{n}}\ ,\ \overline{X_n} + \sqrt{\frac{5}{n}}\right]$$

est un intervalle de confiance de p au niveau de confiance 0,95.

Exercice 8 ⋆⋆

On suppose que la charge maximale supportée par un câble, exprimée en tonnes, est une v.a. qui suit une loi normale d'écart-type 0,5 et de moyenne inconnue. Une étude portant sur 50 câbles a donné une moyenne des charges maximales supportées de 12,2 tonnes.

1. Déterminer un intervalle de confiance à 99% de la charge maximale de tous les câbles fabriqués par l'usine.
2. Déterminer la taille minimale de l'échantillon étudié pour que la longueur de l'intervalle de confiance à 99% soit inférieure ou égale à 0,2.

Exercice 9 ⋆⋆

Lors d'un sondage pré-électoral, 1000 électeurs choisis au hasard ont été interrogés. 520 d'entre eux se sont déclarés favorables à Mme Durand.

1. Déterminer un intervalle de confiance à 95% de la proportion p d'électeurs favorables à Mme Durand dans la population.
2. Quel doit être le nombre minimal d'électeurs interrogés pour que l'intervalle de confiance que l'on puisse en déduire soit de longueur inférieure ou égale à 0,02 ?

Exercice 10 ⋆⋆

Soient $X_1, \dots, X_n$ des variables aléatoires i.i.d. suivant la loi normale d'espérance θ inconnue et de variance 1. On pose $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$.

1. Quelle est la loi de $\overline{X_n}$?
2. Soit $\alpha \in]0,1[$ et soit t_α l'unique réel tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.
   Montrer que $\left[\overline{X_n} - \frac{t_\alpha}{\sqrt{n}}, \overline{X_n} + \frac{t_\alpha}{\sqrt{n}}\right]$ est un intervalle de confiance de θ au niveau de confiance $1 - \alpha$.
3. On suppose à présent que les X_i suivent la loi $\mathcal{N}(\theta, \sigma^2)$ où σ^2 est connu. Sur le même principe, proposer un intervalle de confiance de θ au niveau de risque α.

Exercice 11 ⋆⋆

On suppose que la probabilité qu'un individu contagieux transmette un virus à un individu sain est $p \in]0,1[$ inconnu et que l'on cherche à évaluer.
Soit $(Y_n)_{n\in\mathbb{N}^\star}$ une suite de variables i.i.d. suivant la loi de Bernoulli de paramètre p.

1. On pose $\overline{Y_n} = \frac{1}{n}\sum_{i=1}^n Y_i$. Montrer que $\overline{Y_n}$ est un estimateur sans biais de p.
2. À l'aide de l'inégalité de Bienaymé-Tchebychev, montrer que $\left[\overline{Y_n} - \sqrt{\frac{5}{n}}, \overline{Y_n} + \sqrt{\frac{5}{n}}\right]$ est un intervalle de confiance de p au niveau de confiance $0,95$.

Exercice 12 ⋆⋆

Soit $(X_i)_{i\in\mathbb{N}^\star}$ une suite de variables i.i.d. d'espérance μ inconnue et de variance σ^2 connue. Déterminer à l'aide du théorème central limite un intervalle de confiance asymptotique de μ au niveau de confiance $1 - \alpha$.

3 Exercices type concours

Exercice 13 - Recap ⋆⋆

Soient θ un réel strictement positif et $(X_1, \dots, X_n)$ un $n-$échantillon d'une variable $X \hookrightarrow \mathcal{U}\big([0,\theta]\big)$.

1. Montrer que $\overline{X_n}$ est un estimateur de θ et calculer son espérance.
2. Proposer un estimateur V_n de θ dont l'espérance est θ.
3. On considère maintenant l'estimateur $M_n = \max(X_1, \dots, X_n)$.
   (a) Déterminer la fonction de répartition F de M_n et en déduire une densité f de M_n.
   (b) Montrer alors que $E(M_n) = \frac{n\theta}{n+1}$
   (c) En déduire un estimateur Z_n de θ dont l'espérance est θ à partir de M_n.
4. (a) Établir que $V(V_n) = \frac{\theta^2}{3n}$
   (b) Montrer que $E\big(Z_n^2\big) = \frac{(n+1)^2}{n^2}\int_0^\theta \frac{nt^{n+1}}{\theta^n}\mathrm{d}t$
   (c) En déduire que $V(Z_n) = \frac{\theta^2}{n(n+2)}$
   (d) Quel estimateur aura-t-on tendance à préférer en pratique pour des grandes valeurs de n ?
5. Soit $(x_1, \dots, x_n) \in \big(\mathbb{R}_+^*\big)^n$ fixé. On note f_θ la densité de X.
   On introduit la fonction de vraisemblance, définie sur $\mathbb{R}_*^+$ par $L_n(\theta) = \prod_{i=1}^n f_\theta(x_i)$.
   (a) Montrer que, pour tout $\theta \geq 0$, on a $L_n(\theta) = \begin{cases} \theta^{-n} & \text{si } \theta \geqslant \max(x_1, \dots, x_n) \\ 0 & \text{sinon} \end{cases}$.

   (b) En déduire que l'estimateur du maximum de vraisemblance pour la loi $\mathcal{U}([0,\theta])$ est donné par $\hat{\theta}_n = \max(X_1, \dots, X_n)$

6. (a) Montrer à l'aide de l'inégalité de Bienaymé-Tchebychev que $P(|V_n - \theta| > \epsilon) \leq \frac{\theta^2}{3n\epsilon^2}$ (V_n est donc un estimateur convergent de θ)
   (b) Montrer que $\theta \in \left[V_n - \sqrt{\frac{\theta^2}{3n\alpha}}; V_n + \sqrt{\frac{\theta^2}{3n\alpha}}\right] \Leftrightarrow \frac{V_n}{1+\frac{1}{\sqrt{3n\alpha}}} \leq \theta \leq \frac{V_n}{1-\frac{1}{\sqrt{3n\alpha}}}$
   (c) En déduire un intervalle de confiance au risque α pour θ.

Exercice 14 - EDHEC ECE 2012 ⋆⋆

On pourra aussi traiter le problème de l'année 2012 de l'EDHEC en remplaçant les questions portant sur Scilab par des questions portant sur Python :

5. (b) En déduire une fonction python dont l'entête est `def vax(lambda):` qui simule la loi de $|X|$.
   (c) Vérifier que la probabilité que X prenne des valeurs positives est égale à la probabilité que X prenne des valeurs positives est égale à la probabilité que X prenne des valeurs négatives.
   En déduire une fonction Python dont l'entête est `def Simulx(lambda):` qui simule la loi de X.

4 Mathématiques approfondies

4.1 Estimation ponctuelle

Exercice 15 ⋆

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes suivant la loi exponentielle $\mathcal{E}(\lambda)$. On pose :

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ et } T_n = \frac{1}{n+1}\sum_{i=1}^{n} X_i.$$

Calculer le biais de chacun de ces estimateurs de $\frac{1}{\lambda}$.

Exercice 16 ⋆⋆

Soit $(X_n)_{n\in\mathbb{N}^\star}$ un échantillon de la loi de Poisson de paramètre λ. Pour tout $n \geq 1$, on pose :

$$S_n = \sum_{i=1}^{n} X_i \text{ et } Y_n = \mathrm{e}^{-\frac{S_n}{n}}.$$

Y_n est-il un estimateur sans biais de $\mathrm{e}^{-\lambda}$? Est-il asymptotiquement sans biais ?

Exercice 17 ⋆⋆

Soit (X_n) une suite de variables aléatoires indépendantes suivant la loi uniforme $[a,b]$ où b est connu et l'on cherche à déterminer a. Pour tout $n \geq 1$, on pose

$$Y_n = \min(X_1, \dots, X_n).$$

1. Montrer que Y_n est une variable à densité, et en donner une densité.
2. Déterminer $E(Y_n)$, puis le biais de Y_n en a. Comment interpréter le signe de ce biais ?
3. En utilisant la question précédente, proposer un estimateur sans biais de a.
4. Montrer que Y_n est un estimateur asymptotiquement sans bias et convergent de a.

Exercice 18 ⋆⋆

Soient $X_1, \dots, X_n$ des variables aléatoires i.i.d. telles que :

$$P(X_1 = -1) = (1-p)^2, \quad P(X_1 = 0) = 2p(1-p),$$
$$P(X_1 = 1) = p^2.$$

On cherche dans la suite à estimer $p \in]0,1[$. Montrer que $Z_n = \sum_{k=1}^{n} \frac{1+X_k}{2n}$ est un estimateur sans biais et convergent de p.

Exercice 19 ⋆⋆

Soit $(X_n)_{n\geq 1}$ une suite i.i.d. de variables aléatoires suivant la loi $\mathcal{B}(p)$ où p est inconnu. On pose :

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k \text{ et } T_n = \frac{2}{n(n+1)}\sum_{k=1}^{n} kX_k.$$

1. Montrer que $\overline{X_n}$ et T_n sont deux estimateurs sans biais de p.
2. Montrer que $\overline{X_n}$ et T_n sont deux estimateurs convergents de p.

Exercice 20 ⋆⋆⋆

Soit X une variable aléatoire suivant la loi normale centrée de variance σ^2 avec σ inconnu. Soit $(X_n)_{n\in\mathbb{N}^\star}$ une suite de variables indépendantes, de même loi que X. On note alors :

$$S_n = \frac{1}{n}\sum_{k=1}^{n} X_k^2.$$

Montrer que S_n est un estimateur sans biais et convergent de σ^2.

Exercice 21 ⋆⋆⋆

Soit θ une réel strictement positif et soient $(X_k)_{k\in\mathbb{N}^\star}$ une suite de variables i.i.d. suivant la loi uniforme sur $[0,2\theta]$. On pose $M_n = \max(X_1,\dots,X_n)$ et $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$.

1. Déterminer la loi de M_n, calculer son espérance et sa variance.
2. En déduire que $U_n = \frac{n+1}{2n}M_n$ est un estimateur sans biais de θ.
3. Calculer $P(|U_n-\theta|\geq\epsilon)$ pour tout n et pour tout $\epsilon>0$. Faire de même pour $\overline{X_n}$.
   De $\overline{X_n}$ et U_n lequel est un meilleur estimateur de θ ?

Exercice 22 ⋆⋆⋆

Soit θ un réel strictement positif, et soit f_θ la fonction définie sur $\mathbb{R}$ par :

$$f_\theta(x) = \begin{cases} 0 & \text{si } x\leq\theta \\ \mathrm{e}^{-(x-\theta)} & \text{si } x>\theta \end{cases}.$$

1. Montrer que f est une densité de probabilités.
2. Soit X une variable aléatoire dont la densité est f_θ. Calculer $E(X)$ et $V(X)$.
3. Dans toute la suite $(X_n)_{n\geq1}$ est une suite de variables aléatoires indépendantes sur $(\Omega,\mathcal{A})$ de densité f_θ. Pour tout $n\geq1$, on pose :

$$\overline{X_n} = \frac{X_1+\cdots+X_n}{n} \text{ et } Y_n = \min(X_1,\dots,X_n).$$

   (a) À partir de $\overline{X_n}$, construire un estimateur T_n sans biais et convergent de θ.
   (b) Montrer que Y_n est un estimateur asymptotiquement sans biais et convergent de θ.
   On pourra commencer par étudier $Y_n-\theta$.
   (c) Construire à partir de Y_n un estimateur U_n sans biais et convergent de θ.
   (d) Comparer les deux estimateurs T_n et U_n.

Exercice 23 ⋆⋆⋆

Soit $(X_n)_{n\geq1}$ une suite de variables aléatoires indépendants suivant la loi exponentielle de paramètre λ. On pose :

$$S_n = \sum_{k=1}^n X_k.$$

1. Déterminer une densité de S_n.
2. Calculer l'espérance de $\frac{1}{S_n}$.
3. En déduire un estimateur sans biais de λ.

Exercice 24 ⋆⋆⋆⋆

Soit $X_1,\dots,X_n$ un n-échantillon de la loi de Poisson de paramètre $\lambda>0$ inconnu. On cherche à estimer $\mathrm{e}^{-\lambda}$. On définit des variables de Bernoulli $Y_1,\dots,Y_n$ par :

$$Y_k = \begin{cases} 1 & \text{si } X_k=0 \\ 0 & \text{sinon} \end{cases}.$$

On pose alors :

$$\overline{Y_n} = \frac{1}{n}\sum_{k=1}^n Y_k \text{ et } S_n = \sum_{k=1}^n X_k.$$

1. (a) Montrer que $\overline{Y_n}$ est un estimateur sans biais de $\mathrm{e}^{-\lambda}$.
   (b) Calculer $V(\overline{Y_n})$ et en déduire que $\overline{Y_n}$ est un estimateur convergent de $\mathrm{e}^{-\lambda}$.
2. Pour $j\in\mathbb{N}$, on pose $\varphi(j) = P_{[S_n=j]}(X_1=0)$. Calculer $\varphi(j)$.
3. On pose à présent $T_n = \varphi(S_n)$.
   (a) Prouver que T_n est un estimateur sans biais de $\mathrm{e}^{-\lambda}$.
   (b) Calculer $V(T_n)$ et en déduire que T_n est un estimateur convergent de $\mathrm{e}^{-\lambda}$.
4. Calculer $E((\overline{Y_n}-\mathrm{e}^{-\lambda})^2)$ et $E((T_n-\mathrm{e}^{-\lambda})^2)$. Quel est le meilleur des deux estimateurs ?

4.2 Estimation par intervalle de confiance

Exercice 25 ⋆⋆⋆

Soit $(X_i)_{i\in\mathbb{N}^\star}$ une suite de variables aléatoires i.i.d. suivant la loi uniforme $\mathcal{U}([0,\theta])$, θ inconnu. On pose : $M_n = \max(X_1,\dots,X_n)$.

1. Déterminer la fonction de répartition de M_n.
2. Montrer que pour $\theta>0$ fixé, il existe deux uniques réels x_1 et x_2, que l'on exprimera en fonction de θ tels que $F_{M_n}(x_1) = 0,025$ et $F_{M_n}(x_1) = 0,975$.
3. En déduire un intervalle de confiance de θ au niveau de confiance $0,95$.
4. Proposer un autre intervalle de θ au niveau de confiance $0,95$ de la forme $[M_n, U_n]$ où U_n est à déterminer. **Plus difficile :** lequel de ces deux intervalles possède l'étendue la plus faible ?

Exercice 26 ⋆⋆⋆

Soient $(X_i)_{i\geq1}$ des variables aléatoires i.i.d. suivant la loi exponentielle $\mathcal{E}(\lambda)$ où $\lambda>0$ est inconnu. On pose $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i$.

1. Montrer que $\lambda\sqrt{n}\overline{X_n}-\sqrt{n}$ converge en loi vers une variable suivant la loi normale centrée réduite.

2. Soit $\alpha \in]0,1[$ et soit t_α l'unique réel tel que $\Phi(t_\alpha) = 1 - \frac{\alpha}{2}$.
Montrer que $\left[\left(1 - \frac{t_\alpha}{\sqrt{n}}\right)\frac{1}{\overline{X_n}}, \left(1 + \frac{t_\alpha}{\sqrt{n}}\right)\frac{1}{\overline{X_n}}\right]$ est un intervalle de confiance asymptotique de λ au niveau de risque α.

Exercice 27 ⋆⋆⋆⋆

Soit $(X_i)_{i\in\mathbb{N}^\star}$ une suite de variables aléatoires i.i.d. suivant la loi uniforme $\mathcal{U}([0,\theta])$, θ inconnu. On pose : $M_n = \max(X_1, \dots, X_n)$.

1. Montrer que $\left(n\left(1 - \frac{M_n}{\theta}\right)\right)$ converge en loi vers une variable aléatoire X suivant la loi exponentielle $\mathcal{E}(1)$.
2. Déterminer un intervalle de confiance asymptotique de θ au niveau de risque α, sous la forme $[M_n, U_n]$.

4.3 Exercices de concours

Exercice 28 - ESSEC ECE 2007 ⋆⋆⋆

Pour $a > 0$, on note $f_a : t \mapsto \begin{cases} 0 & \text{si } t < 1 \\ \frac{a}{t^{a+1}} & \text{si } t \geq 1 \end{cases}$. On admet que f_a est une densité.
On considère alors la suite $(X_n)_{n\in\mathbb{N}^\star}$ de variables aléatoires i.i.d. de densité f_a où a est inconnu. Pour $n \in \mathbb{N}^\star$, on définit une fonction L_n (appelée fonction de vraisemblance) :

$$L_n : [1,+\infty[^n \times \mathbb{R}_+^\star,\ (x_1,\dots,x_n,a) \mapsto \prod_{i=1}^n f_a(x_i).$$

1. Pour $(x_1,\dots,x_n)$ fixés, montrer que :
$a \mapsto L_n(x_1,\dots,x_n,a)$
possède un maximum atteint en un unique réel a que l'on exprimera en fonction de $x_1,\dots,x_n$.

On pourra étudier $a \mapsto \ln(L_n(x_1,\dots,x_n,a))$.

On a donc trouvé une expression $a = \varphi(x_1,\dots,x_n)$ où φ_n est une fonction définie sur $[1,+\infty[^n$. On pose alors $T_n = \varphi_n(X_1,\dots,X_n)$ et on appelle T_n estimateur du maximum de vraisemblance.

2. Montrer que pour $k \in [\![1,n]\!]$, $\ln(X_k)$ suit une loi exponentielle dont on précisera le paramètre. En déduire une densité de $S_n = \ln(X_1) + \dots + \ln(X_k)$.
3. Exprimer T_n en fonction de S_n. En déduire $E(T_n)$ et $V(T_n)$.
4. Montrer que T_n est un estimateur asymptotiquement sans biais et convergent de a.

Exercice 29 - QSP ESCP 2007 ⋆⋆⋆⋆⋆

Soient $n \geq 1$ et $(X_1,\dots,X_n)$ un échantillon identiquement distribué indépendant de la loi de Poisson de paramètre $\lambda > 0$ inconnu. On pose :

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i \text{ et } T_n = \sqrt{n}\frac{\overline{X_n} - \lambda}{\sqrt{\lambda}}.$$

En utiliser T_n, déterminer un intervalle de confiance asymptotique de λ au niveau de risque α.

Exercice 30 - Oral ESCP 2012 ⋆⋆⋆⋆⋆

On cherche à évaluer le nombre N de lions d'Asie, espèce en voie de disparition, encore en vie dans la forêt de Gir. Pour cela, on capture d'abord en une seule fois m lions ($m \in \mathbb{N}^\star$) que l'on tatoue avant de les relacher dans la nature et on admet que pendant toute la durée de l'étude, il n'y a ni naissance ni décès, puis l'on utilise l'une des deux méthodes suivantes.

1. **Méthode 1 :** On capture successivement au hasard (donc avec équiprobabilité) et avec remise en liberté après observation du sujet, n lions. Soit Y_n le nombre de lions tatoués parmi eux.
   (a) Déterminer la loi de Y_n. En déduire que $\frac{1}{nm}Y_n$ est un estimateur sans biais et convergent de $\frac{1}{N}$.
   (b) Pourquoi ne peut-on pas prendre $\frac{nm}{Y_n}$ comme estimateur de N ?
   (c) On pose $B_n = \frac{m(n+1)}{Y_n+1}$. Calculer l'espérance de B_n et montrer que B_n est un estimateur asymptotiquement sans biais de N.
2. **Méthode 2 :** On se donne $n \in \mathbb{N}^\star$. On capture également, un par un, et avec remise en liberté après observation du sujet, des lions de Gir. On note X_n la variable aléatoire égale au nombre de lions qu'il a été juste nécessaire de capturer pour en obtenir n tatoués.
   On pose $D_1 = X_1$ et pour tout $i \in [\![2,n]\!]$, $D_i = X_i - X_{i-1}$. On admet que les D_i sont mutuellement indépendantes.
   (a) Pour tout $i \in [\![2,n]\!]$, que représente concrètement D_i ?
   (b) Déterminer, pour tout $i \in [\![2,n]\!]$, la loi de D_i son espérance et sa variance. En déduire l'espérance et la variance de X_n.
   (c) On pose $A_n = \frac{m}{n}X_n$. Montrer que A_n est estimateur sans biais et convergent de N.