II LA METHODE DU MAXIMUM DE VRAISEMBLANCE

Ceci n'est pas au programme mais c'est très courant.

X est une variable aléatoire réelle et $\mathcal{H}$ l'ensemble de ses valeurs. θ est un paramètre réel associé à X que l'on souhaite estimer.

$(X_1, X_2, \ldots, X_n)$ étant un n-échantillon indépendant de X nous allons donner une méthode pour construire un estimateur de θ c'est à dire une variable aléatoire $T_n = T(X_1, X_2, \ldots, X_n)$ fonction de $(X_1, X_2, \ldots, X_n)$.

Supposons que X est une variable aléatoire réelle discrète.

Si $(x_1, x_2, \ldots, x_n)$ est une réalisation de $(X_1, X_2, \ldots, X_n)$, c'est à dire un élément de $\mathcal{H}^n$, par indépendance on a : $P(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \ldots \cap \{X_n = x_n\}) = P(X_1 = x_1)\, P(X_2 = x_2) \ldots P(X_n = x_n)$.

Observons que $P(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \ldots \cap \{X_n = x_n\}) = P(X_1 = x_1)\, P(X_2 = x_2) \ldots P(X_n = x_n)$ est une fonction de θ.

On pose : $L(x_1, x_2, \ldots, x_n, \theta) = P(X_1 = x_1)\, P(X_2 = x_2) \ldots P(X_n = x_n) = \displaystyle\prod_{k=1}^{n} P(X = xk)$

$\theta \to L(x_1, x_2, \ldots, x_n, \theta)$ est la fonction de vraisemblance (L comme likelihood).

L'idée consiste alors à $\boxed{\textbf{trouver la valeur } \hat{\theta} \textbf{ qui rend maximum la fonction } \theta \to \mathbf{L(x_1, x_2, \ldots, x_n, \theta)}}$.

C'est à dire la valeur de θ qui assure la plus grande probabilité d'obtenir ces observations ; c'est à dire encore qui rend le plus vraissemblable la réalisation de l'événement $\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap \ldots \cap \{X_n = x_n\})$.

Clairement $\hat{\theta}$ dépend de x_1, x_2, ..., x_n. $\hat{\theta}$ est une fonction de x_1, x_2, ..., x_n.

Ainsi $\hat{\theta} = \varphi(x_1, x_2, \ldots, x_n)$ où φ est une fonction de n variables.

La seconde idée est alors de poser $\mathbf{T_n} = \varphi(\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_n})$ et de vérifier que $\mathbf{T_n}$ est un estimateur raisonnable de θ.

Le plus souvent on ne travaille pas sur $\theta \to L(x_1, x_2, \ldots, x_n, \theta)$ mais sur son logarithme pour se ramener à une somme.

Notons que les valeurs de θ qui rendent maximum $\theta \to L(x_1, x_2, \ldots, x_n, \theta)$ sont les mêmes que celles qui rendent maximum $\theta \to \ln\Big(L(x_1, x_2, \ldots, x_n, \theta)\Big)$

Si X est une variable aléatoire à densité de densité f : $L(x_1, x_2, \ldots, x_n, \theta) = \displaystyle\prod_{k=1}^{n} f(x_k)$.

Exercice 1 **PC et CC** ★ Construire un estimateur du paramètre λ d'une variable aléatoire suivant une loi de Poisson de paramètre λ en utilisant la méthode du maximum de vraisemblance.

Soit $(x_1, x_2, \dots, x_n) \in \mathbb{N}^n$. Cherchons la valeur qui rend maximum $\lambda \mapsto \prod_{k=1}^{n} P(X_k = x_k)$ où $(X_1, X_2, \dots, X_n)$ est un n-échantillon i.i.d de la loi de Poisson de paramètre λ. Nous supposerons que $\exists i \in [\![1, n]\!]$, $x_i \neq 0$.

Cela revient à trouver la valeur qui rend maximum $\varphi : \lambda \to \ln \prod_{k=1}^{n} P(X_k = x_k)$.

$\forall \lambda \in \mathbb{R}_+^*$, $\varphi(\lambda) = \sum_{k=1}^{n} \ln\left(P(X_k = x_k)\right) = \sum_{k=1}^{n} \ln\left(\frac{\lambda^{x_k}}{(x_k)!} e^{-\lambda}\right)$.

$\forall \lambda \in \mathbb{R}_+^*$, $\varphi(\lambda) = \sum_{k=1}^{n} \left(x_k \ln \lambda - \ln((x_k)!) - \lambda\right)$. φ est dérivable sur $\mathbb{R}_+^*$.

et $\forall \lambda \in \mathbb{R}_+^*$, $\varphi'(\lambda) = \sum_{k=1}^{n}\left(x_k \frac{1}{\lambda} - 1\right) = \frac{1}{\lambda}\left[\sum_{k=1}^{n} x_k - n\lambda\right] = \frac{n}{\lambda}\left[\frac{1}{n}\sum_{k=1}^{n} x_k - \lambda\right]$.

Posons $m_n = \frac{1}{n}\sum_{k=1}^{n} x_k$. $\forall \lambda \in]0, m_n[$, $\varphi'(\lambda) > 0$, $\forall \lambda \in]m_n, +\infty[$, $\varphi'(\lambda) < 0$ et $\varphi'(m_n) = 0$.

Alors φ est strictement croissante sur $]0, m_n]$ et strictement décroissante sur $[m_n, +\infty[$.

Ainsi φ possède un maximum atteint en le seul point $m_n = \frac{x_1 + x_2 + \cdots + x_n}{n}$.

Posons alors $\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$. $\forall i \in [\![1, n]\!]$, $E(X_i)$ existe et vaut λ.

Alors $E(\overline{X}_n)$ existe et vaut $\frac{1}{n}\sum_{i=1}^{n} E(X_i)$ donc λ.

$\overline{X}_n$ est un estimateur sans biais de λ.

"les X_i" sont indépendantes, ont même espérance λ et même variance λ.

La loi faible des grands nombres montre alors que $\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)$ converge en probabilité vers λ. Finalement $\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$ est un estimateur sans biais et convergent de λ.

EX 20

Exercice Construire un estimateur du paramètre λ d'une variable aléatoire suivant la loi exponentielle de paramètre λ en utilisant la méthode du maximum de vraisemblance.

Montrer que cet estimateur est asymptotiquement sans biais et convergent.

Posons $\forall t \in \mathbb{R}$, $f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \in [0, +\infty[\\ 0 & \text{sinon} \end{cases}$. f est une densité d'une variable aléatoire qui suit la loi exponentielle de paramètre λ.

Soit $n \in \mathbb{N}^*$. Soit $(x_1, x_2, \dots, x_n) \in ([0, +\infty[)^n$ tel que $\sum_{k=1}^n x_k \neq 0$.

Posons $\forall \lambda \in \mathbb{R}_+^*$, $\varphi(\lambda) = \ln\left(\prod_{k=1}^n f(x_k)\right)$.

$\forall \lambda \in \mathbb{R}_+^*$, $\varphi(\lambda) = \sum_{k=1}^n \ln f(x_k) = \sum_{k=1}^n (\ln \lambda - \lambda x_k)$. φ est dérivable sur $\mathbb{R}_+^*$.

$\forall \lambda \in \mathbb{R}_+^*$, $\varphi'(\lambda) = \sum_{k=1}^n \left(\frac{1}{\lambda} - x_k\right) = \frac{n}{\lambda} - \sum_{k=1}^n x_k = n\left[\frac{1}{\lambda} - \frac{1}{n}\sum_{k=1}^n x_k\right]$.

$\forall \lambda \in \mathbb{R}_+^*$, $\varphi'(\lambda) \geq 0 \iff \lambda \leq \dfrac{n}{\sum_{k=1}^n x_k}$.

φ possède un maximum sur $\mathbb{R}_+^*$ atteint en $\dfrac{n}{\sum_{k=1}^n x_k}$.

Soit $(X_1, X_2, \dots, X_n)$ un n échantillon i.i.d de la loi exponentielle de paramètre λ. Posons $V_n = \dfrac{n}{\sum_{k=1}^n X_k}$.

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

$X_1, X_2, \dots, X_n$ sont indépendantes et $\forall k \in [\![1, n]\!]$, $X_k \hookrightarrow \mathcal{E}(\lambda)$ ou $X_k \hookrightarrow \Gamma(\frac{1}{\lambda}, 1)$.

Alors d'après le cours : $S_n \hookrightarrow \Gamma(\frac{1}{\lambda}, n)$. Supposons désormais que $\underline{\underline{n \geq 3}}$.

Montrons que V_n possède une espérance et un moment d'ordre 2.

Pour tout r dans $\mathbb{N}^*$ posons $\forall t \in \mathbb{R}$, $f_r(t) = \begin{cases} \dfrac{e^{-\lambda t}\, t^{r-1}}{(1/\lambda)^r\, \Gamma(r)} & \text{si } t \in]0, +\infty[\\ 0 & \text{sinon} \end{cases}$

Pour tout r dans $\mathbb{N}^*$, f_r est une densité d'une variable aléatoire qui suit la loi gamma de paramètres $\frac{1}{\lambda}$ et r.

En particulier f_n est une densité de S_n.

Soit $k \in [\![0, n-1]\!]$. Montrons que $V_n = \frac{n}{S_n}$ possède un moment d'ordre k c'est à dire que $\left(\frac{n}{S_n}\right)^k$ possède une espérance. D'après le théorème de transfert il suffit de montrer que $\int_{-\infty}^{+\infty} \frac{n^k}{t^k} f_n(t)\,dt$ est absolument convergente ou que $\int_0^{+\infty} \frac{n^k}{t^k} f_n(t)\,dt$ est absolument convergente ou encore que $\int_0^{+\infty} \frac{n^k}{t^k} f_n(t)\,dt$ converge ($\forall t \in]0,+\infty[$, $\frac{n^k}{t^k} f_n(t) \geq 0$).

$$\forall t \in]0,+\infty[,\ \frac{n^k}{t^k} f_n(t) = \frac{n^k}{t^k} \frac{e^{-\lambda t} t^{n-1}}{(1/\lambda)^n \Gamma(n)} = \frac{\Gamma(n-k)}{\Gamma(n)} \frac{e^{-\lambda t} t^{n-k-1}}{(1/\lambda)^{n-k} \Gamma(n-k)} \times (n\lambda)^k.$$

$\forall t \in]0,+\infty[$, $\frac{n^k}{t^k} f_n(t) = (n\lambda)^k \frac{\Gamma(n-k)}{\Gamma(n)} f_{n-k}(t)$. Ainsi $\int_0^{+\infty} \frac{n^k}{t^k} f_n(t)\,dt$ existe et vaut $(n\lambda)^k \frac{\Gamma(n-k)}{\Gamma(n)} \int_0^{+\infty} f_{n-k}(t)\,dt$ c'est à dire $(n\lambda)^k \frac{\Gamma(n-k)}{\Gamma(n)} \times 1$.

Ceci achève de montrer que $E\left(\left(\frac{n}{S_n}\right)^k\right)$ existe et vaut $(n\lambda)^k \frac{\Gamma(n-k)}{\Gamma(n)}$.

$\frac{n}{S_n}$ possède alors une espérance qui vaut $n\lambda \cdot \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{n\lambda}{n-1}$ ... car $1 \in [\![0, n-1]\!]$

$\frac{n}{S_n}$ possède un moment d'ordre 2 qui vaut $(n\lambda)^2 \frac{\Gamma(n-2)}{\Gamma(n)} = \frac{(n\lambda)^2}{(n-1)(n-2)}$ ... car $2 \in [\![0, n-1]\!]$

Alors $V\left(\frac{n}{S_n}\right)$ existe et vaut $\frac{(n\lambda)^2}{(n-1)(n-2)} - \left(\frac{n\lambda}{n-1}\right)^2$ ou $\frac{(n\lambda)^2}{(n-1)^2(n-2)}$.

$\lim_{n\to+\infty} E(V_n) = \lim_{n\to+\infty} E\left(\frac{n}{S_n}\right) = \lim_{n\to+\infty} \left(\frac{n\lambda}{n-1}\right) = \lambda$.

V_n est un estimateur asymptotiquement sans biais de λ.

V_n donc $V_n - \lambda$ possède un moment d'ordre 2. Rappelons donc alors

$$\forall \varepsilon \in \mathbb{R}_+^*,\ 0 \leq P(|V_n - \lambda| \geq \varepsilon) \leq \frac{E((V_n-\lambda)^2)}{\varepsilon^2} = \frac{V(V_n-\lambda) + (E(V_n-\lambda))^2}{\varepsilon^2} = \frac{V(V_n) + (E(V_n)-\lambda)^2}{\varepsilon^2}$$

$\lim_{n\to+\infty} V(V_n) = 0$ et $\lim_{n\to+\infty} (E(V_n)-\lambda)^2 = 0$. Alors $\lim_{n\to+\infty} P(|V_n-\lambda| \geq \varepsilon) = 0$ pour tout $\varepsilon \in \mathbb{R}_+^*$.

Finalement $\frac{n}{X_1+X_2+\dots+X_n}$ est un estimateur asymptotiquement sans biais et convergent de λ.

Remarque.. Il serait judicieux de passer de V_n à $\frac{n-1}{n} V_n$ ...

Ex 21

Exercice Q1. Construire un estimateur de l'espérance m d'une variable aléatoire suivant une loi normale de paramètres m et σ^2 ($\boxed{\sigma \text{ est donné}}$) en utilisant la méthode du maximum de vraisemblance.

Q2. Construire un estimateur de la **variance** σ^2 d'une variable aléatoire suivant une loi normale de paramètres m et σ^2($\boxed{m \text{ est donné}}$) en utilisant la méthode du maximum de vraisemblance ; on pourra étudier

$$v \to \ln\left(\prod_{k=1}^{n}\left(\frac{1}{\sqrt{2\pi v}}\,e^{-\frac{(x_k-m)^2}{2v}}\right)\right)$$

L'estimateur obtenu est-il sans biais ? convergent ?

Pour $\forall x \in \mathbb{R}$, $f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\,e^{-\frac{(x-m)^2}{2\sigma^2}}$. f est une densité d'une variable aléatoire qui suit la loi normale de paramètres m et σ^2.

Soit $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

$\ln \prod_{k=1}^{n} f(x_k) = \sum_{k=1}^{n} \ln f(x_k) = \sum_{k=1}^{n}\left(\ln\frac{1}{\sqrt{2\pi}} + \ln\frac{1}{\sigma} - \frac{(x_k-m)^2}{2\sigma^2}\right)$

$\ln \prod_{k=1}^{n} f(x_k) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-m)^2.$

Q1 Pour $\forall m \in \mathbb{R}$, $\varphi(m) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-m)^2$.

φ est dérivable sur $\mathbb{R}$ et $\forall m \in \mathbb{R}$, $\varphi'(m) = +\frac{1}{2\sigma^2}\times 2 \times \sum_{k=1}^{n}(x_k-m)$.

$\forall m \in \mathbb{R}$, $\varphi'(m) \geq 0 \Leftrightarrow \sum_{k=1}^{n}(x_k-m) \geq 0 \Leftrightarrow \sum_{k=1}^{n} x_k \geq n m \Leftrightarrow m \leq \frac{1}{n}\sum_{k=1}^{n} x_k$.

Alors φ admet un maximum sur $\mathbb{R}$ atteint en $\frac{1}{n}\sum_{k=1}^{n} x_k$.

Soit $(X_1, X_2, \dots, X_n)$ un n-échantillon i.i.d de la loi normale de paramètres m et σ^2. $\frac{X_1+X_2+\dots+X_n}{n}$ est un estimateur de m sans biais et convergent (loi faible des grands nombres).

Q2 Ici on cherche à estimer $v=\sigma^2$. Soit $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

On pose alors $\forall v \in \mathbb{R}_+^*$, $\psi(v) = \ln\left(\prod_{k=1}^{n}\frac{1}{\sqrt{2\pi}\sqrt{v}}\,e^{-\frac{(x_k-m)^2}{2v}}\right)$.

Alors $\forall v \in \mathbb{R}_+^*$, $\psi(v) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln v - \frac{1}{2v}\sum_{k=1}^{n}(x_k-m)^2$. ψ est dérivable sur $\mathbb{R}_+^*$

et $\forall v \in \mathbb{R}_+^*$, $\psi'(v) = -\frac{n}{2v} + \frac{1}{2v^2}\sum_{k=1}^{n}(x_k-m)^2 = \frac{n}{2v^2}\left[\frac{1}{n}\sum_{k=1}^{n}(x_k-m)^2 - v\right]$.

$\forall v \in \mathbb{R}_+^*$, $\psi'(v) \geqslant 0 \Leftrightarrow \frac{1}{n}\sum_{k=1}^{n}(x_k-m)^2 \geqslant v$. Alors ψ admet un maximum en $\frac{1}{n}\sum_{k=1}^{n}(x_k-m)^2$.

Si l'on soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon i.i.d de loi normale de paramètres m et σ^2. Posons $U_n = \frac{1}{n}\sum_{k=1}^{n}(X_k-m)^2$. Montrons que U_n est un estimateur sans biais et convergent de σ^2.

Posons $\forall k \in [\![1,n]\!]$, $Y_k = (X_k-m)^2$. $\quad U_n = \dfrac{Y_1+Y_2+\cdots+Y_n}{n}$.

. les variables aléatoires Y_k → sont indépendantes
→ ont même loi
→ possède une espérance σ^2 et une variance

(X_k possède un moment d'ordre r, pour tout $r \in \mathbb{N}$)

Ainsi 1) $E(U_n)$ existe et vaut $\frac{1}{n}\times n\times\sigma^2 = \sigma^2$

2) la suite de terme général U_n converge en probabilité vers la variable certaine σ^2 d'après la loi faible des grands nombres.

$U_n = \frac{1}{n}\sum_{k=1}^{n}(X_k-m)^2$ est un estimateur sans biais et convergent de σ^2.

EX 9.2

Exercice θ est réel non nul. $\forall t \in]-\infty, 1[$, $f(t) = 0$ et $\forall t \in [1, +\infty[$, $f(t) = \dfrac{1}{\theta\, t^{1+1/\theta}}$.

Q1. Montrer que f est une densité de probabilité si et seulement si θ est strictement positif. Ce que nous supposerons dans la suite.

Q2. X est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ de densité f. Etudier $T = \dfrac{1}{\theta} \ln X$.

Q3. Utiliser la méthode du maximum de vraisemblance pour construire un estimateur de θ

(prendre $(x_1, x_2, \ldots, x_n) \in ([1, +\infty[)^n$ et étudier $\theta \to \ln(\prod_{k=1}^{n} f(x_k))$).

L'estimateur obtenu est-il sans biais ? convergent ?

Q1 → si $\theta < 0$, f est strictement négative sur $[1,+\infty[$!! supposons $\theta > 0$

• f est positive sur $\mathbb{R}$ et au moins continue sur $\mathbb{R}-\{1\}$.

• $\int_{-\infty}^{0} f(t)dt$ existe et vaut 0. f est continue sur $[1,+\infty[$ et $\forall A \in [1,+\infty[$,

$$\int_1^A f(t)dt = \frac{1}{\theta}\int_1^A t^{-1-\frac{1}{\theta}}dt = \frac{1}{\theta}\left[\frac{t^{-\frac{1}{\theta}}}{-\frac{1}{\theta}}\right]_1^A = 1 - \frac{1}{A^{1/\theta}} \text{ et } \lim_{A\to+\infty}\left(1 - \frac{1}{A^{1/\theta}}\right) = 1.$$

Alors $\int_1^{+\infty} f(t)dt$ existe et vaut 1. Donc $\int_{-\infty}^{+\infty} f(t)dt$ existe et vaut 1.

Ainsi f est une densité de probabilité si et seulement si θ est strictement positif.

Q2 X prend ses valeurs dans $[1,+\infty[$ donc T prend ses valeurs dans $\mathbb{R}_+$

$\forall x \in]-\infty, 0[$, $F_T(x) = 0$. Soit $x \in [0,+\infty[$.

$$F_T(x) = P(T \leq x) = P\left(\frac{1}{\theta}\ln X \leq x\right) = P(X \leq e^{\theta x}) = \int_1^{e^{\theta x}} \frac{1}{\theta}\frac{1}{t^{1+1/\theta}}dt = \left[-\frac{1}{t^{1/\theta}}\right]_1^{e^{\theta x}}$$

$F_T(x) = 1 - \dfrac{1}{(e^{\theta x})^{1/\theta}} = 1 - e^{-x}$. $\forall x \in \mathbb{R}$,

$\forall x \in \mathbb{R}$, $F_T(x) = \begin{cases} 0 & \text{si } x \in]-\infty, 0[\\ 1 - e^{-x} & \text{si } x \in [0,+\infty[\end{cases}$. T suit la loi exponentielle de paramètre 1.

Q3 Soit $(x_1, \ldots, x_n) \in ([1,+\infty[)^n$. Posons $\forall \theta \in]0,+\infty[$, $\varphi(\theta) = \ln \prod_{k=1}^{n} f(x_k)$.

$\forall \theta \in \mathbb{R}_+^*$, $\varphi(\theta) = \sum_{k=1}^{n} \ln\left(\dfrac{1}{\theta\, x_k^{1+\frac{1}{\theta}}}\right) = -n\ln\theta - \left(1 + \frac{1}{\theta}\right)\sum_{k=1}^{n}\ln x_k$. φ est dérivable sur $]0,+\infty[$.

$\forall \theta \in \mathbb{R}_+^*$, $\varphi'(\theta) = -\dfrac{n}{\theta} + \dfrac{1}{\theta^2}\sum_{k=1}^{n}\ln x_k = \dfrac{n}{\theta^2}\left[\dfrac{1}{n}\sum_{k=1}^{n}\ln x_k - \theta\right]$. Alors φ est maximum pour

$\theta = \dfrac{1}{n}\sum_{k=1}^{n}\ln x_k$. Soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon iid de la loi de X.

Nous retiendrons comme estimateur de θ : $\dfrac{\ln X_1 + \ln X_2 + \cdots + \ln X_n}{n}$. La loi faible des grands nombres montrent qu'il est sans biais et convergent (... $E\left(\dfrac{\ln X_k}{\theta}\right) = 1$ d'après Q2 donc $E(\ln X_k) = \theta$ ...)

Ex 23

Exercice Et si on estimait un paramètre vectoriel la méthode du maximum de vraisemblance ?
X est une variable aléatoire réelle suivant une loi normale de paramètres m et σ. On se propose donc d'estimer simultanément m et σ^2donc le paramètre vectoriel (m, σ^2).
Dès lors on se donne $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ et on considère :

$$L(x_1, x_2, \ldots, x_n, m, \sigma^2) = \left(\prod_{k=1}^{n}\left(\frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{(x_k-m)^2}{2\sigma^2}}\right)\right)$$

Le but est de trouver (m, σ^2) pour que cette quantité soit maximale.
On pose alors :

$$\forall (m, v) \in \mathbb{R} \times \mathbb{R}^{+*},\ \varphi(m, v) = \ln\left(L(x_1, x_2, \ldots, x_n, m, (\sqrt{v})^2)\right)$$

Q1. Montrer que $\mathbb{R} \times \mathbb{R}^{+*}$ est un ouvert de $\mathbb{R}^2$.
Q2. Montrer que φ est de classe $\mathcal{C}^1$ sur $\mathbb{R} \times \mathbb{R}^{+*}$.
Q3. Montrer que φ admet un point critique et un seul (m^*, v^*) avec :

$$m^* = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad \text{et} \quad v^* = \frac{1}{n}\sum_{k=1}^{n}(x_k - m^*)^2.$$

Q4. Montrer que $\varphi(m^*, v^*) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln v^* - \frac{n}{2}$.
Montrer que si $(m, v) \in \mathbb{R} \times \mathbb{R}^{+*}$:

$$\varphi(m, v) - \varphi(m^*, v^*) = \frac{n}{2}\left[\ln\frac{v^*}{v} + 1\right] - \sum_{k=1}^{n}\frac{(x_k - m)^2}{2v} \leqslant 0.$$

$(\ln t \leqslant t - 1\ldots)$
Proposer un estimateur de m et un estimateur de σ^2.

Q1 $\mathbb{R}$ et $\mathbb{R}_+^* =]0, +\infty[$ sont deux ouverts de $\mathbb{R}$ donc $\mathbb{R} \times \mathbb{R}_+^*$ est un ouvert de $\mathbb{R}^2$.

Q2 $\forall (m,v) \in \mathbb{R} \times \mathbb{R}_+^*$, $\varphi(m,v) = \ln\left(\prod_{k=1}^{n} \frac{1}{\sqrt{2\pi v}}\, e^{-\frac{(x_k-m)^2}{2v}}\right)$

$\forall (m,v) \in \mathbb{R} \times \mathbb{R}_+^*$, $\varphi(m,v) = \sum_{k=1}^{n}\left(-\frac{1}{2}\ln(2\pi v) - \frac{(x_k-m)^2}{2v}\right) = \sum_{k=1}^{n}\left(-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln v - \frac{(x_k-m)^2}{2v}\right)$

Soit $k \in [\![1, n]\!]$; $(m,v) \longrightarrow -\frac{1}{2}\ln(2\pi) - \frac{(x_k-m)^2}{2v}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$ car elle est rationnelle.

$(m,v) \to v$ est de classe $\mathcal{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$ et y prend des valeurs strictement positives ; comme $\ln$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$, $(m,v) \to -\frac{1}{2}\ln v$ est de classe $\mathcal{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$

donc $(m,v) \longrightarrow -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln v - \frac{(x_k-m)^2}{2v}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$ et ceci pour tout $k \in [\![1, n]\!]$.

Donc φ est de classe $\mathcal{C}^1$ sur $\mathbb{R} \times \mathbb{R}_+^*$.

Q3 Soit $(m,v)\in\mathbb{R}\times\mathbb{R}_+^*$.

$$\frac{\partial\varphi}{\partial m}(m,v)=\sum_{k=1}^{n}\left(-2\,\frac{(-1)(x_k-m)}{2v}\right)=\sum_{k=1}^{n}\frac{(x_k-m)}{v}=\frac{1}{v}\left(\sum_{k=1}^{n}x_k-nm\right)$$

$$\frac{\partial\varphi}{\partial m}(m,v)=0\iff m=\frac{1}{n}\sum_{k=1}^{n}x_k\,.$$

$$\frac{\partial\varphi}{\partial v}(m,v)=\sum_{k=1}^{n}\left(-\frac12\,\frac1v-\frac{(x_k-m)^2}{2}\left(-\frac{1}{v^2}\right)\right)=\frac{1}{2v^2}\left[\sum_{k=1}^{n}\left((x_k-m)^2-v\right)\right]=\frac{1}{2v^2}\left(\sum_{k=1}^{n}(x_k-m)^2-nv\right)$$

$$\frac{\partial\varphi}{\partial v}(m,v)=0\iff\sum_{k=1}^{n}(x_k-m)^2=nv\iff v=\frac1n\sum_{k=1}^{n}(x_k-m)^2.$$

φ admet un point critique et un seul (m^*,v^*) où :

$$m^*=\frac{x_1+x_2+\dots+x_n}{n}\qquad\text{et}\qquad v^*=\frac1n\sum_{k=1}^{n}(x_k-m^*)^2$$

Q4
$$\varphi(m^*,v^*)=\sum_{k=1}^{n}\left(-\frac12\ln(2\pi)-\frac12\ln v^*-\frac{(x_k-m^*)^2}{2v^*}\right)$$

$$\varphi(m^*,v^*)=-\frac n2\ln(2\pi)-\frac n2\ln(v^*)-\frac{1}{2v^*}\sum_{k=1}^{n}(x_k-m^*)^2=-\frac n2\ln(2\pi)-\frac n2\ln(v^*)-\frac{1}{2v^*}\times nv^*$$

donc $\varphi(m^*,v^*)=-\frac n2\ln(2\pi)-\frac n2\ln(v^*)-\frac n2$.

Soit $(m,v)\in\mathbb{R}\times\mathbb{R}_+^*$. $\underbrace{\varphi(m,v)-\varphi(m^*,v^*)}_{\Delta(m,v)}=\sum_{k=1}^{n}\left(-\frac12\ln(2\pi)-\frac12\ln v-\frac{(x_k-m)^2}{2v}\right)+\frac n2\ln 2\pi+\frac n2\ln v^*+\frac n2$.

$$\Delta(m,v)=-\frac n2\ln(2\pi)-\frac n2\ln v-\frac{1}{2v}\sum_{k=1}^{n}(x_k-m)^2+\frac n2\ln(2\pi)+\frac n2\ln(v^*)+\frac n2=\frac n2\left[\ln\left(\frac{v^*}{v}\right)+1\right]-\frac{1}{2v}\sum_{k=1}^{n}(x_k-m)^2$$

$\Delta(m,v)\underset{\ln t+1\le t}{\le}\frac n2\,\frac{v^*}{v}-\frac{1}{2v}\sum_{k=1}^{n}(x_k-m)^2=\frac{n}{2v}\left(v^*-\frac1n\sum_{k=1}^{n}(x_k-m)^2\right)$. Montrons dès lors que $\Delta(m,v)\le 0$,

c'est à dire que $v^*\le\frac1n\sum_{k=1}^{n}(x_k-m)^2$ (1)

$(1)\iff\frac1n\sum_{k=1}^{n}(x_k-m^*)^2\le\frac1n\sum_{k=1}^{n}(x_k-m)^2\iff\cancel{\sum_{k=1}^{n}x_k^2}-m^*\underbrace{2\sum_{k=1}^{n}x_k}_{2nm^*}+nm^{*2}\le\cancel{\sum_{k=1}^{n}x_k^2}-2m\sum_{k=1}^{n}x_k+nm^2$

$(1)\iff-2nm^{*2}+nm^{*2}\le-2mnm^*+nm^2$

$(1)\iff 0\le nm^{*2}-2nmm^*+nm^2=n(m^*-m)^2$; ceci achève de prouver que $\Delta(m,v)\le 0$

donc $\forall(m,v)\in\mathbb{R}\times\mathbb{R}_+^*$, $\varphi(m,v)\le\varphi(m^*,v^*)$. φ admet un maximum (strict !) en (m^*,v^*)

Nous retiendrons comme estimateur de m : $\overline{X}_n=\frac{X_1+X_2+\dots+X_n}{n}$ et comme estimateur de v :
$\frac1n\sum_{k=1}^{n}(X_k-\overline{X})^2$.

Ex 25

Exercice Construire un estimateur du paramètre p ($p \in]0,1[$) d'une variable aléatoire suivant une loi binômiale de paramètres r et p en utilisant la méthode du maximum de vraisemblance.

L'estimateur obtenu est-il sans biais ? convergent ?

$n \in \mathbb{N}^*$. $(x_1, x_2, \dots, x_n) \in [\![0, r]\!]^n$.

$(X_1, X_2, \dots, X_n)$ est un n-échantillon i.i.d de la loi binômiale de paramètres r et p.

$P(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\}) = \prod_{k=1}^{n} P(X_k = x_k) = \prod_{k=1}^{n} \left[\binom{r}{x_k} p^{x_k} (1-p)^{r-x_k} \right]$.

$\ln\left(\prod_{k=1}^{n} P(X = x_k)\right) = \sum_{k=1}^{n} \left[\ln \binom{r}{x_k} + x_k \ln p + (r - x_k) \ln(1-p) \right]$.

Etudions $\varphi : p \mapsto \sum_{k=1}^{n} \left[\ln \binom{r}{x_k} + x_k \ln p + (r-x_k) \ln(1-p) \right]$ sur $]0,1[$.

φ est dérivable sur $]0,1[$ et $\forall p \in]0,1[$, $\varphi'(p) = \sum_{k=1}^{n} \left[\frac{x_k}{p} - \frac{r - x_k}{1-p} \right]$.

$\forall p \in]0,1[$, $\varphi'(p) = \frac{1}{p(1-p)} \left[(1-p) \sum_{k=1}^{n} x_k - p \sum_{k=1}^{n} (r - x_k) \right] = \frac{1}{p(1-p)} \left[(1-p) S - p(nr - S) \right]$

où $S = x_1 + x_2 + \dots + x_n$.

$\forall p \in]0,1[$, $\varphi'(p) \geqslant 0 \iff (1-p) S - p(nr - S) \geqslant 0 \iff S \geqslant p(S + nr - S) = pnr$

$\forall p \in]0,1[$, $\varphi'(p) \geqslant 0 \iff p \leqslant \frac{S}{nr} = \frac{1}{rn} \sum_{k=1}^{n} x_k$.

φ possède en $\frac{1}{rn} \sum_{k=1}^{n} x_k$ un maximum sur $]0,1[$.

Posons alors $U_n = \frac{1}{rn} \sum_{k=1}^{n} X_k$. U_n possède une espérance qui vaut $\frac{1}{rn} \times n \times (rp) = p$ et une variance qui vaut $\frac{1}{(rn)^2} n \times r \times p \times (1-p) = \frac{p(1-p)}{rn}$.

U_n est un estimateur sans biais de p.

$\forall \varepsilon \in \mathbb{R}_+^*$, $0 \leqslant P(|U_n - p| \geqslant \varepsilon) = P(|U_n - E(U_n)| \geqslant \varepsilon) \leqslant \frac{V(U_n)}{\varepsilon^2} = \frac{p(1-p)}{rn\varepsilon^2}$ et $\lim_{n \to +\infty} \frac{p(1-p)}{rn\varepsilon^2} = 0$

par encadrement $\forall \varepsilon \in \mathbb{R}_+^*$, $\lim_{n \to +\infty} P(|U_n - p| \geqslant \varepsilon) = 0$.

Finalement : $U_n = \frac{1}{nr} \sum_{k=1}^{n} X_k$ est un estimateur sans biais et convergent de p.