D.M. n°9 – Correction – ESSEC II 2009

Rappels importants :

_ Soit X une VAR. Si X admet une espérance et si $(X - E(X))^2$ admet une espérance, on dit que X admet une variance et on appelle variance le réel $\mathbf{V(X) = E\big((X - E(X))^2\big)}$

_ Théorème de transfert (version VAR discrète)

Soit X une VAR discrète, avec $X(\Omega) = \{x_i,\ i \in I\}$ et g une fonction de $X(\Omega)$ dans $\mathbb{R}$.

La VAR Y = g(X) admet une espérance si et seulement si $X(\Omega)$ est fini ou si la série $\sum_i g(x_i)P(X = x_i)$ converge absolument.

Dans ce cas, $E(Y) = \mathbf{E(g(X)) = \sum_{i \in I} g(x_i)P(X = x_i)}$

_ Théorème de transfert (version VAR à densité)

Soit X une V.A.R. qui admet une densité f_X nulle en dehors d'un intervalle]a;b[(avec $-\infty \le a < b \le +\infty$) et soit g une fonction continue sur]a;b[sauf en un nombre fini de points. La variable aléatoire Y = g(X) admet une espérance si et seulement si l'intégrale $\int_a^b g(x)f_X(x)dx$ est absolument convergente.

Dans ce cas, $E(Y) = \mathbf{E(g(X)) = \int_a^b g(x)f_X(x)dx}$.

_ Inégalité de Cauchy-Schwarz (méthode à connaitre)

Soit $(x_1, \dots, x_n)$ et $(y_1, \dots, y_n)$ deux familles de n réels. $\forall\, t \in \mathbb{R}$, on pose $P(t) = \sum_{k=1}^{n} (t.x_k + y_k)^2$.

En **développant P**, on voit facilement qu'il s'agit d'un polynôme du second degré.

Comme **P(t) est positif quel que soit le réel t**, on a $\mathbf{\Delta \le 0}$ (si $\Delta > 0$, P change de signe).

On en déduit que : $\left|\sum_{k=1}^{n} x_k y_k\right| \le \sqrt{\sum_{k=1}^{n} x_k^2}\sqrt{\sum_{k=1}^{n} y_k^2}$

I. Information de Fisher

I.A.1 $\forall\, \theta \in\,]0\,;1[$, $\ln(p(\theta,1)) = \ln(\theta)$ donc $\partial_1\big(\ln(p(\theta,1))\big) = \dfrac{1}{\theta}$

et $\ln(p(\theta,0)) = \ln(1 - \theta)$ donc $\partial_1\big(\ln(p(\theta,0))\big) = -\dfrac{1}{1-\theta}$

Donc $I_X(\theta) = \Big(\partial_1\big(\ln\big(p(\theta, 1)\big)\big)\Big)^2 p(\theta, 1) + \Big(\partial_1\big(\ln\big(p(\theta, 0)\big)\big)\Big)^2 p(\theta, 0)$

$= \dfrac{1}{\theta^2}\theta + \dfrac{1}{(1-\theta)^2}(1-\theta) = \dfrac{1}{\theta} + \dfrac{1}{1-\theta} = \dfrac{1-\theta+\theta}{\theta(1-\theta)} = \dfrac{1}{\theta(1-\theta)}$.

I.A.2 a) $\forall\, k \in \{0, \dots, N\}$, $p(\theta, k) = \binom{N}{k}\theta^k(1-\theta)^{N-k}$ donc $\ln(p(\theta, k)) = \ln\left(\binom{N}{k}\right) + k.\ln(\theta) + (N - k)\ln(1 - \theta)$

Donc $\partial_1\big(\ln(p(\theta, k))\big) = \dfrac{k}{\theta} + (N-k)\dfrac{-1}{1-\theta} = \dfrac{k(1-\theta) - \theta(N-k)}{\theta(1-\theta)} = \dfrac{k - N\theta}{\theta(1-\theta)}$

Donc $I_x(\theta) = \sum_{k=0}^{N}\left(\dfrac{k-N\theta}{\theta(1-\theta)}\right)^2\binom{N}{k}\theta^k(1-\theta)^{N-k} = \dfrac{1}{(\theta(1-\theta))^2}\sum_{k=0}^{N}(k-N\theta)^2\binom{N}{k}\theta^k(1-\theta)^{N-k}$

b) Or par définition, $V(X) = E\big((X - E(X))^2\big) = E((X - N\theta)^2)$

D'après le théorème de transfert, $V(X) = \sum_{k=0}^{N} (k - N\theta)^2 P(X = k) = \sum_{k=0}^{N} (k - N\theta)^2 \binom{N}{k} \theta^k (1 - \theta)^{N-k}$.

Donc $I_X(\theta) = \dfrac{V(X)}{(\theta(1-\theta))^2} = \dfrac{N\theta(1-\theta)}{(\theta(1-\theta))^2} = \dfrac{N}{\theta(1-\theta)}$.

I.A.3 $\forall k \in \mathbb{N}$, $p(\theta, k) = \dfrac{\theta^k e^{-\theta}}{k!}$ donc $\ln(p(\theta, k)) = \ln(\theta^k) + \ln(e^{-\theta}) - \ln(k\,!) = k.\ln(\theta) - \theta - \ln(k\,!)$.

Donc $\partial_1\big(\ln(p(\theta, k))\big) = \dfrac{k}{\theta} - 1 = \dfrac{k-\theta}{\theta}$

Donc $\sum_{k \in \mathbb{N}} \big(\partial_1\big(\ln\big(p(\theta, k)\big)\big)\big)^2 p\,(\theta, k) = \sum_{k \in \mathbb{N}} \left(\dfrac{k-\theta}{\theta}\right)^2 \dfrac{\theta^k e^{-\theta}}{k!} = \dfrac{1}{\theta^2} \sum_{k \in \mathbb{N}} (k-\theta)^2 P(X = k)$

Or, comme $E(X) = \theta$, on reconnait comme dans le cas précédent, la série définissant la variance de X.
Comme X admet une variance, les séries convergent (donc $I_X(\theta)$ existe) et

$I_X(\theta) = \sum_{k=0}^{+\infty} \big(\partial_1\big(\ln\big(p(\theta, k)\big)\big)\big)^2 p\,(\theta, k) = \dfrac{1}{\theta^2} V(X) = \dfrac{\theta}{\theta^2} = \dfrac{1}{\theta}$.

b. $\sum_{k \in \mathbb{N}} \big(\partial_1\big(\ln\big(p(\theta, k)\big)\big)\big)^2 P(X = k) = \sum_{k \in \mathbb{N}} \big(\partial_1\big(\ln\big(p(\theta, k)\big)\big)\big)^2 p(\theta, k)$

Donc la série converge (absolument car tout est positif).
Donc d'après le théorème de transfert, la variable aléatoire $Y = \big(\partial_1\big(\ln\big(p(\theta, X)\big)\big)\big)^2$ admet une espérance et

$E(Y) = E\Big(\big(\partial_1\big(\ln\big(p(\theta, X)\big)\big)\big)^2\Big) = \sum_{k=0}^{+\infty} \big(\partial_1\big(\ln\big(p(\theta, k)\big)\big)\big)^2 p\,(\theta, k) = I_X(\theta)$.

B. Cas d'une variable gaussienne

I.B.1 $\forall \theta \in \mathbb{R}$, $\forall x \in \mathbb{R}$, $f(\theta, x) = \dfrac{1}{\sqrt{2\pi}} \exp\left(-\dfrac{(x-\theta)^2}{2}\right)$ (qui est bien strictement positif sur $\mathbb{R}$)

$\ln(f(\theta, x)) = \ln\left(\dfrac{1}{\sqrt{2\pi}}\right) - \dfrac{(x-\theta)^2}{2}$ donc $\partial_1\big(\ln(f(\theta, x))\big) = -(-1)(x-\theta) = x - \theta$.

Donc, sous réserve d'existence, $I_X(\theta) = \int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)dx$.

I.B.2 On a d'après le théorème de transfert et sous réserve d'existence,
$\int_{-\infty}^{+\infty} (x-\theta)^2 f(\theta, x)dx = E\big((X-\theta)^2\big) = E\,\big((X - E(X))^2\big) = V(X)$.

Or X admet une variance qui vaut 1. Donc toutes ces intégrales convergent et $I_X(\theta) = V(X) = 1$.
I.B.3 D'après le théorème de transfert, on a également
$I_X(\theta) = \int_{-\infty}^{+\infty} \big(\partial_1\big(\ln\big(f(\theta, x)\big)\big)\big)^2 f(\theta, x)dx = E\,\Big(\big(\partial_1\big(\ln\big(f(\theta, X)\big)\big)\big)^2\Big)$

II. Inégalité de Cramer-Rao

1. X est une variable aléatoire à valeurs dans $\{0, \dots, N\}$ donc $\sum_{k=0}^{N} P(X = k) = 1$

Donc $\forall \theta \in I$, $\sum_{k=0}^{N} p(\theta, k) = 1$. Donc en dérivant, par linéarité de la dérivée : $\sum_{k=0}^{N} \partial_1(p(\theta, k)) = 0$

2. $\forall k \in \{0, .., N\}$, $\forall \theta \in I$, $\partial_1(\ln(p(\theta, k))) = \dfrac{\partial_1(p(\theta,k))}{p(\theta, k)}$ (*). $(\ln(u) \longrightarrow u'/u)$

D'après le théorème de transfert,

$E\left(\partial_1 \ln(p(\theta, X))\right) = \sum_{k=0}^{N} \partial_1(\ln(p(\theta, k)))P(X = k) = \sum_{k=0}^{N} \frac{\partial_1(p(\theta,k))}{p(\theta, k)} p(\theta, k) = \sum_{k=0}^{N} \partial_1(p(\theta, k)) = 0$

3. $\forall \theta \in I$, $\sum_{k=0}^{N} \partial_1(\ln(p(\theta, k)))\, p(\theta, k) = 0$

En dérivant cette somme par rapport à θ, on obtient :

$\sum_{k=0}^{N} \left(\partial_1(\partial_1(\ln(p(\theta, k))))p(\theta, k) + \partial_1(\ln(p(\theta, k)))\, \partial_1(p(\theta, k))\right) = 0$

Donc $\sum_{k=0}^{N} \partial^2_{1,1}(\ln(p(\theta, k)))p(\theta, k) = - \sum_{k=0}^{N} \left(\partial_1(\ln(p(\theta, k)))\right)p(\theta, k)\, \partial_1(\ln(p(\theta, k)))$ (d'après (*))

$\sum_{k=0}^{N} \partial^2_{1,1}(\ln(p(\theta, k)))P(X = k) = - \sum_{k=0}^{N} \left(\partial_1(\ln(p(\theta, k)))\right)^2 P(X = k)$

Donc d'après le théorème de transfert, $E\left(\partial^2_{1,1} \ln(p(\theta,X))\right) = - E\left(\left(\partial_1 \ln (p (\theta,X))\right)^2\right)$

4. $\forall \theta \in I$, $E(f(X)) = g(\theta)$ donc $\sum_{k=0}^{N} f(k)P(X = k) = g(\theta)$ $\quad \sum_{k=0}^{N} f(k)p(\theta, k) = g(\theta)$

$\forall k \in \{0, \dots, N\}$, $\theta \longmapsto p(\theta, k)$ est dérivable sur I et g aussi, donc en dérivant cette égalité :

$\forall \theta \in I$, $\sum_{k=0}^{N} f(k)\partial_1(p(\theta, k)) = g'(\theta)$ donc $\sum_{k=0}^{N} f(k)p(\theta, k)\, \partial_1(\ln(p(\theta, k))) = g'(\theta)$

Toujours d'après le théorème de transfert,

$E\left((f (X) - g (\theta))\partial_1 \ln(p (\theta ,X))\right) = \sum_{k=0}^{N} \left(f(k) - g(\theta)\right)\partial_1(\ln(p(\theta, k)))P(X = k)$

$= \sum_{k=0}^{N} f(k).\, \partial_1(\ln(p(\theta, k)))p(\theta,k) - g(\theta)\sum_{k=0}^{N} \partial_1(\ln(p(\theta, k)))p(\theta,k)$

$= g'(\theta) - g(\theta) \times 0 = g'(\theta)$.

5. *(On reconnaitra la méthode de l'inégalité de Cauchy-Schwarz)*

Posons $Y = f(X) - g(\theta)$ et $Z = \partial_1 \ln(p(\theta,X))$

$\forall t \in \mathbb{R}$, $L(t) = E\left((Y + tZ)^2\right) = E(Y^2 + 2tYZ + t^2Z^2) = E(Z^2)\, t^2 + 2E(YZ).t + E(Y^2)$ par linéarité de l'espérance.

b) $\Delta = 4E(YZ)^2 - 4E(Y^2)E(Z^2) = 4(E(YZ)^2 - E(Y^2)E(Z^2))$

$\forall t \in \mathbb{R}$, la variable aléatoire $(Y + tZ)^2$ est à valeurs positives, donc $L(t) = E\left((Y + tZ)^2\right) \geq 0$.

L est un polynôme du second degré qui ne change pas de signe sur $\mathbb{R}$. Donc $\Delta \leq 0$.

c) Donc $E(YZ)^2 \leq E(Y^2)E(Z^2)$

avec $E(YZ) = E\left((f (X) - g (\theta))\partial_1 \ln(p (\theta ,X))\right) = g'(\theta)$

$E(Y^2) = E\left((f(X) - g(\theta))^2\right) = E\left((f(X) - E(f(X)))^2\right) = V(f(X))$

$E(Z^2) = E\left(\partial_1 \ln(p(\theta,X))^2\right) = I_X(\theta)$ d'après la formule de transfert.

Donc $g'(\theta)^2 \leq V(f(X)).\, I_X(\theta)$

$I_X(\theta)$ est l'espérance d'une variable positive donc $I_X(\theta) \geq 0$ et on suppose dans l'énoncé qu'elle est non nulle, donc $I_X(\theta) > 0$.

Donc $\frac{g'(\theta)^2}{I_X(\theta)} \leq V(f(X))$.