Elements de correction MIP

Table des matières

1 QCM

1.1 Question 1

Énoncé : On considère deux suites de variables aléatoires, $\{X_n : n \in \mathbb{N}\}$ et $\{Y_n : n \in \mathbb{N}\}$, définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. Laquelle de ces trois assertions est vraie ?

(a) Si $X_n \xrightarrow[n\to\infty]{\text{loi}} X$ et $Y_n \xrightarrow{\text{p.s.}} Y$, alors $X_n Y_n \xrightarrow[n\to\infty]{\text{loi}} XY$.

(b) Si $X_n \xrightarrow[n\to\infty]{\text{loi}} X$ et $Y_n \xrightarrow{\mathbb{P}} a$ où a est un nombre réel, alors $X_n Y_n \xrightarrow[n\to\infty]{\text{loi}} aX$.

(c) Si $X_n \xrightarrow[n\to\infty]{\text{loi}} X$, $Y_n \xrightarrow[n\to\infty]{\text{loi}} Y$ et X est indépendant de Y, alors $X_n Y_n \xrightarrow[n\to\infty]{\text{loi}} XY$.

Réponse

Assertion (a) : *Fausse*

Explication : La convergence en loi de X_n vers X et la convergence presque sûre de Y_n vers Y ne suffisent pas, en général, pour assurer que le produit $X_n Y_n$ converge en loi vers XY. Pour donner un contre-exemple simple, considérons le cas où $X_n = X \sim \mathcal{N}(0,1)$ et $Y_n = Y \sim \mathcal{N}(0,1)$ pour tout n, avec X et Y indépendants. Alors $X_n \xrightarrow[n\to\infty]{\text{loi}} Y$ et $Y_n \xrightarrow[n\to\infty]{\text{p.s.}} Y$ mais $X_n Y_n = XY$ ne converge pas en loi vers Y^2. En effet, $P(XY > 0) = 1/2$ alors que $P(Y^2 > 0) = 1$. Ce contre-exemple illustre bien que le produit ne se comporte pas nécessairement comme on pourrait le souhaiter.

Assertion (b) : *Vraie*

Explication : Par le théorème de Slutsky, si $X_n \xrightarrow[n\to\infty]{\text{loi}} X$ et $Y_n \xrightarrow[n\to\infty]{\mathbb{P}} a$, alors $X_n Y_n \xrightarrow[n\to\infty]{\text{loi}} aX$.

Assertion (c) : *Fausse*

Explication : La convergence en loi de X_n vers X et de Y_n vers Y n'implique pas la convergence en loi de $X_n Y_n$ vers XY, même si X et Y sont indépendants.

Par exemple, prenons deux variables indépendants $X \sim \mathcal{N}(0,1)$ et $Y \sim \mathcal{N}(0,1)$ et définissons $X_n = Y_n = X$ pour tout n. Alors $X_n \xrightarrow[n\to\infty]{\text{loi}} X$ et $Y_n \xrightarrow[n\to\infty]{\text{loi}} Y$ mais $X_n Y_n = X^2 \xrightarrow[n\to\infty]{\text{loi}} X^2$, qui n'a pas la même loi que XY.

$\boxed{\textbf{Conclusion :}\ \textit{La seule assertion vraie est la } \textbf{\textit{(b)}}.}$

1.2 Question 2

Énoncé :

Soit $\widehat{\theta}_n : \Omega \to \mathbb{R}^2$ un estimateur fortement consistant de $\theta^* \in \mathbb{R}^2$. Laquelle de ces assertions est vraie ?

(a) Le risque quadratique de $\widehat{\theta}_n$ tend nécessairement vers zéro lorsque $n \to \infty$.

(b) Le biais de $\widehat{\theta}_n$ tend nécessairement vers zéro lorsque $n \to \infty$.

(c) $\|\widehat{\theta}_n\|$ est nécessairement un estimateur fortement consistant de $\|\theta^*\|$.

Réponse

Assertion (a) : *Fausse*

Explication : La consistance forte n'implique pas nécessairement que le risque quadratique tende vers zéro. Considérons un modèle de Bernoulli avec paramètre θ^*. Définissons un estimateur $\widehat{\theta}_n = \bar{X}_n + \frac{Y}{n}$ où $\bar{X}_n$ est la moyenne empirique des variables iid de loi de Bernoulli de paramètre θ^* et Y est une variable aléatoire indépendante du reste, de loi de Cauchy. L'estimateur $\widehat{\theta}_n$ est fortement consistant (car $\bar{X}_n \to \theta^*$ p.s. et $\frac{Y}{n} \to 0$ p.s.), mais son risque quadratique étant égal à $+\infty$ ne tend pas vers zéro.

Assertion (b) : *Fausse*

Explication : Reprenons le même exemple. L'espérance de $\widehat{\theta}_n$ peut ne pas exister (loi de Cauchy), ce qui rend le biais mal défini ou infini. Même s'il existait, on pourrait avoir un biais constant. Ainsi, la consistance forte n'assure pas la décroissance du biais.

Assertion (c) : *Vraie*

Explication : La norme est une fonction continue. Puisque $\widehat{\theta}_n \xrightarrow[n\to\infty]{\text{p.s.}} \theta^*$ et que $x \mapsto \|x\|$ est continue, on obtient $\|\widehat{\theta}_n\| \xrightarrow[n\to\infty]{\text{p.s.}} \|\theta^*\|$. Ainsi, $\|\widehat{\theta}_n\|$ est un estimateur fortement consistant de $\|\theta^*\|$.

Conclusion : *La seule assertion vraie est la **(c)**.*

1.3 Question 3

Énoncé : Soit F_n la fonction de répartition empirique de n variables iid $X_1, \ldots, X_n$. Soit F la fonction de répartition de X_1. Deux assertions sont fausses, lesquelles ? Quel que soit $x \in \mathbb{R}$:

(a) $nF_n(x)$ suit la loi binomiale de paramètres $(n, F(x)(1 - F(x)))$.

(b) $\sqrt{n}\,(F_n(x) - F(x)) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, F(x))$.

(c) Le risque quadratique de $F_n(x)$, considéré comme estimateur du paramètre $\theta = F(x)$, est égal à $\dfrac{F(x)(1-F(x))}{n}$.

(d) $F_n(x)$ est un estimateur faiblement consistant de $F(x)$.

Réponse

Assertion (a) : *Fausse*

Explication : La variable $nF_n(x) = \sum_{i=1}^n \mathbf{1}_{\{X_i \le x\}}$ suit une loi binomiale de paramètres $(n, F(x))$, non $(n, F(x)(1 - F(x)))$.

Assertion (b) : *Fausse*

Explication : Par le théorème central limite, on a :

$$\sqrt{n}\,(F_n(x) - F(x)) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}\big(0, F(x)(1 - F(x))\big).$$

L'assertion (b) indique une variance asymptotique de $F(x)$, ce qui est incorrect. La variance asymptotique correcte est $F(x)(1 - F(x))$.

Assertion (c) : *Vraie*

Explication : Le risque quadratique de $F_n(x)$ est :

$$\mathbb{E}\left[(F_n(x) - F(x))^2\right] = \mathrm{Var}\,(F_n(x)) = \frac{F(x)(1-F(x))}{n}.$$

Ainsi, l'assertion (c) est vraie.

Assertion (d) : *Vraie*
Explication : Par la loi faible des grands nombres, $F_n(x)$ converge en probabilité vers $F(x)$. Donc, $F_n(x)$ est un estimateur faiblement consistant de $F(x)$.

Conclusion : *Les deux assertions fausses sont les assertions **(a)** et **(b)**.*

1.4 Question 4

Énoncé : Pour tout $n \in \mathbb{N}$, soit X_n une variable aléatoire de loi Gaussienne $\mathcal{N}\left(\frac{2}{\sqrt{n}}, \frac{1}{n}\right)$. Laquelle des assertions suivantes est vraie ?

(a) $X_n \xrightarrow[n\to\infty]{\text{loi}} \frac{2}{\sqrt{n}}$.

(b) $\sqrt{n}X_n \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(2,1)$.

(c) $nX_n - 2\sqrt{n} \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0,1)$.

(d) Les 3 assertions ci-dessus sont fausses.

Réponse

Assertion (a) : *Fausse*
Explication : La limite d'une suite de variables indexées par n ne peut pas dépendre de n.
Assertion (b) : *Vraie*
Explication : Calculons la loi de $\sqrt{n}X_n$:

$$\sqrt{n}X_n \sim \mathcal{N}\left(\sqrt{n} \times \frac{2}{\sqrt{n}}, n \times \frac{1}{n}\right) = \mathcal{N}(2,1).$$

Ainsi, $\sqrt{n}X_n$ converge en loi vers $\mathcal{N}(2,1)$. L'assertion (b) est donc vraie.
Assertion (c) : *Fausse*
Explication : Calculons $nX_n - 2\sqrt{n}$:

$$nX_n - 2\sqrt{n} = n\left(\frac{2}{\sqrt{n}} + \frac{1}{\sqrt{n}}Z\right) - 2\sqrt{n} = 2\sqrt{n} + \sqrt{n}Z - 2\sqrt{n} = \sqrt{n}Z,$$

où $Z \sim \mathcal{N}(0,1)$. Ainsi, $nX_n - 2\sqrt{n} \sim \mathcal{N}(0,n)$. Comme la variance tend vers l'infini, cela ne peut pas converger en loi vers $\mathcal{N}(0,1)$.
Assertion (d) : *Fausse*
Explication : Puisque l'assertion (b) est vraie, l'assertion (d) est fausse.

Conclusion : *La seule assertion vraie est la **(b)**.*

Cours Question 1

Énoncé : Soit $\{P_\theta : \theta \in \Theta\}$ un modèle paramétrique, avec $\Theta \subset \mathbb{R}^p$ un ouvert. On suppose qu'il existe une fonction $\varphi : X \to \mathbb{R}^p$ telle que la fonction $M : \Theta \to \mathbb{R}^p$, définie par

$$M(\theta) = \mathbb{E}_\theta[\varphi(X_1)],$$

est continue et injective. Proposer un estimateur fortement consistant du paramètre θ, en justifiant sa consistance forte. Vous pouvez introduire des conditions supplémentaires si nécessaire.

Réponse

Estimateur proposé :

$$\widehat{\theta}_n = M^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(X_i)\right).$$

Justification : Posons $\widehat{M}_n = \frac{1}{n}\sum_{i=1}^{n}\varphi(X_i)$. Par la loi forte des grands nombres, si $\mathbb{E}_\theta[\|\varphi(X_1)\|] < \infty$,

$$\widehat{M}_n = \frac{1}{n}\sum_{i=1}^{n}\varphi(X_i) \xrightarrow{\text{p.s.}} M(\theta).$$

Sous l'hypothèse que $\widehat{M}_n$ se trouve dans l'image de M, et que M est injective et continue, il existe une application continue inversible M^{-1} sur cette image. Ainsi,

$$\widehat{\theta}_n = M^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\varphi(X_i)\right) \xrightarrow{\text{p.s.}} M^{-1}(M(\theta)) = \theta.$$

Question 2

Énoncé : Donner la définition d'un modèle dominé et d'un modèle discret. Donner un exemple de modèle statistique $\{P_\theta : \theta \in [0,1]\}$ qui n'est pas discret, alors que toutes les lois P_θ sont discrètes.

Réponse

Modèle dominé : Un modèle $\{P_\theta : \theta \in \Theta\}$ est dominé s'il existe une mesure σ-finie μ telle que $P_\theta \ll \mu$ pour tout $\theta \in \Theta$.

Modèle discret : Un modèle est discret s'il existe un sous-ensemble au plus dénombrable $E \subset \mathcal{X}$ tel que pour tout $\theta \in \Theta$, $P_\theta(E) = 1$.

Exemple : Considérons $\mathcal{X} = [0,1]$ et, pour $\theta \in [0,1]$,

$$P_\theta = \delta_\theta,$$

la mesure de Dirac en θ. Chaque P_θ est une loi discrète (soutenue par un singleton $\{\theta\}$). Cependant, si on tente de trouver un ensemble dénombrable E tel que $P_\theta(E) = 1$ pour tous les θ, on échoue, car chaque P_θ est concentré sur un point différent. Ainsi, ce modèle n'est pas discret au sens défini ci-dessus, malgré le fait que chaque P_θ soit une mesure discrète.

2 Correction de l'exercice

2.1 Proposer un estimateur sans biais et fortement consistant $\widehat{\theta}_n$ de θ

Le paramètre à estimer est $\theta = \begin{pmatrix} F(t) \\ F(s) \end{pmatrix}$, où F est la fonction de répartition commune des variables $X_1, \ldots, X_n$.

Estimateur proposé : Soit $\widehat{F}_n(x)$ la fonction de répartition empirique définie par :

$$\widehat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}.$$

On propose d'estimer θ par

$$\widehat{\theta}_n = \begin{pmatrix} \widehat{F}_n(t) \\ \widehat{F}_n(s) \end{pmatrix}.$$

Propriétés :

— Sans biais : Pour chaque composante,

$$\mathbb{E}[\widehat{F}_n(x)] = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}\right] = \frac{1}{n} \sum_{i=1}^{n} F(x) = F(x).$$

— Consistance forte : Par la loi forte des grands nombres (LFGN), on a pour tout $x \in \mathbb{R}$:

$$\widehat{F}_n(x) \xrightarrow{\text{p.s.}} F(x) \quad \text{lorsque } n \to \infty.$$

Ainsi, $\widehat{\theta}_n$ est fortement consistant.

2.2 Calculer le risque de cet estimateur. Peut-on affirmer que ce risque est inférieur ou égal à $1/(2n)$?

Le risque quadratique (ou erreur quadratique moyenne) de l'estimateur $\widehat{\theta}_n$ est donné par :

$$R_n = \mathbb{E}\left[\|\widehat{\theta}_n - \theta\|^2\right],$$

où $\|\cdot\|$ est la norme euclidienne.

Calcul du risque :

$$R_n = \mathbb{E}\left[\left(\widehat{F}_n(t) - F(t)\right)^2 + \left(\widehat{F}_n(s) - F(s)\right)^2\right].$$

Chaque composante est la variance de $\widehat{F}_n(x)$, donnée par :

$$\mathrm{Var}[\widehat{F}_n(x)] = \frac{1}{n} F(x)(1 - F(x)).$$

Ainsi :

$$R_n = \frac{1}{n}\left(F(t)(1 - F(t)) + F(s)(1 - F(s))\right).$$

Majorant du risque : Le maximum de $F(x)(1-F(x))$ est atteint pour $F(x) = \frac{1}{2}$, où $F(x)(1-F(x)) = \frac{1}{4}$. Donc :

$$R_n \leq \frac{1}{n}\left(\frac{1}{4}+\frac{1}{4}\right) = \frac{1}{2n}.$$

On peut donc affirmer que $R_n \leq \frac{1}{2n}$.

2.3 Cet estimateur $\widehat{\theta}_n$ est-il asymptotiquement normal ? Si oui, déterminer sa matrice de covariance limite

Asymptotique normale : Chaque composante $\widehat{F}_n(x)$ suit une distribution asymptotiquement normale grâce au théorème central limite :

$$\sqrt{n}\left(\widehat{F}_n(x) - F(x)\right) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, F(x)(1-F(x))).$$

Matrice de covariance limite : Les covariances entre $\widehat{F}_n(t)$ et $\widehat{F}_n(s)$ sont :

$$\mathrm{Cov}(\widehat{F}_n(t), \widehat{F}_n(s)) = \frac{1}{n}F(t)(1-F(s)).$$

Ainsi, la matrice de covariance asymptotique est donnée par :

$$\Sigma = \begin{pmatrix} F(t)(1-F(t)) & F(t)(1-F(s)) \\ F(t)(1-F(s)) & F(s)(1-F(s)) \end{pmatrix}.$$

Conclusion :

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow[n\to\infty]{\text{loi}} \mathcal{N}(0, \Sigma).$$

2.4 Voir MIP

Estimation par la méthode des moments :

$$\mathbb{E}_\lambda[X_1] = \int_0^\infty x \cdot \lambda^2 x e^{-\lambda x} dx = \frac{1}{\lambda}\int_0^\infty y^2 e^{-y} dy = \frac{2}{\lambda}.$$

Par conséquent, on a :

$$\lambda = \frac{2}{\mathbb{E}_\lambda[X_1]}.$$

L'estimateur par la méthode des moments est donné par :

$$\widehat{\lambda}_n^{\mathsf{MM}} = \frac{2}{\bar{X}_n} = \frac{2n}{\sum_{i=1}^n X_i}.$$

Consistance forte :

La loi forte des grands nombres implique la consistance de $\widehat{\lambda}_n^{\mathsf{MM}}$, combinée avec le 1^{er} théorème de continuité pour la fonction $g(x) = \frac{1}{x}$ sur $]0, +\infty[$.

2.5 Voir MIP

Calcul de $F(t)$:

$$F(t) = \int_0^t \lambda^2 x e^{-\lambda x} dx = -\int_0^t \lambda x \, d(e^{-\lambda x}).$$

Par intégration par parties :

$$F(t) = -\lambda t e^{-\lambda t} + \int_0^t \lambda e^{-\lambda x} dx,$$

$$F(t) = -\lambda t e^{-\lambda t} + \left[1 - e^{-\lambda t}\right] = 1 - (1 + \lambda t) e^{-\lambda t}.$$

Expression de θ :

$$\theta = \begin{bmatrix} F(t) \\ F(s) \end{bmatrix} = \begin{bmatrix} 1 - (1 + \lambda t) e^{-\lambda t} \\ 1 - (1 + \lambda s) e^{-\lambda s} \end{bmatrix}.$$

Cela peut être réécrit comme :

$$\theta = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} (1 + \lambda t) e^{-\lambda t} \\ (1 + \lambda s) e^{-\lambda s} \end{bmatrix}.$$

Estimateur $\widehat{\theta}_n^{\mathsf{MM}}$:
Pour simplifier l'écriture, on écrit $\hat{\lambda}_n$ au lieu de $\hat{\lambda}_n^{\mathsf{MM}}$. En utilisant les réponses aux questions précédentes, on peut définir :

$$\widehat{\theta}_n^{\mathsf{MM}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} (1 + \widehat{\lambda}_n t) e^{-\widehat{\lambda}_n t} \\ (1 + \widehat{\lambda}_n s) e^{-\widehat{\lambda}_n s} \end{bmatrix}.$$

Consistance forte de $\widehat{\theta}_n^{\mathsf{MM}}$:
La continuité des fonctions :

$$\lambda \mapsto (1 + \lambda t), \quad \lambda \mapsto (1 + \lambda s), \quad \lambda \mapsto e^{-\lambda t}, \quad \lambda \mapsto e^{-\lambda s}$$

et la consistance forte de $\widehat{\lambda}_n^{\mathsf{MM}}$ impliquent la consistance forte de $\widehat{\theta}_n^{\mathsf{MM}}$ comme estimateur de θ.

2.6 Montrer que $\widehat{\theta}_n^{\mathsf{MM}}$ est asymptotiquement normal

On sait que $\widehat{\theta}_n^{\mathsf{MM}} = h(\widehat{\lambda}_n^{\mathsf{MM}})$, où $h :]0, +\infty[\to \mathbb{R}^2$ est une fonction de classe $\mathcal{C}^1$. Le second théorème de continuité (ou encore la méthode delta que l'on a vu en TD dans les slides) implique que :

$$\sqrt{n} \left(\widehat{\theta}_n^{\mathsf{MM}} - \theta \right) \xrightarrow{\text{loi}} \mathcal{N}_2(0, \Sigma),$$

où :

$$\Sigma = h'(\lambda) \cdot \sigma^2 \cdot \left(h'(\lambda) \right)^\top,$$

avec σ^2 la variance limite de $\widehat{\lambda}_n^{\mathsf{MM}}$.

2.7 Calcul de la matrice de covariance limite Σ

On a :

$$h'(\lambda) = \begin{pmatrix} te^{-\lambda t} - te^{-\lambda t} + \lambda t^2 e^{-\lambda t} \\ se^{-\lambda s} - se^{-\lambda s} + \lambda s^2 e^{-\lambda s} \end{pmatrix} = \lambda \begin{pmatrix} t^2 e^{-\lambda t} \\ s^2 e^{-\lambda s} \end{pmatrix}.$$

D'autre part, le second théorème de continuité implique que :

$$\sqrt{n}\left(\widehat{\lambda}_n^{\mathsf{MM}} - \lambda\right) \xrightarrow{\text{loi}} \mathcal{N}\left(0, g'(\mathbb{E}[X_1])^2 \cdot \mathrm{Var}(X_1)\right).$$

Calculs intermédiaires : On sait que :

$$\mathrm{Var}(X_1) = \frac{2}{\lambda^2}, \quad \mathbb{E}[X_1] = \frac{2}{\lambda}.$$

La fonction $g(x)$ est donnée par $g(x) = \frac{2}{x}$, donc :

$$g'(x) = -\frac{2}{x^2}.$$

D'où :

$$g'(\mathbb{E}[X_1])^2 \cdot \mathrm{Var}(X_1) = \left(-\frac{2}{(2/\lambda)^2}\right)^2 \cdot \frac{2}{\lambda^2} = \frac{\lambda^2}{2}.$$

Par conséquent :

$$\sqrt{n}\left(\widehat{\lambda}_n^{\mathsf{MM}} - \lambda\right) \xrightarrow{\text{loi}} \mathcal{N}\left(0, \frac{\lambda^2}{2}\right).$$

Matrice de covariance Σ : On en déduit que :

$$\sqrt{n}\left(\widehat{\theta}_n^{\mathsf{MM}} - \theta\right) \xrightarrow{\text{loi}} \mathcal{N}_2(0, \Sigma),$$

avec :

$$\Sigma = \frac{\lambda^2}{2} \begin{pmatrix} t^4 e^{-2\lambda t} & t^2 s^2 e^{-\lambda(t+s)} \\ t^2 s^2 e^{-\lambda(t+s)} & s^4 e^{-2\lambda s} \end{pmatrix}.$$

3 Problème

3.1 Voir MIP

On a :

$$\mathbb{E}_\theta\left[\|\bar{\theta}_n - \theta\|^2\right] = \mathbb{E}_\theta\left[\|\bar{\theta}_n - \bar{\theta}'_n + \bar{\theta}'_n - \theta\|^2\right].$$

Développons cette expression :

$$\mathbb{E}_\theta\left[\|\bar{\theta}_n - \theta\|^2\right] = \mathbb{E}_\theta\left[\|\bar{\theta}_n - \bar{\theta}'_n\|^2\right] + \mathbb{E}_\theta\left[\|\bar{\theta}'_n - \theta\|^2\right] + 2\mathbb{E}_\theta\left[\langle \bar{\theta}_n - \bar{\theta}'_n, \bar{\theta}'_n - \theta\rangle\right].$$

On remarque que :

$$\bar{\theta}'_n - \theta = \mathbb{E}_\theta[\bar{\theta}_n \mid Y] - \theta$$

est une variable aléatoire Y-mesurable. Donc :

$$\mathbb{E}_\theta\left[\langle \bar{\theta}_n - \bar{\theta}'_n, \bar{\theta}'_n - \theta\rangle \mid Y\right] = \langle \mathbb{E}_\theta[\bar{\theta}_n - \bar{\theta}'_n \mid Y], \bar{\theta}'_n - \theta\rangle.$$

Cependant :

$$\mathbb{E}_\theta[\bar{\theta}_n - \bar{\theta}'_n \mid Y] = \bar{\theta}'_n - \bar{\theta}'_n,$$

ce qui donne :

$$\langle \mathbb{E}_\theta[\bar{\theta}_n - \bar{\theta}'_n \mid Y], \bar{\theta}'_n - \theta\rangle = \langle 0, \bar{\theta}'_n - \theta\rangle = 0.$$

Ainsi :

$$\mathbb{E}_\theta\left[\langle \bar{\theta}_n - \bar{\theta}'_n, \bar{\theta}'_n - \theta\rangle\right] = \mathbb{E}_\theta\left[\mathbb{E}_\theta\left[\langle \bar{\theta}_n - \bar{\theta}'_n, \bar{\theta}'_n - \theta\rangle \mid Y\right]\right] = \mathbb{E}_\theta[0] = 0.$$

En conclusion, nous avons :

$$\mathbb{E}_\theta\left[\|\bar{\theta}_n - \theta\|^2\right] = \mathbb{E}_\theta\left[\|\bar{\theta}_n - \bar{\theta}'_n\|^2\right] + \mathbb{E}_\theta\left[\|\bar{\theta}'_n - \theta\|^2\right].$$

Or :

$$\mathbb{E}_\theta\left[\|\bar{\theta}_n - \bar{\theta}'_n\|^2\right] \geq 0,$$

ce qui implique :

$$\mathbb{E}_\theta\left[\|\bar{\theta}_n - \theta\|^2\right] \geq \mathbb{E}_\theta\left[\|\bar{\theta}'_n - \theta\|^2\right].$$

3.2 Voir MIP

Comme $\bar{\theta}'_n = \mathbb{E}_\theta[\bar{\theta}_n \mid Y]$, la variable aléatoire $\bar{\theta}'_n$ dépend à première vue de θ. En effet, le calcul de l'espérance conditionnelle $\mathbb{E}_\theta[\bar{\theta}_n \mid Y]$ semble nécessiter la connaissance de θ, ce qui empêcherait d'utiliser $\bar{\theta}'_n$ comme estimateur.

Cependant, l'énoncé précise que la loi conditionnelle de X sachant Y ne dépend pas de θ. Cela implique que $\mathbb{E}_\theta[\bar{\theta}_n \mid Y]$ peut être calculée sans dépendre de θ, même si elle utilise les observations Y. Par conséquent, $\bar{\theta}'_n$ est bien un estimateur.