DM2 – Group 50

1] . On a $f_{n,\theta} = f_{n,\theta'} \iff \theta^T x_i = \theta'^T x_i \quad \forall i = \overline{1,n}$

$$\iff X_n \theta = X_n \theta'$$

X_n est de rang $p \Rightarrow \theta = \theta'$ D'où le modèle est identifiable

2] $F_n(\theta) = \sum_{i=1}^n h(\theta^T x_i)\, x_i x_i^T$

. Soit $u \in \mathbb{R}^p$, $u \neq 0$ on a : $u^T F_n u = \sum_{i=1}^n h(\theta^T x_i)\, u^T x_i x_i^T u = \sum_{i=1}^n h(\theta^T x_i) \|x_i^T u\|^2 \geq$

car $h(x) \geq 0 \ \forall x$

• $u^T F_n u = 0 \iff x_i^T u = 0 \ \forall i = \overline{1,n} \iff u = 0$ (Contradiction)

D'où $F_n(\theta)$ est définie ~~p~~ positive

3] $h(t) = \dfrac{e^t}{(1+e^t)^2} \Rightarrow h'(t) = \dfrac{e^t(1+e^t)^2 - 2(1+e^t)\, e^t . e^t}{(1+e^t)^4} = \dfrac{e^t(1-e^t)}{(1+e^t)^3}$

On a : $\left|\dfrac{e^t}{(1+e^t)^2}\right| < 1$, $\left|\dfrac{1-e^t}{1+e^t}\right| < 1 \Rightarrow |h'(t)| < 1 \quad \forall t$

D'où h est 1-lipschitzienne sur $\mathbb{R}$

4] . Vraisemblance : $L_n(\theta) = \prod_{i=1}^n \varphi(\theta^T x_i)^{Y_i} \left(1 - \varphi(\theta^T x_i)\right)^{1-Y_i}$

On a $1 - \varphi(t) = 1 - \dfrac{1}{1+e^{-t}} = \dfrac{e^{-t}}{1+e^{-t}} = \varphi(-t)$

$\Rightarrow L_n(\theta) = \prod_{i=1}^n \varphi(\theta^T x_i)^{Y_i}\, \varphi(-\theta^T x_i)^{1-Y_i}$

• Log-Vraisemblance : $\ell_n(\theta) = \sum_{i=1}^n Y_i \log(\varphi(\theta^T x_i)) + \sum_{i=1}^n (1-Y_i) \log(\varphi(-\theta^T x_i))$

$$= \sum_{i=1}^n Y_i \left[\theta^T x_i - \ln(1+e^{\theta^T x_i})\right] + \sum_{i=1}^n (1-Y_i)\left[-\ln(1+e^{\theta^T x_i})\right]$$

$$= \sum_{i=1}^n Y_i\, \theta^T x_i - \sum_{i=1}^n \ln(1+e^{\theta^T x_i})$$

5] . $\nabla \ell_n(\theta) = \sum_{i=1}^n Y_i x_i - \sum_{i=1}^n \dfrac{x_i e^{\theta^T x_i}}{1+e^{\theta^T x_i}} = \sum_{i=1}^n \{Y_i - \varphi(\theta^T x_i)\}\, x_i = X_n^T \{Y_n - \Phi_n(\theta)\}$

• $\nabla^2 \ell_n(\theta) = -X_n^T \dfrac{\partial}{\partial \theta} \Phi_n(\theta) = -(x_1 \cdots x_n) \begin{pmatrix} h(\theta^T x_1)\, x_1^T \\ \vdots \\ h(\theta^T x_n)\, x_n^T \end{pmatrix} = -\sum_{i=1}^n h(\theta^T x_i)\, x_i x_i^T$

$$= -F_n(\theta)$$

• $E_{n,\theta}[\nabla \ell_n(\theta) \nabla \ell_n(\theta)^T] = E_{n,\theta}[X_n^T \{Y_n - \Phi_n(\theta)\}\{Y_n - \Phi_n(\theta)\}^T X_n]$

$= X_n^T E_{n,\theta}[\{Y_n - \Phi_n(\theta)\}\{Y_n - \Phi_n(\theta)\}^T] X_n = X_n^T \operatorname{Cov}(Y_n) X_n = X_n^T \operatorname{diag}(h(\theta^T x_i)) X_n$

$$= \sum_{i=1}^{n} h(\theta^T x_i) x_i x_i^T = F_n(\theta)$$

• $F_n(\theta)$ est positive définie $\Rightarrow \nabla^2 l_n(\theta) = -F_n(\theta)$ est négative définie $\rightarrow l_n(\theta)$ est strictement concave presque surement.

6) • Si $Y_k = 1$ on a $\theta_*^T x_k > 0 \Rightarrow \lim_{\lambda\to\infty} \lambda \theta_*^T x_k = \infty \Rightarrow \lim_{\lambda\to\infty} \varphi(\lambda \theta_*^T x_k) = 1$

$\Rightarrow \lim_{\lambda\to\infty} \ln(\varphi(\lambda\theta_*^T x_k)) = 0 \Rightarrow \lim_{\lambda\to\infty} Y_k \ln(\varphi(\lambda\theta_*^T x_k)) = 0$

$\bullet (1-Y_k) = 0 \Rightarrow (1-Y_k)\ln(\varphi(-\lambda\theta_*^T x_k)) = 0$

$\Rightarrow Y_k \ln(\varphi(\lambda\theta_*^T x_k)) + (1-Y_k)\ln(\varphi(-\lambda\theta_*^T x_k)) \xrightarrow[\lambda\to\infty]{\text{limit}} 0$

• Si $Y_k = 0$ on a $\theta_*^T x_k < 0 \Rightarrow Y_k \ln(\varphi(\lambda\theta_*^T x_k)) + (1-Y_k)\ln(\varphi(-\lambda\theta_*^T x_k)) \xrightarrow[\lambda\to\infty]{} 0$

$\Rightarrow l_n(\lambda\theta_*) = \sum_k Y_k \ln(\varphi(\lambda\theta_*^T x_k)) + (1-Y_k)\ln(\varphi(-\lambda\theta_*^T x_k)) \xrightarrow[\lambda\to\infty]{} 0$

+) On suppose qu'il existe $\bar\theta$ t.q $\bar\theta = \arg\max_{\theta\in\mathbb{R}^p} l_n(\theta)$.

Considerer $f(\lambda) = l_n(\bar\theta + \lambda\theta_*)$

•) Si $Y_k = 1 \Rightarrow \theta_*^T x_k > 0 \Rightarrow (\bar\theta + \lambda\theta_*)^T x_k \xrightarrow{\lambda\to\infty} \infty \Rightarrow Y_k \ln(\varphi(\bar\theta+\lambda\theta_*)^T x_k)) + (1-Y_k) \ln\varphi(-(\bar\theta+\lambda\theta_*)^T x_k) \xrightarrow[\lambda\to\infty]{} 0$

• Si $Y_k = 0 \Rightarrow \theta_*^T x_k < 0 \Rightarrow (\bar\theta+\lambda\theta_*)^T x_k \to -\infty \Rightarrow Y_k \ln(\varphi(\bar\theta+\lambda\theta_*)^T x_k) + (1-Y_k)\ln\varphi(-(\bar\theta+\lambda\theta_*)^T x_k) \to 0$

D'où $f(\lambda) \xrightarrow{\lambda\to\infty} 0$ et on a $f(\lambda)$ strictement concave ; $0 = \arg\max_\lambda f(\lambda)$

Soit $\lambda_0 > 0 \Rightarrow f(\lambda_0) < f(0) \Rightarrow (1-\gamma) f(0) + \gamma f\left(\frac{\lambda_0}{\gamma}\right) \leq f(\lambda_0)$

$$\Rightarrow f\left(\frac{\lambda_0}{\gamma}\right) \leq \frac{f(\lambda_0) - (1-\gamma) f(0)}{\gamma} \quad \forall \gamma > 0, \gamma < 1$$

$$\lim_{\gamma\to 0} \frac{f(\lambda_0) - (1-\gamma) f(0)}{\gamma} \rightarrow \frac{f(\lambda_0) - f(0)}{\gamma} \rightarrow -\infty$$

$\Rightarrow f\left(\frac{\lambda_0}{\gamma}\right) \longrightarrow -\infty$ (Contradiction avec $f(\lambda) \xrightarrow{\lambda\to\infty} 0$)

Donc l'estimateur MV n'existe pas.

7) • On suppose qu'il existe $\bar\theta$ tq $\bar\theta = \arg\max_{\theta\in\mathbb{R}^p} l_n(\theta)$

Considerer $f(\lambda) = l_n(\bar\theta + \lambda\theta_*)$; f est strictement concave, $f(0) = \max_{\lambda\in\mathbb{R}} f(\lambda)$

+) Si $\begin{cases} k \notin \mathcal{E} \\ Y_k = 1 \end{cases}$ $\theta_*^T x_k > 0 \Rightarrow Y_k \ln(\varphi(\lambda\theta_* + \bar\theta)^T x_k) + (1-Y_k)\ln\varphi(-(\bar\theta+\lambda\theta_*)^T x_k) \to 0$

+) Si $\begin{cases} k \notin \mathcal{E} \\ Y_k = 0 \end{cases}$ $\theta_*^T x_k < 0 \rightarrow Y_k \ln(\varphi(\bar\theta+\lambda\theta_*)^T x_k) + (1-Y_k)\ln\varphi(-(\bar\theta+\lambda\theta_*)^T x_k) \to 0$

+) Si $k \in \mathcal{E} \Rightarrow \theta_*^T X_k = 0 \Rightarrow Y_k \ln(\varphi(\bar{\theta} + \lambda\theta_*)^T X_k) + (1-Y_k)\ln(\varphi(-(\bar{\theta}+\lambda\theta_*)^T X_k)$

$= Y_k \ln(\varphi(\bar{\theta}^T X_k)) + (1-Y_k)\ln \varphi(-\bar{\theta}^T X_k)$

(ne dépends pas à λ)

$\Rightarrow$ Il existe c t.q $\lim_{\lambda\to\infty} f(\lambda) = c$

Soit $\lambda_0 > 0$, on a $f(\lambda_0) < f(0)$. Comme dans question précédente, $f\left(\frac{\lambda_0}{\gamma}\right) \underset{\gamma\to 0}{\to} -\infty$

C'est contradiction avec $\lim_{\lambda\to\infty} f(\lambda) = c \Rightarrow$ le maximum de vraisemblance n'existe pas

8) Soit $\xi > 0$ t.q $\forall \theta \in S(0,1)$, $\exists$ $k_{1,\theta}$ et $k_{2,\theta}$ t.q

$\theta^T X_{k_{1,\theta}} > \xi$, $\theta^T X_{k_{2,\theta}} < -\xi$ et $Y_{k_1} = Y_{k_2}$

Soit I_0, I_1 deux ensembles d'indices t.q $\begin{cases} Y_i = 0 \ \forall i \in I_0 \\ Y_i = 1 \ \forall i \in I_1 \end{cases}$

Soit $k_{0,1,\theta} = \underset{k\in I_0}{\arg\max}\ \theta^T X_k$ $\qquad k_{0,2,\theta} = \underset{k\in I_0}{\arg\min}\ \theta^T X_k$

$k_{1,1,\theta} = \underset{k\in I_1}{\arg\max}\ \theta^T X_k$ $\qquad k_{1,2,\theta} = \underset{k\in I_1}{\arg\min}\ \theta^T X_k$

$\Rightarrow (\theta^T X_{k_{0,1,\theta}}, \theta^T X_{k_{0,2,\theta}}, \theta^T X_{1,1,\theta}, \theta^T X_{1,2,\theta})$ continue

Soit $O_0 = \{\theta \in S(0,1) \mid \theta^T X_{k_{0,1,\theta}}\ \theta^T X_{0,2,\theta} < 0\}$ $\Rightarrow O_0, O_1$ ouvert

$O_1 = \{\theta \in S(0,1) \mid \theta^T X_{k_{1,1,\theta}}\ \theta^T X_{1,2,\theta} < 0\}$

Par l'hypothèse, $O_0 \cup O_0 = S^P$

$\Rightarrow \exists F_0 \subset O_0$, $F_1 \subset O_1$ tq $\begin{cases} F_1, F_0 \text{ fermé, compact} \\ F_1 \cup F_0 = S^P \end{cases}$

On choisit $\xi = \min\left\{\min_{\theta\in F_0} \theta^T X_{k_{0,1,\theta}}, -\max_{\theta\in F_0} \theta^T X_{k_{0,2,\theta}}, \min_{\theta\in F_1} \theta^T X_{k_{1,1,\theta}}, -\max_{\theta\in F_1} \theta^T X_{k_{1,2,\theta}}\right\}$

- $\forall M > 0, \exists \lambda_M$ tq $\forall \theta \in S(0,1)$, $\forall \lambda > \lambda_M$: $\ln(\lambda\theta) \leq -M$:
- Car $\ln(\varphi(-\lambda\xi)) \underset{\lambda\to\infty}{\longrightarrow} -\infty$, on peut choisir λ_M tq $\ln(\varphi(-\lambda\xi)) \leq -M \quad \forall \lambda > \lambda_M$
- Soit $\theta \in S(0,1)$, $\exists k_1, k_2$ t.q $\theta^T X_{k_1,\theta} > \xi$, $\theta^T X_{k_2,\theta} < -\xi$ et $Y_{k_1} = Y_{k_2}$

+ Si $Y_{k_1} = Y_{k_2} = 0$, on a : $Y_{k_1}\ln(\varphi(\lambda\theta^T X_{k_1})) + (1-Y_{k_1})\ln(\varphi(-\lambda\theta^T X_{k_1}))$

$= \ln(\varphi(-\lambda\theta^T X_{k_1})) \leq \ln(\varphi(-\lambda\xi)) \leq -M \quad \forall \lambda > \lambda_M$

+ Si $Y_{k_1} = Y_{k_2} = 1$, on a : $Y_{k_2} \ln \varphi(\lambda \theta^T x_{k_2}) + (1 - Y_{k_2}) \ln \varphi(-\lambda\theta^T x_{k_2})$

$$= \ln \varphi(\lambda \theta^T x_{k_2}) \leq \ln \varphi(-\lambda \xi) \leq -M \ \forall \lambda > \lambda_M$$

- On a aussi $Y_k \ln \varphi(\lambda\theta^T x_k) + (1 - Y_k) \ln \varphi(-\lambda\theta^T x_k) \leq 0 \quad \forall k.$

En tous cas, on a $l_n(\lambda\theta) \leq -M$, $\forall \theta \in S(0,1)$, $\lambda > \lambda_M$

- On choisit $M > -l_n(0)$, on a $\sup\limits_{\overline{B(0,M)}} l_n(\theta) > l_n(0) > -M \geq \sup\limits_{\|\theta\| > M} l_n(\theta)$

$\Rightarrow l_n(\theta)$ atteint son maximum dans $\overline{B(0,M)}$ ($l_n(\theta)$ continue, $\overline{B(0,M)}$ compacte)

- Car $l_n \theta$ est strictement concave, il n'existe qu'un point maximal de $l_n(\theta)$.

9] On a $\| F_n(\theta) - F_n(\vartheta) \| = \| \sum_{i=1}^n [h(\theta^T x_i) - h(\vartheta^T x_i)] x_i x_i^T \|$ $(*)$

On a h est 1-lipschitzienne $\Rightarrow (*) \leq \sum_{i=1}^n |h(\theta^T x_i) - h(\vartheta^T x_i)| \, \| x_i x_i^T \|$

$$\leq \sum_{i=1}^n \|\theta - \vartheta\| \, \|x_i\| \, \|x_i x_i^T\| \leq n \|\theta - \vartheta\| \left(\frac{1}{n} \sum_{i=1}^n \|x_i\|^3\right)$$

$$\leq C n \|\theta - \vartheta\| \quad (\text{Hypothèse } H_1)$$

10] $\exists\, \bar{\theta}_n \in [\theta, \hat{\theta}_n^{MV}]$ tq $\nabla l_n(\hat{\theta}_n^{MV}) - \nabla l_n(\theta) = -F_n(\bar{\theta}_n) \cdot (\hat{\theta}_n^{MV} - \theta)$

$$\Leftrightarrow \frac{\nabla l_n(\hat{\theta}_n^{MV}) - \nabla l_n(\theta)}{\sqrt{n}} = -\frac{F_n(\bar{\theta}_n)}{\sqrt{n}\sqrt{n}} \sqrt{n} \, (\hat{\theta}_n^{MV} - \theta)$$

- $$-\frac{F_n(\bar{\theta}_n)}{n} = -\underbrace{\frac{F_n(\bar{\theta}_n) - F_n(\theta)}{n}}_{R_n} - \frac{F_n(\theta)}{n}$$

$$R_n \text{ où } \|R_n\| \leq \frac{Cn \|\bar{\theta}_n - \theta\|}{n} \leq C \|\bar{\theta}_n - \theta\|$$

On a $\hat{\theta}_n^{MV} \xrightarrow{P_{n,\theta}\text{ proba}} \theta \Rightarrow \bar{\theta}_n \xrightarrow{P_{n,\theta}} \theta \Rightarrow R_n \xrightarrow{P_{n,\theta}\text{ proba}} 0$ (Par Slutsky)

Donc $$\frac{\nabla l_n(\hat{\theta}_n^{MV}) - \nabla l_n(\theta)}{\sqrt{n}} = \left(-\frac{F_n(\theta)}{n} + R_n\right) \sqrt{n} \, (\hat{\theta}_n - \theta)$$ avec $R_n \xrightarrow{P_{n,\theta}\text{ proba}} 0$

11] D'après q5, on a : $$\frac{1}{\sqrt{n}} \nabla l_n(\theta) = \sum_{i=1}^n \underbrace{\frac{1}{\sqrt{n}} (Y_i - \varphi(\theta^T x_i)) x_i}_{T_{n,i}}$$

Vu que $\{T_{n,i}\}_{i,n}$ est un tableau triangulaire de variables aléatoires défini sur le même espace de probabilité.

On vérifie les conditions du théorème L.F :

- $E[T_{n,i}] = 0$

- $\|T_{n,i}\|^2 = \frac{1}{n}\|x_i\|^2 |Y_i - \varphi(\theta^T x_i)|^2$

Donc $\mathbb{E}[\|T_{n,i}\|^2] = \frac{1}{n} \mathrm{Var}(Y_i)\|x_i\|^2 = \frac{1}{n}\|x_i\|^2 h(\theta^T x_i) \leq \infty$

- Par l'hypothèse 1 on a : $\sum_{i=1}^{n} \mathrm{Var}(T_{n,i}) = \sum_{i=1}^{n} \frac{1}{n} x_i h(\theta^T x_i) x_i^T = \frac{1}{n} F_n(\theta) \xrightarrow{n\to\infty} Q(\theta)$

Soit $\varepsilon > 0$ $\{\|T_{n,i}\| > \varepsilon\} = \{\frac{\sqrt{n}\,\varepsilon}{\|x_i\|} < |Y_i - \varphi(\theta^T x_i)|\}$ C'est improbale car

$|Y_i - \varphi(\theta^T x_i)| < 1 \Rightarrow \mathbb{1}_{\{\|T_{n,i}\| > \varepsilon\}} = 0$ presque partout

$$\lim_{n\to\infty} \sum_{i=1}^{n} E[\|T_{n,i}\|^2 \mathbb{1}_{\{\|T_{n,i}\| > \varepsilon\}}] = 0$$

- Donc, D'après theoreme Linderberge-Feller : $\frac{1}{\sqrt{n}} \nabla l_n \theta \xrightarrow{P_{n,\theta}} \mathcal{N}(0, Q(\theta))$

<u>12</u> On a $\cancel{\frac{\nabla l_n(\hat{\theta}_n^{MV}) - \nabla l_n \theta}{\sqrt{n}} = \left(-\frac{F_n(\theta)}{n} + R_n\right)\sqrt{n}(\hat{\theta}_n - \theta)}$

$\cancel{\Leftrightarrow \sqrt{n}(\hat{\theta}_n^{MV} - \theta) = \nabla l_n(\theta)/\sqrt{n}}$ On a : $\frac{\nabla l_n(\theta)}{\sqrt{n}} = \left(\frac{F_n(\theta)}{n} - R_n\right)\sqrt{n}(\hat{\theta}_n^{MV} - \theta)$

On a $\frac{F_n(\theta)}{n} \xrightarrow{n\to\infty} Q(\theta)$; $\frac{1}{\sqrt{n}} \nabla l_n(\theta) \xrightarrow{P_{n,\theta}} \mathcal{N}(0, Q(\theta))$; $R_n \xrightarrow{P_{n,\theta}\text{-proba}} 0$

Lemme de Slutsky : $Q(0)\sqrt{n}(\hat{\theta}_n^{MV} - \theta) \xrightarrow{P_{n,\theta}} \mathcal{N}(0, Q(\theta))$

$$\Rightarrow \sqrt{n}(\hat{\theta}_n^{MV} - \theta) \xrightarrow{P_{n,\theta}} \mathcal{N}(0, Q(\theta)^{-1})$$

<u>13</u> D'après question précédente : $\sqrt{n}(\hat{\theta}_{n,k}^{MV} - \theta_k) \xrightarrow{P_{n,\theta}} \mathcal{N}(0, \beta_k)$

Où β_k est le kième coefficient de la diagonale de $\{F_n(\hat{\theta}_n^{MV})/n\}^{-1}$

$\Rightarrow \sqrt{\frac{n}{\beta_k}}(\hat{\theta}_{n,k}^{MV} - \theta_k) \xrightarrow{P_{n,\theta}} \mathcal{N}(0,1)$. On a aussi $\sqrt{\frac{\beta_k}{\beta_{n,k}}} \to 1$ d'après H1

Par Slutsky

$\to \sqrt{\frac{n}{\beta_{n,k}}}(\hat{\theta}_{n,k}^{MV} - \theta_k) \xrightarrow{P_{n,\theta}} \mathcal{N}(0,1)$

<u>14</u> Soit $z_{1-\frac{\alpha}{2}}$ la quantille niveau α de $\mathcal{N}(0,1)$, on a

$$\lim_{n\to\infty} \left(\left|\sqrt{\frac{n}{\beta_{n,k}}}(\hat{\theta}_{n,k}^{MV} - \theta_k)\right| < z_{1-\frac{\alpha}{2}}\right) = \alpha$$

$$\Rightarrow \lim_{n\to\infty} \Bigg(\underbrace{\hat{\theta}_{n,k}^{MV} - \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n/\beta_{n,k}}} \leq \theta_k \leq \hat{\theta}_{n,k}^{MV} + \frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n/\beta_{n,k}}}}_{\text{L'intervalle assymptotique pour } \theta_k}\Bigg) = \alpha$$

15| $H_0 : \theta_k = 0$ contre $H_1 : \theta_k \neq 0$

On prend le test $\mathbb{1}\left\{\left|\sqrt{\frac{n}{\beta_{n,k}}} \cdot \hat{\theta}^{MV}_{n,k}\right| > z_{1-\frac{\alpha}{2}}\right\}$ est un test de niveau asymptotique α

16| La p-valeur asymptotique de ce test satisfait:

$$|\hat{\theta}^{MV}_{n,k}| = z_{1-\frac{\hat{\alpha}}{2}} \sqrt{\frac{\beta_{n,k}}{n}} \Rightarrow z_{1-\frac{\hat{\alpha}}{2}} = \sqrt{\frac{n}{\beta_{n,k}}} \, |\hat{\theta}^{MV}_{n,k}|$$

$$\Rightarrow 1 - \frac{\hat{\alpha}}{2} = \Phi\left(\sqrt{\frac{n}{\beta_{n,k}}} \, |\hat{\theta}^{MV}_{n,k}|\right)$$ où Φ est la fonction répartition de $\mathcal{N}(0,1)$

$$\Rightarrow \hat{\alpha} = 2\left(1 - \Phi\left(\sqrt{\frac{n}{\beta_{n,k}}} \, |\hat{\theta}^{MV}_{n,k}|\right)\right)$$