Stat Fonda TD2

Exo 1

1- On a $\bar{X}_n \overset{(loi)}{=} \mathcal{N}(\theta, \frac{1}{n})$, donc

$\left[\bar{X}_n - \frac{q_{95\%}}{\sqrt{n}}, +\infty\right[$ est un IC (5%) pour θ.

~~$]-\infty, \bar{X}_n + \frac{q}{\sqrt{n}}]$ IC 95%.~~

2- D'après 1-, $P_0\left(0 \geq \bar{X}_n - \frac{q_{95\%}}{\sqrt{n}}\right) = 95\%$,

donc $1_{\bar{X}_n > \frac{q_{95\%}}{\sqrt{n}}}$ est un test de niveau 5%

pour $H_0 : \theta = 0$ contre $H_1 : \theta > 0$.

3- Notons p_θ les densités (le modèle est dominé par λ^n).

~~La vraisemblance s'écrit $V_\theta(x_{1:n})$~~

Le rapport de vraisemblance s'écrit

$$R(x) = e^{-\frac{1}{2}\left[\sum_{i=1}^n (x_i-\mu)^2 - (x_i)^2\right]}$$

$$= e^{-\frac{1}{2}\sum_{i=1}^n (\mu^2 - 2\mu x_i)} = e^{-\frac{n}{2}\mu(\mu - 2\bar{x})}.$$

Donc $\{R(x) \geq t\} \iff (2\bar{x} - \mu) \geq g(t)$, pour $t > 0$ et $g(t) = 2\frac{\log(t)}{n\mu}$.

Le test du RV est de la forme $1_{\bar{X} \geq u}$, pour un u à déterminer.

On veut $P_0(\bar{X} \geq u) = 5\%$ et on trouve $u = q_{95\%}$.

4- On prend $t = q_{95\%}$. On a

$$\int_{\mathbb{R}} (\varphi - T_{0,\mu})(f_\mu - t f_0)(x)\,dx$$

$$= \int_{\mathbb{R}} \left(\varphi - 1_{\frac{f_\mu}{f_0} \geq t}\right)(f_\mu - t f_0(x))\,dx \leq 0$$

$$= \int_{\mathbb{R}} \varphi f_\mu - t\underbrace{\int_{\mathbb{R}} \varphi f_0(x)}_{\leq \alpha} - \int_{\mathbb{R}} T_{0,\mu} f_\mu(x) + t\underbrace{\int_{\mathbb{R}} T_{0,\mu} f_0(x)}_{= \alpha}$$

$$\geq E_\mu(\varphi) - E_\mu(T_{0,\mu}).$$

Donc $T_{0,\mu}$ est plus puissant que φ.

5. On a $\begin{cases} V_{H_0}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\|x\|^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{n}{2}\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 + \bar{x}^2\right)} \\ V_{H_1}(x) = \sup_{\theta>0} \frac{1}{\sqrt{2\pi}} e^{-\frac{n}{2}\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 + (\bar{x}-\theta)^2\right)} \end{cases}$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{x})^2} \left(e^{-(\bar{x})^2 \frac{n}{2}} \mathbb{1}_{\bar{x}\leq 0} + \mathbb{1}_{\bar{x}>0} \right)$$

Donc $R_{H_1|H_0}(x) = \dfrac{e^{-\bar{x}^2 \frac{n}{2}} \mathbb{1}_{\bar{x}\leq 0} + \mathbb{1}_{\bar{x}>0}}{e^{-\frac{n}{2}\bar{x}^2}}$

$$= \mathbb{1}_{\bar{x}\leq 0} + e^{\frac{n}{2}\bar{x}^2} \mathbb{1}_{\bar{x}>0}.$$

Donc $T_{0,0^+}$ sera de la forme $\mathbb{1}_{\bar{X}\geq t}$, avec $t = q_{95\%}$.

Exo 2)

1. $P_1(M_n \leq 1-t) = (1-t)^n$. Si $\alpha>0$, $(1-t)^n = \alpha$ $\Leftrightarrow$ $t = 1-\alpha^{1/n}$. Donc $T = \mathbb{1}_{M_n \leq \alpha^{1/n}}$ est de niveau α.

2. $V_0(x) = 1$

pour $\theta<1$, $V_\theta(x) = \prod_{i=1}^{n} \frac{1}{\theta} \mathbb{1}_{x_i\leq\theta}$, donc

$V_{H_1}(x) = \frac{1}{(M_n)^n}$.

On en déduit $R(x) = \frac{1}{(M_n)^n}$. Le test du rapport de vraisemblance sera alors de la forme $\mathbb{1}_{M_n\leq a}$, et donc $T_{1,1^-} = \mathbb{1}_{M_n\leq\alpha^{1/n}}$.

3. $P_\theta(T_{1,1^-}=1) = P_\theta(M_n \leq \alpha^{1/n})$

$= P_0(\theta M_n \leq \alpha^{1/n})$

$= \frac{\alpha}{\theta^n} \xrightarrow[\theta\to 1]{} \alpha.$

Donc $\inf_{\theta\leq 1} P_\theta(T_{1,1^-}=1) = \alpha$.

4- $V_\theta(x) = \frac{1}{\theta^n} \mathbb{1}_{m_n \leq \theta}$, donc, pour $H_1 = \{\theta > 1\}$

$$\begin{cases} V_{H_1}(x) = \frac{1}{M_n^n} \mathbb{1}_{M_n > 1} + \mathbb{1}_{M_n \leq 1} \\ V_{H_0}(x) = \mathbb{1}_{M_n \leq 1} \end{cases}$$

Donc $R_{H_0|H_1}(x) = \mathbb{1}_{M_n \leq 1} + (+\infty) \mathbb{1}_{M_n > 1}$.

Le test du rapport de vraisemblance au niveau α vaut $\mathbb{1}_{M_n > 1}$ (et est stupide).

Si $\theta > 1$, $P_\theta(\{M_n > 1\}) = 1 - P_\theta(\{M_n \leq 1\})$

$= 1 - \frac{1}{\theta^n} \underset{\theta \to 1}{\longrightarrow} 0$

Donc $T_{1,1^+}$ est de puissance minimale 0.

5- On prend T de la forme $\mathbb{1}_{M_n \geq t}$.

On a alors $P_1(M_n \geq t) = \alpha \Leftrightarrow 1 - t^n = \alpha$

$\Leftrightarrow t = (1-\alpha)^{1/n}$.

T est bien de niveau α. Pour $\theta > 1$,

$$\begin{cases} P_\theta(T_{1,1^+} = 1) = 1 - 1/\theta^n \\ P_\theta(T = 1) = P_\theta(M_n \geq t) \geq P_\theta(M_n > 1) \end{cases}$$

De plus $\inf_{\theta > 1} P_\theta(M_n \geq t) = \inf_{\theta > 1} \left(1 - \left(\frac{t}{\theta}\right)^n\right) = \alpha > 0$

Exo 3

1- Sous H_0, $S \sim \mathcal{B}(n, P_m(X_1 > m))$
$\sim \mathcal{B}(n, \frac{1}{2})$.

2- $\begin{cases} H_0 : m = 38 \\ H_1 : m < 38 \end{cases}$. On cherche une région de rejet pour S de la forme.

$\{S \leq q_\alpha\}$. On cherche q_α ; $\begin{cases} \mathbb{P}_{38}(S \leq q_\alpha) \leq \alpha \\ \mathbb{P}_{38}(S \leq q_\alpha + 1) > \alpha \end{cases}$

pour un test de niveau α.

p-valeur : + gros niveau α pour lequel

On rejette H_0 à partir de $S_{obs} = 4$

$\longrightarrow \mathbb{P}_{38}(S \leq S_{obs}) = \mathbb{P}_{38}(S \leq 4)$

$\mathbb{P}_{38}(S \leq 4) = 0.0013$. (On rejette H_0).

Exo 4

1- On va tester $H_0 : p = \frac{1}{2}$ contre $H_1 : p \neq \frac{1}{2}$. Notons X le nb de naissances de garçons observées.

On a, pour $\theta \in [0,1]$, $V_\theta(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}$

Donc $V_{H_0}(x) = \binom{n}{x}/2^n$.

$\frac{\partial}{\partial\theta} V_\theta(x) = \binom{n}{x}\left(x\theta^{x-1}(1-\theta)^{n-x} - (n-x)(1-\theta)^{n-x-1}\theta^x\right)$

$\hat\theta_{EMV}$ vérifie $x(1-\hat\theta) - (n-x)\hat\theta = 0$

$\Leftrightarrow \hat\theta = x/n$. ($n$ impair donc $\hat\theta_{EMV} \neq \frac{1}{2}$)

$$V_{\hat\theta_{EMV}}(x) = \binom{n}{x}\left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x}$$

On a donc

$$R(x) = \left(\frac{x}{n}\right)^x\left(1-\frac{x}{n}\right)^{n-x} \times 2^n = 2^n\left[\left(\frac{x}{n}\right)^{\frac{x}{n}}\left(1-\frac{x}{n}\right)^{1-\frac{x}{n}}\right]$$

Avec $u = x/n$, $R(x) \geq t$ est de la forme

$$u\log(u) + (1-u)\log(1-u) \geq s.$$

avec $v = u - \frac{1}{2}$

$$u\log(u) + (1-u)\log(1-u) = \left(v+\tfrac{1}{2}\right)\log\left(v+\tfrac{1}{2}\right) + \left(\tfrac{1}{2}-v\right)\log\left(\tfrac{1}{2}-v\right)$$

$F(u) = u\log(u) + (1-u)\log(1-u)$

F sym / $\frac{1}{2}$.

$F'(u) = \log(u) + 1 - 1 - \log(1-u) = \log\left(\frac{u}{1-u}\right) \leq 0$ sur $]0, \frac{1}{2}]$.

Donc

F	0	$\frac{1}{2}$	1
u	$0 \searrow$	$\log(\frac{1}{2})$	$\nearrow 0$

Donc $\{R(x) \geq t\}$ est de la forme $\{|\frac{x}{n} - \frac{1}{2}| \geq s\}$.

Sous H_0, $\sqrt{n}\left(\frac{\cancel{x}}{n} - \frac{1}{2}\right) \sim \mathcal{N}(0, \frac{1}{4})$.

Si q est quantile d'ordre 97,5% de $\mathcal{N}(0,1)$, le test du RV au niveau 5% de $H_0 : p = \frac{1}{2}$ contre $H_1 : p \neq \frac{1}{2}$ est

$$\mathbb{1}_{|\bar{X}_n - \frac{1}{2}| > \frac{q}{2\sqrt{n}}}.$$

2) On note $\bar{X}_{n_1}$ la moyenne du 1er éch.

$\bar{X}_{n_2}$ —— 2ème éch.

Sous H_0 : $p_{1920} = p_{2015} = p$.

$\cancel{\sqrt{n}\ \bar{X}_{n_1} \simeq \mathcal{N}(}$ $\cancel{\sqrt{n_1}(\bar{X}_{n_1} - p) \simeq \mathcal{N}(0, p(1-p))}$

$\cancel{\sqrt{n_2}(\bar{X}_{n_2} - p) \simeq \mathcal{N}(0, p(1-p))}$

Donc $\cancel{\sqrt{n_1}(\bar{X}_{n_1} - p) - \sqrt{n_2}(\bar{X}_{n_2} - p) \simeq \mathcal{N}(0, 2p(1-p))}$

$\cancel{\sqrt{n_1}}$

$\bar{X}_{n_1} \simeq$

On a, sous H_0 $\bar{X}_{n_1} - \bar{X}_{n_2} \sim \mathcal{N}\left(0, p(1-p)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right)$

$\rightarrow$ μLemme $\left[\dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \xrightarrow[\substack{n_1 \to +\infty \\ n_2 \to +\infty}]{} \mathcal{N}(0, p(1-p)) \right.$

$$T = \mathbb{1}_{\frac{\bar{X}_{n_1} - \bar{X}_{n_2}}{\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \in \left[-\sqrt{p(1-p)}\, q_{97,5\%},\ \sqrt{p(1-p)}\, q_{97,5\%} \right]}.$$

est de niveau 5%.

En utilisant $\sqrt{p(1-p)} \leq \frac{1}{2}$, ou un estimateur de

$$\hat{\sigma} \left(\text{type} \sqrt{\frac{\sum_{i=1}^{n_1} (X_i^1 - \bar{X}_{n_1})^2}{2(n_1-1)} + \frac{\sum_{j=1}^{n_2} (X_j^2 - \bar{X}_{n_2})^2}{2(n_2-1)}} \right). \Big]$$

H

Ex 5

1. On a $V_{\theta_1}(X_{1:n}) = \binom{n}{n\bar{X}} \theta_1^{n\bar{X}_n} (1-\theta_1)^{n(1-\bar{X}_n)}$

Donc $R_{\theta_0|\theta_1}(x) = \dfrac{\theta_1^{n\bar{x}_n} (1-\theta_1)^{n(1-\bar{x}_n)}}{\theta_0^{n\bar{x}_n} (1-\theta_0)^{n(1-\bar{x}_n)}}$

$\ln(R_{\theta_0,\theta_1}(x)) \geq s \Leftrightarrow \underbrace{n\bar{x}_n}_{S}\left(\ln\left(\frac{\theta_1}{\theta_0}\right) + (n-S)\ln\left(\frac{1-\theta_1}{1-\theta_0}\right)\right) \geq s$

$\Leftrightarrow S \underbrace{\ln\left(\frac{\theta_1(1-\theta_0)}{\theta_0(1-\theta_1)}\right)}_{>0} \geq s - n\ln\left(\frac{1-\theta_1}{1-\theta_0}\right)$.

Donc $T_{RV,1}$ est de la forme $\mathbb{1}_{S \geq q}$, où

$$\begin{cases} P_{H_0}(S \geq q) \leq \alpha \\ P_{H_0}(S \geq q-1) > \alpha \end{cases}$$

2. $V_{H_0} = \binom{n}{S} \theta_0^S (1-\theta_0)^{n-S}$

$V_{H_1} = \binom{n}{S} \theta_0^S (1-\theta_0)^{n-S}$ si $S/n \leq \theta_0$

$= \binom{n}{S} \left(\frac{S}{n}\right)^S \left(1-\frac{S}{n}\right)^{n-S}$ si $S/n > \theta_0$

$$R(X) = \mathbb{1}_{S/n \leq \theta_0} + \underbrace{\frac{\cancel{\theta_0(1-\theta)} \left(\frac{S}{n}\right)^S \left(1-\frac{S}{n}\right)^{n-S}}{\theta_0^S (1-\theta_0)^{n-S}}}_{h(S)} \mathbb{1}_{\frac{S}{n} > \theta_0}$$

$$g(S) = \cancel{\$}\ S \log\left(\frac{S}{n\theta_0}\right) + (n-S)\log\left(\frac{1-S/n}{1-\theta_0}\right)$$

$$g'(s) = \log\left(\frac{s}{n\theta_0}\right) + 1 - \log\left(\frac{1-S/n}{1-\theta_0}\right)$$

$$+ (n-S) \times \frac{\frac{-1/n}{1-\theta_0}}{\frac{1-S/n}{1-\theta_0}}$$

$$= \log\left(\frac{s}{n\theta_0}\right) + 1 - \log\left(\frac{1-S/n}{1-\theta_0}\right) - 1$$

$$= \underbrace{\log\left(\frac{s/n(1-\theta_0)}{\theta_0(1-s/n)}\right)}_{\geq 0} \text{ si } \frac{s}{n} > \theta_0.$$

Avec ~~$g(\theta_0)$~~ $h(\theta_0) = 1$.

Donc $\{R(X) \geq t\}$ est de la forme $S \geq q$.

Avec $P_{H_0}(S \geq q) \leq \alpha$, $P_{H_0}(S \geq q-1) > \alpha$.

On suppose $P_{H_0}(S \geq q) = \alpha$.

~~Si φ est de niveau α, et $\theta > \theta_0$.~~

$$\cancel{E_\theta((\mathbb{1}_{S \geq q} - \varphi))}$$

$$\cancel{= E_{\theta_0}\left(\frac{\theta^S (1-\theta)^{n-S}}{\theta_0^S (1-\theta_0)^{n-S}} (\mathbb{1}_{S \geq q} - \varphi)\right)}$$

On a alors $T_{RV,2} = T_{RV,1}$ (celui du 1-)

Si $\theta > \theta_0$, d'après 1-, si φ est de niveau α, $E_\theta(\varphi) \leq E_\theta(T_{RV,1}) = E_\theta(T_{RV,2})$ □

3-| On a $V_{H_0} = \binom{n}{S} \theta_0^S (1-\theta_0)^{n-S}$ si $\frac{S}{n} > \theta_0$

$$= \binom{n}{S} \left(\frac{S}{n}\right)^S \left(1-\frac{S}{n}\right)^{n-S} \text{ si } \frac{S}{n} \leq \theta_0$$

IV

$$V_{H_1} = \begin{cases} \binom{n}{S} \theta_0^S (1-\theta_0)^{n-S} & \text{si } \frac{S}{n} \leq \theta_0 . \\ \binom{n}{S} \left(\frac{S}{n}\right)^S (1-S/n)^{n-S} & \text{si } S/n > \theta_0 \end{cases}$$

On a $RV_3 = \mathbb{1}_{S/n \leq \theta_0} \left(\frac{\theta_0}{S/n}\right)^S \left(\frac{1-\theta_0}{1-S/n}\right)^{n-S}$) $\nearrow S$ d'après 2.

$$+ \mathbb{1}_{S/n > \theta_0} \left(\frac{S/n}{\theta_0}\right)^S \left(\frac{1-S/n}{1-\theta_0}\right)^{n-S}$$

$\nearrow S$ d'après 2.

Donc $T_{RV,3} = \mathbb{1}_{S \geq q} = T_{RV,1}$ et on a encore un test UPP.

4-) En admettant Lemme 6.1 du cours : si T' de niveau exactement α et de même puissance qu'un $RV(\alpha)$ (entre θ_1 et θ_0), alors $T' \mathbb{1}_{RV \neq \tau_\alpha}$

$$= T \mathbb{1}_{RV \neq \tau_\alpha}$$

(τ_α seuil qui définit T).

Supposons que T soit un test UPP pour $\Theta_0 = [\theta_0 - \varepsilon, \theta_0 + \varepsilon]$, $\Theta_1 = [0,1] \setminus \Theta_0$.

Si $\theta_1 > \theta_0$, on a T domine T_{θ_0,θ_1}.

Donc T est de la forme $\mathbb{1}_{S \geq q_{\alpha-}}$ ($T = T_{\theta_0,\theta_1}$ vu qu'il est pile-poil de niveau α).

Inversement, si $\theta_1 < \theta_0$, T domine T_{θ_0,θ_1}.

Donc T est de la forme $\mathbb{1}_{S \leq q_{\alpha+}}$.

Donc $T = 1$ et $\alpha = 1$.