

1. $(\mathbb{N}^n, \mathcal{P}(\mathbb{N}^n), \mathbb{B} \mathcal{P}(\theta)^{\otimes n}, \theta \in \mathbb{R}^+)$

2. Vraisemblance:

$$\begin{aligned} \mathbb{P}_\theta(Y_1=y_1, \dots, Y_n=y_n) &= \frac{\theta^{y_1} e^{-\theta}}{y_1!} \dots \frac{\theta^{y_n} e^{-\theta}}{y_n!} \\ &= \frac{\theta^{n\bar{y}} e^{-n\theta}}{y_1! \dots y_n!} \end{aligned}$$

$$l_\theta = n\bar{y} \log \theta - n\theta$$

on dérive

$$\frac{n\bar{y}}{\theta} - n$$

$$\hat{\theta} = \bar{y}$$

$$\begin{aligned} 3. E(e^{sN}) &= \sum_k \frac{\theta^k e^{-\theta}}{k!} e^{sk} = \exp(\theta e^s - 1) \\ &= \exp(\theta(e^s - 1)) \end{aligned}$$

$$\begin{aligned} 4. E(e^{s(N_1 + N_2)}) &= E(e^{sN_1}) E(e^{sN_2}) \\ &= e^{\theta_1(e^s - 1)} e^{\theta_2(e^s - 1)} \text{ donc } N_1 + N_2 \sim \mathcal{P}(\theta_1 + \theta_2) \end{aligned}$$

5. soit $\theta_1 < \theta_2$

$$\mathbb{P}_{\theta_1}(N \leq t)$$

$$\leq \mathbb{P}_{\theta_1}(N+M \geq t)$$

$$= \mathbb{P}_{\theta_2}(N \geq t)$$

$$\mu = \theta_2 - \theta_1$$

$$M \sim \mathcal{P}(\mu) \perp N.$$

$$\text{alors } N+M \sim \mathcal{P}(\theta_2)$$

6. Par récurrence $Y_1 + \dots + Y_n \sim \mathcal{P}(n\theta)$.

$$\begin{aligned} 7. E(e^{i\frac{N-1}{\sqrt{n}}}) &= E(e^{i\frac{N}{\sqrt{n}} - \frac{i}{\sqrt{n}}}) = e^{-\frac{i}{\sqrt{n}}} E(e^{i\frac{N}{\sqrt{n}}}) = e^{-\frac{i}{\sqrt{n}}} (e^{\frac{i\theta}{\sqrt{n}} - 1}) = e^{-\frac{i}{\sqrt{n}}} e^{\frac{i\theta}{\sqrt{n}}} e^{-1} \\ &\xrightarrow{n \rightarrow \infty} e^{-\frac{\theta^2}{2} + o(n^2)} \text{ i.e. } E(e^{i\frac{N-1}{\sqrt{n}}}) \rightarrow e^{-\frac{\theta^2}{2}} \end{aligned}$$

8. du côté du propriétaire

$$H_0: \theta \leq 100$$

$$H_1: \theta > 100$$

donc on rejette qd $\hat{\theta}$ trop grand $\hat{\theta} > c$

maintenant il faut calibrer c

$$\begin{aligned} \mathbb{P}_{100}(\hat{\theta} > c) &= \mathbb{P}_{100}(Y_1 + \dots + Y_n > nc) = \mathbb{P}_{100}(N > nc) \\ &= \mathbb{P}\left(\frac{N - n100}{\sqrt{n100}} > \frac{nc - n100}{\sqrt{n100}}\right) \\ &\approx \mathbb{P}(U > \frac{nc - n100}{\sqrt{n100}}) \end{aligned}$$

$$\nearrow nc = t_{\alpha, n100}$$

donc si $\frac{nc - n100}{\sqrt{n100}} = \Phi_{\alpha, n100} \quad \Phi_{\alpha}$ quantile gaussienne

$$ie \quad c = \frac{(\Phi_{\alpha, n100} \times \sqrt{n100}) + n100}{n}$$

Verifions test de taille α

$\forall \theta \leq 100$

$$P_{\theta}(\hat{\theta} > c) = P_{n\theta}(Y_1 + \dots + Y_n > nc) \quad f^o T \text{ de } n\theta$$

$$\leq P_{n100}(Y_1 + \dots + Y_n > nc) = \alpha$$

donc test de taille α (atteint en $\theta = 100$)

puis

$$\hat{\theta} = \bar{Y} = 103,4 =$$

on cherche α tq $c = 103,4$
= valeur

$$ie \quad \frac{\Phi_{\alpha} \times \sqrt{n100} + n100}{n} = 103,4$$

$$\begin{aligned} \Phi_{\alpha} &= \frac{103,4 \times n - n100}{\sqrt{n100}} \\ &= \frac{n \cdot 3,4}{\sqrt{n100}} = 3,4 \frac{\sqrt{30}}{\sqrt{100}} = 3,4 \frac{\sqrt{3}}{\sqrt{10}} = 1,862 \end{aligned}$$

donc $2,5\% \leq \alpha \leq 5\%$

la p-value θ est entre 2,5% et 5% on rejette le test : l'eau est polluée.

9. Test du rapport de vraisemblance :

$$\frac{L_1^{n\bar{Y}} e^{-n\lambda_1} / Y_1! \dots Y_n!}{L_0^{n\bar{Y}} e^{-n\lambda_0} / Y_1! \dots Y_n!} = \left(\frac{\lambda_1}{\lambda_0}\right)^{n\bar{Y}} e^{-n(\lambda_1 - \lambda_0)}$$

Comme $\lambda_1 / \lambda_0 > 1$ rejeter qd rapport grand

$\Rightarrow$ rejeter $\bar{Y}$ grand.

et la calibration précédente dit de taille α

10. On verifie vraisemblance mancrone par 3. du th sur la partie randomisation.

II.1. $Y = X \begin{pmatrix} a \\ b \end{pmatrix} + \varepsilon$
avec $Y \in \mathbb{R}_n$

$$X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \text{ de } \mathbb{R}^2 \text{ donc } \text{rang } X = 2 \text{ si } x_i \neq x_j$$

$$\varepsilon \sim \mathcal{CN}_n(0, \sigma^2 I_n)$$

2. $\hat{b} = \frac{\text{Cov}_n(x, Y)}{\text{Var}_n x} = \frac{\overline{xy} - \bar{x} \bar{y}}{\text{Var}_n x} = -1,14$

$$\hat{a} = \bar{y} - \bar{x} \hat{b} = 121,05$$

$$\hat{\sigma}^2 = \frac{\sum (y_i - (\hat{a} + \hat{b} x_i))^2}{n-2}$$

3.

IC en $a + b x_{30}$

$$d = \begin{pmatrix} 1 \\ x_{30} \end{pmatrix}$$

$$\frac{d^* \hat{\theta} - d^* \theta}{\sqrt{\hat{\sigma}^2 (d^* (X^* X)^{-1} d)}} \sim T(28)$$

$$\sqrt{\hat{\sigma}^2 (d^* (X^* X)^{-1} d)}$$

donc $d^* \theta \in d^* \hat{\theta} \pm t_{28}^{2/2} \sqrt{\hat{\sigma}^2} \sqrt{(1, 30) \begin{pmatrix} \bar{x}^2 - \bar{x} \\ -\bar{x} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 30 \end{pmatrix}}$

$$\sqrt{(1, 30) \begin{pmatrix} \bar{x}^2 - 30\bar{x} \\ -\bar{x} + 30 \end{pmatrix}}$$

$$\sqrt{\frac{\bar{x}^2 - 60\bar{x} + 900}{n(\bar{x}^2 - (\bar{x})^2)}}$$

$$[\hat{a} + \hat{b} x_{30} \pm 2,048 \sqrt{\hat{\sigma}^2} \sqrt{\bar{x}^2 - 60\bar{x} + 900}]$$

$$= [81,24; 92,4] \quad 86,8 \pm \frac{2,048 \times 7,62}{5,56} \times \sqrt{16,89}$$

III

1. $Y_1, \dots, Y_n$ i.i.d de f_0 de réf F
On teste $H_0: F = F_0$ contre $H_1: F \neq F_0$.

avec $D_n = \sqrt{n} \sup |F_n(n) - F_0(n)|$ où $F_0 \sim F$ de $\mathcal{D}(1, 1, 1)$
on rejette qd $D_n = \sqrt{n} \sup | | \geq c_\alpha$.

On tabulé on ne peut pas la voir F_0
et $\rightarrow$ rep. fixé

or D_n ici = 1.131

donc proba $\geq 15\%$ on ne peut pas vraiment
conclure

2.

$H_0: F = F_0$ contre $H_1: F \neq F_0$

sous H_0 si $N_1 = \# \text{ données} \leq 100$

$N_2 = \text{---} > 100$

alors (N_1, N_2) binomiale $\mathcal{B}_n(\frac{1}{2})$.

or ici $N_1 = 10$ $N_2 = 20$.

$np_i = 15 \geq 5$ donc test χ^2 valable

$$\chi^2 = \frac{(N_1 - np_1)^2}{np_1} + \frac{(N_2 - np_2)^2}{np_2}$$

$$= \frac{(10 - 15)^2}{15} + \frac{(20 - 15)^2}{15} = 2 \times \frac{5^2}{15} = \frac{2 \times 5}{3} = 3.33$$

$$\chi^2 \xrightarrow{\alpha} \chi^2(1)$$

d'où proba $< 7\%$

et donc proba de voir ces données sous H_0

$\leq$ proba de voir $N_1 = 10$ et $N_2 = 20$.

$\leq$ proba $(\chi^2(n) \geq \chi^2_{\alpha})$.

$\leq 7\%$.

3. on forme $x_i = \frac{Y_i - \bar{Y}}{\sqrt{\hat{\sigma}^2}}$ ce n'est plus ind ms sous H_0

pour même loi donc tabulable

et on forme $\sqrt{n} \sup |F_n - F_0|$ où $F_0 = \text{---}$

IV.

$$f(x, m) = \frac{e^{-\frac{\sum (x_i - m)^2}{2}}}{(\sqrt{2\pi})^n}$$

$$1. \log l_m(x_1, \dots, x_n) = -\frac{\sum (x_i - m)^2}{2} + n \log \sqrt{2\pi}$$

$$\frac{\partial}{\partial m} = -(\sum x_i - nm) \text{ donc max en } \hat{m} = \bar{x}$$

$$2. E(\bar{Y}) = m.$$

2. ~~forme~~ Info de Fisher

$$I(m) = E\left[\left(\frac{\partial}{\partial m}\right)^2\right] = n.$$

or Var $\bar{Y} = \frac{1}{n}$ donc atteint la borne

ie estimateur sans biais de + petite variance

$$3. d_0, d_1, \dots, d_n$$

$$E(e^{i(d_0 m + d_1 Y_1 + \dots + d_n Y_n)})$$

$$= E(e^{i d_0 m} E(Y_1 | m))$$

$$= E(e^{i d_0 m} e^{i d_1 m} e^{-\frac{d_1^2}{2}} \dots e^{-i d_n m} e^{-\frac{d_n^2}{2}})$$

$$= e^{i(d_0 + d_1 + \dots + d_n) \eta} e^{-\frac{(d_0 + d_1 + \dots + d_n)^2 \sigma^2}{2}} e^{-\frac{d_1^2}{2} - \dots - \frac{d_n^2}{2}}$$

soit la forme est gaussien de V.G.

$$\text{densité } f(y_1, \dots, y_n, m) = \frac{e^{-\frac{\sum (y_i - m)^2}{2}} e^{-\frac{(m - \eta)^2}{2\sigma^2}}}{(\sqrt{2\pi})^n \sqrt{2\pi}\sigma^2} dy_1 \dots dy_n dm$$

4. donc loi de $m | Y_1, \dots, Y_n$.

$$\frac{f(y_1, \dots, y_n, m) dm}{\int f(y_1, \dots, y_n, m') dm'} = \frac{e^{-\left(m - \frac{\eta + n\bar{Y}\sigma^2}{n\sigma^2 + 1}\right)^2 / \frac{\sigma^2}{n\sigma^2 + 1}}}{\int -}$$

5.

donc $T_{\eta, z}$ est l'estimateur de risque bayésien minimum

$$\begin{aligned} \text{donc } T_{\eta, z} &= E(m | x_1, \dots, x_n) \\ &= \frac{\eta + n\bar{y}z^2}{nz^2 + 1} \quad \text{v.u.} \end{aligned}$$

$$\begin{aligned} r_{\eta, z} &= E \left((T_{\eta, z} - m)^2 \right) \\ &= E \left(E \left(E(m | x_1, \dots, x_n) - m \right)^2 | x_1, \dots, x_n \right) \\ &= E \left(\text{Var } m | x_1, \dots, x_n \right) = \frac{\sigma^2}{nz^2 + 1} \end{aligned}$$

6. $\sup_{m \in \mathbb{R}} E_m((\bar{y} - m)^2) = \frac{1}{n}$ car le sup atteint est $\mathbb{R}$.

$$r_{\eta, z} \xrightarrow{z \rightarrow \infty} \frac{1}{n}$$

ou si T autre estimateur

$$r_{\eta, z} \leq E((T - m)^2) \leq \sup_{m \in \mathbb{R}} E_m((T - m)^2)$$

$$\downarrow$$

$$\sup_{m \in \mathbb{R}} E_m((\bar{y} - m)^2)$$

et donc $\bar{y}$ bayésien.