Page 2

Question N°1

$$\mathcal{D} = \{x_i, y_i\} = i = 1, \dots, n \text{ et } p = 1$$

1-a) Détermination de l'estimateur de maximum de vraisemblance

$$f(y_i|\theta, x_i) = \begin{cases} 2(x_i\theta)^2 y_i^3 \exp\{-\theta x_i y_i^2\} & \text{si } y_i \geq 0,\ x_i \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

$$L(\theta) = \prod_{i=1}^{n} f(y_i|\theta, x_i)$$

$$= \prod_{i=1}^{n} 2(x_i\theta)^2 y_i^3 \exp\{-\theta x_i y_i^2\}$$

$$\mathcal{L}(\theta) = \ln \prod_{i=1}^{n} 2(x_i\theta)^2 y_i^3 e^{-\theta x_i y_i^2}$$

$$= \sum_{i=1}^{n} \left[\ln(2) + 2\ln(x_i\theta) + 3\ln y_i - \theta x_i y_i^2\right.$$

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = \sum_{i=1}^{n} \left[\frac{2}{\theta} - x_i y_i^2\right]$$

$$= \frac{2n}{\theta} - \sum_{i=1}^{n} x_i y_i^2$$

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = 0 \Rightarrow \boxed{\hat{\theta}_{emv} = \frac{2n}{\sum_{i=1}^{n} x_i y_i^2}}$$

$$\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta^2} = -\frac{2n}{\theta^2} < 0$$; nous sommes donc à un maximum.

• Information de Fisher

$$I(\theta) = -E\left[\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta^2}\right]$$

$$= -E\left[-\frac{2n}{\theta^2}\right] = \frac{2n}{\theta^2}$$

Donc $\boxed{I(\theta) = \frac{2n}{\theta^2}}$

La variance asymptotique de θ est

$$V(\hat{\theta}_{evm}) = \frac{1}{I(\theta)} = \frac{1}{\frac{2n}{\theta^2}} = \frac{\theta^2}{2n}$$

Alors lorsque n est grand on a :

$$\boxed{\hat{\theta}_{evm} \rightsquigarrow \mathcal{N}\left(\theta x_i, \frac{\theta^2}{2n}\right)}$$

1-b) Détermination de l'estimateur de maximum de vraisemblance θ_{evm} de θ.

$$D = \{x_i ; y_i\} \sim P(x_i, \theta)$$

$$P(y_i | \theta, x_i) = \begin{cases} \binom{y_i - 1}{x_i - 1} \theta^{x_i} (1-\theta)^{y_i - x_i} & \text{si } y_i = x_i, x_{i+1}, \dots \\ 0 & \text{sinon.} \end{cases}$$

$$L(\theta) = \prod_{i=1}^{n} \binom{y_i - 1}{x_i - 1} \theta^{x_i} (1-\theta)^{y_i - x_i}$$

$$\mathcal{L}(\theta) = \ln \prod_{i=1}^{n} \binom{y_i - 1}{x_i - 1} \theta^{x_i} (1-\theta)^{y_i - x_i}$$

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

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = \sum_{i=1}^{n} \left[\frac{x_i}{\theta} - \frac{(y_i - x_i)}{(1-\theta)} \right]$$

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = 0 \Rightarrow \sum_{i=1}^{n} \frac{x_i}{\theta} - \sum_{i=1}^{n} \frac{(y_i - x_i)}{(1-\theta)} = 0$$

Après transformation et résolution d'equation $\Rightarrow$ $$\boxed{\hat{\theta}_{evm} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i}}$$

$$\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta^2} = \sum_{i=1}^{n} \left[-\frac{x_i}{\theta^2} - \frac{(y_i - x_i)}{(1-\theta)^2} \right] < 0$$

Nous somme donc à un maximum

- Information de Fisher

$$I(\theta) = -E\left[\frac{\partial^2 \mathcal{L}(\theta)}{\partial \theta^2}\right]$$

$$= -E\left[\sum_{i=1}^{n}\left[-\frac{x_i}{\theta^2} - \frac{(y_i - x_i)}{(1-\theta)^2}\right]\right]$$

$$= \sum_{i=1}^{n}\left[E\left[\frac{x_i}{\theta^2}\right] + E\left[\frac{y_i - x_i}{(1-\theta)^2}\right]\right]$$

$$E(Y_i) = \frac{x_i}{\theta}$$

Alors $I(\theta) = \sum_{i=1}^{n}\left[\frac{x_i}{\theta^2} + \frac{1}{(1-\theta)^2} \times \left(\frac{x_i}{\theta} - x_i\right)\right]$

$$\boxed{I(\theta) = \sum_{i=1}^{n} \frac{x_i}{\theta^2(1-\theta)}} \quad V(\hat{\theta}_{emv}) = \frac{1}{n I(\theta)}$$

lors n devient grand,

$$\boxed{\hat{\theta}_{emv} \rightsquigarrow \mathcal{N}\left(\theta, \frac{1}{n\sum_{i=1}^{n}\frac{x_i}{\theta^2(1-\theta)}}\right)}$$

1-c) On donne :

$$\pi(\theta) \begin{cases} 2(1-\theta) & \text{si } 0<\theta<1 \\ 0 & \text{sinon} \end{cases}$$

Déterminons l'estimateur MAP de θ

$$L(\theta|y_i) = \prod_{i=1}^{n} \binom{y_i - 1}{x_i - 1} \theta^{x} (1-\theta)^{y_i - x_i}$$

* Recherchons π à postériori

$$\pi(\theta|y_i) = \frac{P(y_i|\theta, x_i)\,\pi(\theta)}{\int_{-\infty}^{\infty} P(y_i|\theta, x_i)\,d\theta}$$

or $\int_{-\infty}^{\infty} P(y_i|\theta, x_i)\,d\theta = \text{Cons}^{te} = c$

Posons $\frac{1}{c} = K = \text{constante}$

Alors $\pi(\theta|y_i) = K \times P(y_i|\theta, x_i)\,\pi(\theta)$

$$= K \times \binom{y_i - 1}{x_i - 1} \theta^{x_i} (1-\theta)^{y_i - x_i} \times 2(1-\theta)$$

Or $\hat{\theta}_{MAP} = \text{Arg}\max\{\pi(\theta|y_i)\}$

$$= \operatorname{Arg\,max}\left\{ \ln\left[2k \binom{y_i - 1}{x_i - 1} \theta^{x_i} (1-\theta)^{y_i - x_i + 1} \right] \right.$$

$$= \sum_{i=1}^{n} \operatorname{Arg\,max} \left\{ C^{ste} + x_i \ln\theta + (y_i - x_i + 1)\ln(1-\theta) \right\}$$

$$\frac{\partial \mathcal{L}(\theta)}{\partial \theta} = 0 \Rightarrow \frac{1}{\theta}\sum_{i=1}^{n} x_i - \frac{1}{1-\theta}\sum_{i=1}^{n}(y_i - x_i + 1) = 0$$

$$\Longleftrightarrow \frac{1}{\theta}\sum_{i=1}^{n} x_i = \frac{1}{1-\theta}\sum_{i=1}^{n}(y_i - x_i + 1)$$

$$(1-\theta)\sum_{i=1}^{n} x_i = \theta\sum_{i=1}^{n}(y_i - x_i + 1)$$

$$\sum_{i=1}^{n} x_i = \theta\left(\sum_{i=1}^{n} y_i + n\right)$$

D'où $$\boxed{\hat{\theta}_{MAP} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i + n}}$$

C-2) Estimateur de Bayes

Considerons la fonction de perte quadratique.

$$\hat{\theta}_{\pi} = \operatorname{arg\,min} E_{\pi}\left[(\theta - \hat{\theta})^2 \mid Y_1 \dots Y_n \right]$$

$$= E_{\pi}(\theta \mid Y_1 \dots Y_n)$$

$$= \int_{-\infty}^{+\infty} \theta \, \pi(\theta | y_i) \, d\theta$$

On sait aussi que :

$$L(\theta) = \prod_{i=1}^{n} \binom{y_i - 1}{x_i - 1} \theta^{x_i} (1-\theta)^{y_i - x_i}$$ qui

est la fonction de vraissemblance

Sa fonction à posteriori est

$$\pi(\theta) = \text{Conste} \times \theta^{\sum x_i} \times (1-\theta)^{\sum_{i=1}^{n}(y_i - x_i) + 1}$$

poson $\alpha - 1 = \sum_{i=1}^{n} x_i$ et $\beta - 1 = \sum_{i=1}^{n}(y_i - x_i) + 1$

$\Rightarrow \alpha = \sum_{i=1}^{n} x_i + 1$ et $\beta = \sum_{i=1}^{n}(y_i - x_i) + 2$

On observe donc que la distribution à priori est de type Beta

$$\beta = \left(\alpha = \sum_{i=1}^{n} x_i + 1 \text{ et } \beta = \sum_{i=1}^{n}(y_i - x_i) + 2\right)$$

Pour une loi Beta(α, β), l'esperance est :

$$E(\pi(\theta | y_i)) = \frac{\alpha}{\alpha + \beta}$$

$$\Rightarrow E\left[\pi(\theta|y_i)\right] = \frac{\sum_{i=1}^{n} x_i + 1}{\left(\sum_{i=1}^{n} x_i + 1\right) + \left(\sum_{i=1}^{n}(y_i - x_i) + 2\right)}$$

$$= \frac{\sum_{i=1}^{n} x_i + 1}{\sum_{i=1}^{n} x_i + 1 + \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i + 2}$$

$$E\left[\pi(\theta|y_i)\right] = \frac{\sum_{i=1}^{n} x_i + 1}{\sum_{i=1}^{n} y_i + 3}$$

Alors $$\boxed{\hat{\theta}_{Bayes} = \frac{\sum_{i=1}^{n} x_i + 1}{\sum_{i=1}^{n} y_i + 3}}$$

Compareison entre $\hat{\theta}_{MAP}$ et $\hat{\theta}_{Bayes}$

$$\hat{\theta}_{map} = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i + 1}$$

$$\hat{\theta}_{Bayes} = \frac{\sum_{i=1}^{n} x_i + 1}{\sum_{i=1}^{n} y_i + 3}$$

- $\hat{\theta}_{map}$ maximise la densité à postériori, tandis que $\hat{\theta}_{Bayes}$ est l'esperance de la distribution à postérioni

- Par ailleur l'estimateur de Bayes est légèrement plus régularisé, car l'ajout de (+1) au numérateur et (+2) au dénominateur, ce qui correspond à une forme de régularisation introduite par la distribution à priori.

En fin θ_{map} est plus concentré sur la valeur qui maximise la probabilité, alors que Bayes prend en compte la forme complete de la distribution.

Ces deux estimateurs tendent à donner des valeurs proches lorsque le nombre d'observation "n" est grand, car l'influence de la distribution à priori diminue par rapport à la vraissemblance basée sur les données.

Question 2

$\{(x_i, y_i); i = 1, \dots, n\}$ $\quad p = 1 \quad \sum_{i=1}^{n} u_i = 1$

2-a) Détermination de l'estimateur des moindres carrés $\hat{\beta}$ definit par

$\hat{\beta} = \arg\min RSS_1(\beta)$ où $RSS_1(\beta) = \sum_{i=1}^{n} (y_i - \beta x_i)^2$

$$\frac{\partial}{\partial \beta} RSS_1(\beta) = \frac{\partial}{\partial \beta}\left[\sum_{i=1}^{n} (y_i - \beta x_i)^2\right]$$

$$= 2\left[\sum_{i=1}^{n} \left[(-x_i)(y_i - \beta x_i)\right]\right]$$

$$= 2\sum_{i=1}^{n} (x_i y_i - \beta x_i^2)$$

$$= 2\sum_{i=1}^{n} x_i y_i + 2\beta \sum_{i=1}^{n} x_i^2$$

$$\frac{\partial}{\partial \beta} RSS_1(\beta) = 0 \Rightarrow \beta \sum_{i=1}^{n} x^2 = \sum_{i=1}^{n} x_i y_i$$

D'où $\boxed{\hat{\beta} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}}$

2-b) Détermination de $\hat{\beta}_0$ et $\hat{\beta}_1$

$(\hat{\beta}_1, \hat{\beta}_2) = \arg\min RSS_2(\beta_0, \beta_1)$

$$RSS_2(\beta_0, \beta_1) = \sum_{i=1}^{n} \omega_i (Y_i - \beta_0 - \beta_1 X_i)^2$$

$$\frac{\partial RSS_2}{\partial \beta_0} = -2 \sum_{i=1}^{n} \omega_i (Y_i - \beta_0 - \beta_1 X_i) = 0 \quad (1)$$

$$\frac{\partial RSS_2}{\partial \beta_1} = -2 \sum_{i=1}^{n} \omega_i X_i (Y_i - \beta_0 - \beta_1 X_i) = 0 \quad (2)$$

De (1), on a :

$$\sum_{i=1}^{n} \omega_i (Y_i - \beta_0 - \beta_1 X_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} \omega_i Y_i - \beta_0 \sum_{i=1}^{n} \omega_i - \beta_1 \sum_{i=1}^{n} \omega_i X_i = 0 \quad (1)$$

De (2), on a :

$$\sum_{i=1}^{n} \omega_i X_i (Y_i - \beta_0 - \beta_1 X_i) = 0$$

$$\sum_{i=1}^{n} \omega_i X_i Y_i - \beta_0 \sum_{i=1}^{n} \omega_i X_i - \beta_1 \sum_{i=1}^{n} \omega_i X_i^2 = 0$$

$\sum_{i=1}^{n} \omega_i = 1$ alors posons :

$$\overline{X}_\omega = \sum_{i=1}^{n} \omega_i X_i \quad \text{et} \quad \overline{Y}_\omega = \sum_{i=1}^{n} \omega_i Y_i$$

Donc $\overline{Y}_\omega - \beta_0 - \beta_1 \overline{X}_\omega = 0$

De ②, on a aussi

$$\sum_{i=1}^{n} \omega_i x_i y_i - \beta_0 \overline{X}_\omega - \beta_1 \sum_{i=1}^{n} \omega_i X_i^2 = 0$$

si $\beta_0 = \overline{Y}_\omega - \beta_1 \overline{X}_\omega$ alors

$$\sum_{i=1}^{n} \omega_i x_i y_i - (\overline{Y}_\omega - \beta_1 \overline{X}_\omega)\overline{X}_\omega + \beta_1 \sum_{i=1}^{n} \omega_i X_i^2$$

$$\sum_{i=1}^{n} \omega_i x_i y_i - \overline{Y}_\omega \overline{X}_\omega = \beta_1 \left(\sum_{i=1}^{n} \omega_i x_i^2 - \overline{X}_\omega^2 \right)$$

$$\boxed{\beta_1 = \frac{\sum_{i=1}^{n} \omega_i x_i y_i - \overline{Y}_\omega \overline{X}_\omega}{\sum_{i=1}^{n} \omega_i X_i^2 - \overline{X}_\omega^2}}$$

$$\boxed{\beta_0 = \overline{Y} - \beta_1 \overline{X}_\omega}$$