

21.

1°. doi uniforme sur $[0, \theta]$

$$\bullet L(x, \theta) = \prod_{i=1}^n f(x_i, \theta) = \frac{1}{\theta^n} \mathbb{1}_{[0, \theta]^n}(x)$$

$$= \frac{1}{\theta^n} \mathbb{1}_{[x_{(n)} \leq \theta]} \mathbb{1}_{[x_{(1)} \geq 0]} \text{ avec}$$

$$x_{(n)} = \max_{i=1}^n x_i \text{ et } x_{(1)} = \min_{i=1}^n x_i$$

$\Rightarrow L(x, \theta) = g(T(x), \theta) \times h(x)$ en posant

$$h(x) = \mathbb{1}_{[x_{(1)} \geq 0]}, \quad g(T, \theta) = \frac{1}{\theta^n} \mathbb{1}_{[x_{(n)} \leq \theta]}$$

$$\text{et } T(x) = \max x_i$$

D'après le théorème de factorisation de Neymann, $x_{(n)}$ est exhaustive pour θ .

• Le critère de minimalité s'applique. Calculons le rapport de vraisemblance $LR(x, y, \theta)$ de 2 échantillons

$$x = (x_1, \dots, x_n) \text{ et } y = (y_1, \dots, y_m)$$

$$LR(x, y, \theta) = \frac{L(x, \theta)}{L(y, \theta)}$$

$$= \mathbb{1}_{[0, \theta]}(x_{(n)}) \mathbb{1}_{[0, \theta]}(y_{(n)})$$

qui ne prend que 2 valeurs 0 ou 1 selon que $0 \leq x_{(n)} \leq \theta$ ou $0 \leq y_{(n)} \leq \theta$

$LR(x, y, \theta)$ vaut 1 $\forall \theta$ ssi $x_{(n)} = y_{(n)}$

car si les 2 valeurs étaient $\neq$ il suffirait de choisir θ entre les deux pour avoir $LR(x, y, \theta) = 0$

On a donc : $LR(x, y, \theta) = cte \neq 0$

ssi $x_{(n)} = y_{(n)}$

$\Rightarrow x_{(n)}$ est exhaustive minimale

• complétude : le modèle n'est pas exponentiel (il n'est pas homogène car le support de x dépend de θ).

Un calcul simple que nous avons souvent fait montre que la densité de $x_{(n)}$ est donnée par

$$g(x) = \frac{n}{\theta^n} x^{n-1} \mathbb{1}_{[0, \theta]}(x)$$

Soit h mesurable / $E_\theta [h(x_{(n)})] = 0$

$$\Rightarrow \int_0^\theta \frac{n}{\theta^n} x^{n-1} h(x) dx = 0 \quad \forall \theta$$

$$\Rightarrow \int_0^{\theta'} x^{n-1} h(x) dx = 0 \quad \forall \theta, \forall \theta'$$

$\Rightarrow h(x) \equiv 0$ p.s. d'après les propriétés de l'intégrale de Lebesgue $\Rightarrow x_{(n)}$ complète

• On sait que $x_{(n)}$ est également l'estimateur du max de vraisemblance de θ (e.m.v) et que

$$E[x_{(n)}] = \frac{n\theta}{n+1} \text{ donc } T = \frac{n+1}{n} x_{(n)}$$

est un estimateur sans biais (ESB) de θ ; comme $x_{(n)}$ est complète,

d'après le théorème de Lehmann-Scheffé, $E[T/x_{(n)}] = T$ est un

ESB préférable à T . Il est efficace s'il est sans biais (c'est le cas!)

et s'il atteint la borne F.D.R.

Ce n'est pas le cas ici, car comme le modèle n'est pas régulier, l'information de Fisher n'existe pas. $\Rightarrow$ non efficace.

2°. doi exponentielle $\mathcal{E}(\theta)$

$$\bullet L(x, \theta) = \frac{1}{\theta^n} e^{-s/\theta} \text{ où } s = \sum_{i=1}^n x_i$$

$$= 1 \times \frac{1}{\theta^n} \exp(-s/\theta)$$

$$g(x) \quad h(s, \theta)$$

le théorème de factorisation de Neymann montre alors que s est exhaustive pour θ .

• On peut montrer la minimalité en calculant le LR ou bien comme conséquence de la complétude

• complétude : on peut la montrer de deux façons : ou bien revenir à la définition ou bien utiliser le fait que la loi de x appartient au modèle exponentiel. la seconde méthode est ici la plus simple :

La décomposition de Neymann montre que la stat canonique est S et la fonction de partition $d(\theta) = -1/\eta$ avec $\theta > 0$.

$d:]0, \infty[\rightarrow]0, \infty[$ qui contient un ouvert (par ex $]0, \infty[$ lui-même).

Ainsi S est complète

3°. Modèle d'échantillonnage de Poisson

de modèle est donné par

$$(IN^m; \mathcal{P}(IN^m); \{ \mathcal{P}(N)^{\otimes m}; N > 0 \})$$

Il est homogène, paramétrique, exponentiel et identifiable.

La vraisemblance d'un échantillon $x = (x_1, \dots, x_n)$ est :

$$L(x, d) = \prod_{i=1}^m \mathbb{P}[X_i = x_i] = e^{-mN} \frac{N^d}{\prod_{i=1}^m x_i!}$$

$$d = \sum_{i=1}^m x_i$$

$$L(x, d) = \exp(-mN + d \cdot \ln N - \sum_{i=1}^m \ln x_i!)$$

qui est de la forme

$$h(x) \exp(\langle \psi(N), T(x) \rangle - \beta(N)) \text{ avec } \beta(N) = -mN, h(x) = (\prod x_i!)^{-1}$$

$\psi(N) = \ln N$ qui est la fonc. de partition et $T(x) = S(x) = \sum_{i=1}^m x_i$ la statistique canonique.

• Exhaustivité, méthode 1

nous montrons que $L(x/S=s)$ ne dépend pas de N .

$$\mathbb{P}[X=x/S=s] = \frac{\mathbb{P}([X=x] \cap [S=s])}{\mathbb{P}[S=s]}$$

$$= 0 \text{ si } s \neq \sum_{i=1}^m x_i \text{ et sinon } \frac{\mathbb{P}[S=s]}{\mathbb{P}[S=s]}$$

$$= e^{-mN} \frac{N^d}{\prod x_i!} \cdot \left[e^{-mN} \frac{(mN)^d}{d!} \right]^{-1} \mathbb{1}_{[s=\sum x_i]}$$

car $S \sim \mathcal{P}(mN)$

$$\mathbb{P}[X=x/S=s] = \frac{s! m^{-s}}{\prod x_i!} \mathbb{1}_{[s=\sum x_i]}$$

et cette formule ne dépend plus de N

$\Rightarrow S$ est exhaustive pour N

• méthode 2: théorème de factorisation

$$L(x, d) = h(x) g(x, N) \quad h \geq 0$$

$$L(x, d) = \underbrace{e^{-mN} e^{d \ln N}}_{g(x, N)} \underbrace{\frac{1}{\prod_{i=1}^m x_i!}}_{h(x)}$$

ainsi S est exhaustive d'après le th. de factorisation de Neymann.

• méthode 3: 1^{ère} exponentielle

$$L(x, d) = h(x) \exp(d \ln N - mN)$$

$$\beta(N) = mN$$

$$d \cdot \ln N = \langle \psi(N), T(x) \rangle$$

$\Rightarrow S$ est la statistique canonique du modèle, elle est donc exhaustive minimale.

• la minimalité peut aussi se démontrer en calculant $LR(x, y, d)$ rapport de vraisemblance de x et y

$$LR(x, y, d) = \text{constante} \Rightarrow S(x) = S(y)$$

• complétude, méthode 1

le modèle est exponentiel et

$\Lambda = \{ \psi(N); N > 0 \} \subset \mathbb{R}$ contient un ouvert; ici $\Lambda =]0, \infty[$ est lui-même ouvert.

• complétude, méthode 2

Soit g , fonc. mesurable de $\mathbb{N}$ dans $\mathbb{R}$. On a

$$\mathbb{E}_N[g(S)] = \sum_{k \geq 0} g(k) e^{-mN} \frac{(mN)^k}{k!}$$

il s'agit d'une série entière en N qui est nulle $\forall N$ si et seulement si tous les coefficients sont nuls.

comme e^{-mN} et $k!$ sont $> 0 \Rightarrow$

$$g(k) = 0 \quad \forall k \Rightarrow g \equiv 0$$

Ainsi, S est complète pour N

4°. Modèle d'échantillonnage gaussien

le paramètre $\theta = (m, \sigma^2)$ est ici de dimension 2

Il s'agit d'un modèle exponentiel et la factorisation de Neymann donnera donc la statistique canonique, qui est exhaustive et minimale.

$$f(x, \theta) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{x_i - m}{\sigma}\right)^2\right)$$

notons $h(x) = (2\pi)^{-n/2}$

$$f(x) = h(x) \exp\left(-\frac{n}{2} \ln \sigma - \frac{1}{2\sigma^2} \left(\sum_{i=1}^n x_i^2 - 2m \sum_{i=1}^n x_i + nm^2\right)\right)$$

l'intérieur de l'exponentielle peut se mettre sous la forme :

$$-\frac{n}{2} \ln \sigma - \frac{n}{2\sigma^2} (\bar{x}^2 - 2m\bar{x} + m^2)$$

$$= -\beta(\theta) + \langle T(x), \eta(\theta) \rangle \text{ avec}$$

$$\beta(\theta) = \frac{n}{2} \ln \sigma + \frac{nm^2}{2\sigma^2}, \quad T(x) = (\bar{x}, \bar{x}^2)$$

$$\eta(\theta) = \left(\frac{nm}{\sigma^2}, -\frac{n}{2\sigma^2}\right)$$

$\Rightarrow T(x)$ est exhaustive minimale

• complétude

le modèle est exponentiel et

$$\eta : \mathbb{R} \times]0, \infty[\rightarrow \mathbb{R} \times]-\infty, 0[\subset \mathbb{R}^2$$

$$\theta(m, \sigma^2) \mapsto \eta(\theta)$$

Λ , en image de $\mathbb{R} \times]0, \infty[$ par η contient comme ouvert de $\mathbb{R}^2$ lui-même $\Rightarrow (\bar{x}, \bar{x}^2)$ complète

5° Modèle de densité

$$f(x) = \frac{\partial e^{\theta x}}{e^{\theta^2} - 1} \mathbb{1}_{[0, \theta]}(x)$$

$$L(x, \theta) = \left(\frac{\theta}{e^{\theta^2} - 1}\right)^n e^{\theta s} \mathbb{1}_{[x_{(1)} \geq 0]} \mathbb{1}_{[x_{(n)} \leq \theta]}$$

avec $s = \sum x_i$

$(\sum_{i=1}^n x_i, \sup_{i=1}^n x_i)$ est donc exhaustive

mais par contre $\sum x_i$ seule ne l'est pas, d'après le théorème de factorisation.