Statistics

Martin Azpillaga

February 23, 2014

6

Block I

Definitions

1. Statistic models

introduction

Statistical model

Let:

- Ω set
- $\mathcal{A}$ sigma-algebra over Ω
- $\mathcal{P} \subset \mathcal{F}(\Omega, \mathbb{R})$

Then, $(\Omega, \mathcal{A}, \mathcal{P})$ is a statistical model if:

- $\forall\ f \in \mathcal{P}$:

  f probability distribution

Parametrized

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ statistical model
- $\Theta \subset \mathbb{R}^d$

Then, $(\Omega, \mathcal{A}, \mathcal{P})$ is parametrized by Θif:

- $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta}$

Likelihood

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ statistical model parametrized by Θ

We name likelihood function of $(\Omega, \mathcal{A}, \mathcal{P})$ to:

- $$\begin{array}{rcl} L : \Omega \times \Theta & \longrightarrow & \mathbb{R}^+ \\ (x, \theta) & \longmapsto & P_\theta(x) \end{array}$$

Exponential model

Let:

· $(\Omega, \mathcal{A}, \mathcal{P})$ statistical model parametrized by Θ

· L likelihood function of $(\Omega, \mathcal{A}, \mathcal{P})$

Then, $(\Omega, \mathcal{A}, \mathcal{P})$ is exponential if:

· $\exists \, \{f_i\}_{i=1}^{r}, f \in \mathcal{F}(\Omega, \mathbb{R})$ such that:

$\forall \; i \in [1, r]_{\mathbb{N}}$:

f_i, f measurable

· $\exists \, \{\phi_i\}_{i=1}^{r}, \phi \in \mathcal{F}(\theta, \mathbb{R})$ such that:

$\forall \; i \in [1, r]_{\mathbb{N}}$:

ϕ_i, ϕ measurable

· $\forall \; x \in \Omega$:

$\forall \; \theta \in \Theta$:

$$L(x, \theta) = \exp(\sum_{i=1}^{r} f_i(x)\phi_i(\theta) + f(x) + \phi(\theta))$$

2. Estimation

Statistic

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ m-D statistical model parametrized by Θ
- (M, Σ) measurable space
- $X : \Omega \to M$ random variable
- $f : M \to \mathbb{R}^m$

Then, f is a statistic if:

- f measurable

We denote:

- $f : T$

Sufficiency

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ m-D statistical model parametrized by Θ
- (M, Σ) measurable space
- $X : \Omega \to M$ random variable
- $T : M \to \mathbb{R}^m$ statistic

Then, T is sufficient if:

- $\forall\ \theta_1, \theta_2 \in \Theta$:

  $\forall\ x \in M$:

  $\forall\ t \in \mathbb{R}^m$:

$$P_{\theta_1}(X = x \mid T \circ X = t) = P_{\theta_2}(X = x \mid T \circ X = t)$$

Estimator

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ statistical model parametrized by Θ
- (M, Σ) measurable space
- $X : \Omega \to M$ random variable
- $T : M \to \mathbb{R}^k$ statistic
- $\theta \in \Theta$

Then, T is an estimator of θif:

- T approaches θ

Loss function

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ statistical model parametrized by Θ
- (M, Σ) measurable space
- $X : \Omega \to M$ random variable
- $T : M \to \mathbb{R}^k$ estimator
- $W : \mathbb{R}^k \times \Theta \to \mathbb{R}^+$

Then, W is a loss function if:

- $W(\theta, \theta) = 0$

Risk function

Let:

- $(\Omega, \mathcal{A}, \mathcal{P})$ statistical model parametrized by Θ
- (M, Σ) measurable space
- $X : \Omega \to M$ random variable
- $T : M \to \mathbb{R}^k$ estimator
- $W : \mathbb{R}^k \times \Theta \to \mathbb{R}^+$ loss function

We name risk function to:

$$\cdot \quad \begin{array}{rcl} R_T : \Theta & \longrightarrow & \mathbb{R}^+ \\ \theta & \longmapsto & E_\theta(W(T, \theta)) \end{array}$$

Block II

Propositions

1. Statistic models

introduction

Characterization of regular exponential models

Let:

· $(\Omega, \mathcal{A}, \mathcal{P})$ 1-D exponential model parametrized by Θ

· $\exp(f(x)\phi(\theta) + f'(x) + \phi'(\theta))$ likelihood function of $(\Omega, \mathcal{A}, \mathcal{P})$

statistical model

Then, holds:

· Θ interval , $\phi, \phi' \in \mathcal{C}^2$

· $\forall\ \theta \in \Theta$:

$\phi'(\theta) \neq 0$

$E_\theta f^2(x) \in \mathbb{R}$

· $\rightarrow (\Omega, \mathcal{A}, \mathcal{P})$ regular

Demonstration:

no demonstration

2. Statistic models

introduction

Characterization of regular exponential models

Let:

· $(\Omega, \mathcal{A}, \mathcal{P})$ 1-D exponential model parametrized by Θ

· $\exp(f(x)\phi(\theta) + f'(x) + \phi'(\theta))$ likelihood function of $(\Omega, \mathcal{A}, \mathcal{P})$

statistical model

Then, holds:

· Θ interval , $\phi, \phi' \in \mathcal{C}^2$

· $\forall\ \theta \in \Theta$:

$\phi'(\theta) \neq 0$

$E_\theta f^2(x) \in \mathbb{R}$

· $\rightarrow (\Omega, \mathcal{A}, \mathcal{P})$ regular

Demonstration:

no demonstration

18

Block III

Examples

1. Statistic models

introduction

Exponential model

Let:

- $\Omega : \mathbb{R}^n$
- $\mathcal{A} : \mathbb{B}(\mathbb{R}^n)$
- $\theta : (\mu, \sigma^2)$
- $\Theta : \mathbb{R} \times \mathbb{R}^+$
- $\mathcal{P} : \{N(\theta)\}_{\theta \in \Theta}$

Then, $(\Omega, \mathcal{A}, \mathcal{P})$ is an exponential model :

$\forall\ x \in \mathbb{R}^n$:

$$f_\theta(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{1}{2\sigma^2}(x_i - \mu)^2)$$

$$= (\tfrac{1}{\sqrt{2\pi\sigma^2}})^n \exp(-\tfrac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2)$$

$$L(x, \theta) = \exp(\tfrac{n}{2} \log(\tfrac{1}{2\pi\sigma^2}) - \tfrac{1}{2\sigma^2} \sum_{i=1}^{n} x_i^2 + \frac{n\bar{x}\mu}{\sigma^2} - \frac{n\mu^2}{2\sigma^2})$$

$$\begin{aligned} \phi\ : \Theta &\longrightarrow \mathbb{R}^2 \\ (\mu, \sigma^2) &\longmapsto (-\tfrac{\mu}{2\sigma^2}, \tfrac{n\mu}{\sigma^2}) \\ \phi'\ : \Theta &\longrightarrow \mathbb{R} \\ (\mu, \sigma^2) &\longmapsto \tfrac{n}{2}\log(\tfrac{1}{2\pi\sigma^2}) - \tfrac{n\mu^2}{2\sigma^2} \\ f\ : \Omega &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto (\sum_{i=1}^{n} x_i^2, \bar{x}) \end{aligned}$$

$$L(x, \theta) = \exp(\phi'(\theta) - \phi(\theta) f(x))$$

$\therefore)$ $(\Omega, \mathcal{A}, \mathcal{P})$ exponential model

Block IV

Problems

1. Statistic models

introduction

go

Block V

Tasks