1. Statistical models & Statistics

introduction

Statistical model

Let:

- Ω set
- $\mathcal{A}$ sigma-algebra over Ω
- $\mathcal{P} \subset \mathcal{F}(\Omega, [0,1])$

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

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

  $(\Omega, \mathcal{A}, P)$ probability space

Parametrization

Let:

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

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

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

We denote:

- $(\Omega, \mathcal{A}, \mathcal{P})\ :\ (\Omega, \mathcal{A}, \{P_\theta\}_{\theta \in \Theta})$ d-D statistical model

Likelihood function

Let:

- $(\Omega, \mathcal{A}, \{P_\theta\}_{\theta \in \Theta})$ d-D statistical model

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

- $\begin{array}{rccc} L6 : & \Omega \times \Theta & \longrightarrow & [0,1] \\ & (x, \theta) & \longmapsto & P_\theta(x) \end{array}$

Exponential model

Let:

- $(\Omega, \mathcal{A}, \{P_\theta\}_{\theta \in \Theta})$ d-D statistical model
- L likelihood function of $(\Omega, \mathcal{A}, \mathcal{P})$

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

- $\exists \ \{f_i\}_{i=0}^r \subset \mathcal{F}(\Omega, \mathbb{R})$:

  $\forall \ i \in [0, r]_{\mathbb{N}}$:

  f_i measurable

- $\exists \ \{\phi_i\}_{i=0}^r \subset \mathcal{F}(\Theta, \mathbb{R})$:

  $\forall \ x \in \Omega$:

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

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

Statistic

Let:

$\cdot\ (\Omega, \mathcal{A}, \{P_\theta\}_{\theta \in \Theta})$ d-D statistical model

$\cdot\ f\ :\ (\Omega, \mathcal{A}) \to (\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$

Then, f is a statistic if:

$\cdot\ f$ measurable

Sufficiency

Let:

$\cdot\ (\Omega, \mathcal{A}, \{P_\theta\}_{\theta \in \Theta})$ d-D statistical model

$\cdot\ (M, \Sigma)$ measurable space

$\cdot\ X\ :\ M \to \Omega$ random variable

$\cdot\ T\ :\ \Omega \to \mathbb{R}^m$ statistic

Then, T is sufficient if:

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

$\forall\ x \in M\ :$

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

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