Jornadas del Instituto de Cálculo

Un *cacho* de estadística

Marzo 2019

Quiénes somos

- Victor Yohai
- Marina Valdora
- Maria Eugenia (Maru) Szretter - Nostre
- Ezequiel Smucler
- Daniela Rodriguez
- Julieta Molina
- Gabriel Martos
- Juan Domingo Gonzalez
- Manuel Benjamin

Quiénes somos

- Victor Yohai
- Marina Valdora
- Maria Eugenia (Maru) Szretter - Nostre
- Ezequiel Smucler
- Daniela Rodriguez
- Julieta Molina
- Gabriel Martos
- Juan Domingo Gonzalez
- Manuel Benjamin

- Graciela Boente
- Ana Bianco
- ...

Qué hacemos

- Robustez
- Modelos de Regresión
- Alta dimensión - Penalización.
- Reducción (suficiente) de dimensión y predicción
- Datos Funcionales
- Colaboraciones
- Docencia
- Gestión

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).
- Alta dimensión. Penalización. p-value($\widehat{\beta}_j$) $< -$ $\tilde{\beta}_j = 0$

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).
- Alta dimensión. Penalización. p-value($\widehat{\beta}_j$) $<$ $-$ $\quad \tilde{\beta}_j = 0$
- Reducción (suficiente) de dimensión y predicción: hallar $R : \mathbb{R}^p \to \mathbb{R}^q$, $q < p$, tal que $y \mid \mathbf{x} \sim y \mid R(\mathbf{x})$,

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).
- Alta dimensión. Penalización. p-value($\widehat{\beta}_j$) $<$ – $\tilde{\beta}_j = 0$
- Reducción (suficiente) de dimensión y predicción: hallar $R : \mathbb{R}^p \to \mathbb{R}^q$, $q < p$, tal que $y \mid \mathbf{x} \sim y \mid R(\mathbf{x})$,
- Datos Funcionales

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).
- Alta dimensión. Penalización. p-value($\widehat{\beta}_j$) $<$ $-$ $\quad \tilde{\beta}_j = 0$
- Reducción (suficiente) de dimensión y predicción: hallar $R : \mathbb{R}^p \to \mathbb{R}^q$, $q < p$, tal que $y \mid \mathbf{x} \sim y \mid R(\mathbf{x})$,
- Datos Funcionales
- Colaboraciones

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).
- Alta dimensión. Penalización. p-value($\widehat{\beta}_j$) $<$ $-$ $\quad \tilde{\beta}_j = 0$
- Reducción (suficiente) de dimensión y predicción: hallar $R : \mathbb{R}^p \to \mathbb{R}^q$, $q < p$, tal que $y \mid \mathbf{x} \sim y \mid R(\mathbf{x})$,
- Datos Funcionales
- Colaboraciones
- Docencia

Qué hacemos

- Robustez: procedimientos que no se ven *muy* afectados ante la presencia de datos atípicos
- Modelos de Regresión: estudiar la relación entre variables explicativas ($\mathbf{x} \in \mathbb{R}^p$) y variable respuesta ($y \in \mathbb{R}, \mathbb{R}^k$).
- Alta dimensión. Penalización. p-value($\widehat{\beta}_j$) $<$ $-$ $\tilde{\beta}_j = 0$
- Reducción (suficiente) de dimensión y predicción: hallar $R : \mathbb{R}^p \to \mathbb{R}^q$, $q < p$, tal que $y \mid \mathbf{x} \sim y \mid R(\mathbf{x})$,
- Datos Funcionales
- Colaboraciones
- Docencia
- Gestión

Modelos de Regresión

$$y = \mathbf{x}^t \beta + \varepsilon$$

$$y = r(\mathbf{x}) + \varepsilon$$

$$y = \mathbf{x}^{\mathbf{t}} \beta + \eta(t) + \varepsilon$$

$$y = r(\mathbf{x}) + \sigma(\mathbf{x})\varepsilon$$

Modelos de Regresión

$$y = \mathbf{x}^t \beta + \varepsilon$$

$$y = r(\mathbf{x}) + \varepsilon$$

$$y = \mathbf{x}^\mathbf{t} \beta + \eta(t) + \varepsilon$$

$$y = r(\mathbf{x}) + \sigma(\mathbf{x})\varepsilon$$

$$\varepsilon \sim \mathcal{N}(0, \sigma^2)$$

Modelos de Regresión

$$y \mid \mathbf{x} \sim \mathcal{N}(\mu(\mathbf{x}), \sigma^2(\mathbf{x}))$$

$$y = \mathbf{x}^t\beta + \varepsilon \ , \quad \mu(\mathbf{x}) = \mathbf{x}^t\beta \ , \sigma^2(\mathbf{x}) = \sigma^2$$

$$y = \mathbf{x^t}\beta + \eta(t) + \varepsilon \ , \quad \mu(\mathbf{x}, t) = \mathbf{x^t}\beta + \eta(t) \ , \sigma^2(\mathbf{x}, t) = \sigma^2$$

Otras Distribuciones

- Gaussiana $\mathcal{N}(\mu, \sigma^2)$
- Bernoulli $\mathcal{B}(1, p)$ $(\theta = \log(p/(1-p)))$
- Poisson $\mathcal{P}(\lambda)$ $(\theta = \log(\lambda))$
- Exponencial
- Gamma
- Gumbel - Weibull (extreme value distributions)

$$Y \sim F_\theta \ , \quad \theta \in \Theta \subseteq \mathbb{R}^k$$

Modelos Generalizados de Regresión

$$y \mid \mathbf{x} \sim F_{\theta(\mathbf{x})}$$

$$\theta(\mathbf{x}) = \mathbf{x}^t \beta$$

$$\theta(\mathbf{x}) = \mathbf{x}^{\mathbf{t}} \beta + \eta(t)$$

Datos Funcionales

