

Modelación Estadística

Charlas de Acercamiento ICMAT



Departamento de Matemática
UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA



Estructura de la presentación

0. Presentación del grupo de Estadística.
1. Esquema de modelamiento estadístico.
2. Motivación mediante dos problemas.
3. Posibles tópicos de investigación.
4. Alumnos y ex-alumnos asociados a la línea.
5. Algunos comentarios finales.



Modelación Estadística: Académicos



(a) Alfredo Alegría



(b) Felipe Osorio



(c) Ronny Vallejos



Modelación Estadística: Alfredo Alegría¹



- ▶ Ph.D. Mathematics, PUCV-UTFSM-UV, Chile (supervisor: Emilio Porcu).
- ▶ Áreas de investigación:
**Multivariate spatial statistics,
Geostatistics for large datasets,
Non-gaussian random fields.**

- ▶ **Red de Colaboración:** Emilio Porcu, Peter Diggle (Reino Unido), Jorge Mateu (España), Reinhard Furrer (Suiza), Stefano Castruccio (USA), Moreno Bevilacqua, Xavier Emery (Chile).
- ▶ **Trabajos relevantes:** Computational Statistics & Data Analysis, Electronic Journal of Statistics, Environmetrics, International Statistical Review, Journal of Multivariate Analysis, SIAM Journal on Scientific Computing, Spatial Statistics, Statistics and Computing.
- ▶ Proyectos FONDECYT.
- ▶ Postdoctorado en la Universidad de Newcastle, UK.

¹Página web: <https://sites.google.com/site/alfredoalegrijimenez/>



Modelación Estadística: Ronny Vallejos²



- ▶ Ph.D. Statistics, University of Maryland Baltimore County, USA. (supervisor: Andrew L. Rukhin).
- ▶ Áreas de investigación:
Spatial statistics,
Robust modelling,
Statistical image modelling,
Time series.

- ▶ **Red de Colaboración:** Daniel Griffith, Aaron Ellison (USA), Silvia Ojeda, Oscar Bustos (Argentina), Hannah Buckley, Bradley Case (New Zealand).
- ▶ **Trabajos relevantes:** Chance, Electronic Journal of Statistics, Journal of Mathematical Imaging and Vision, Journal of Statistical Planning and Inference, Natural Resource Modeling, Spatial Statistics, Stochastic Environmental Research and Risk Assessment.
- ▶ Proyectos FONDECYT, de cooperación internacional (CECYT, Math-AmSud), PIA.
- ▶ Ex Editor en Jefe de la revista *Chilean Journal of Statistics*, miembro de los centros AM2V y AC3E.

²Página web: <http://rvallejos.mat.utfsm.cl>



Modelación Estadística: Felipe Osorio³

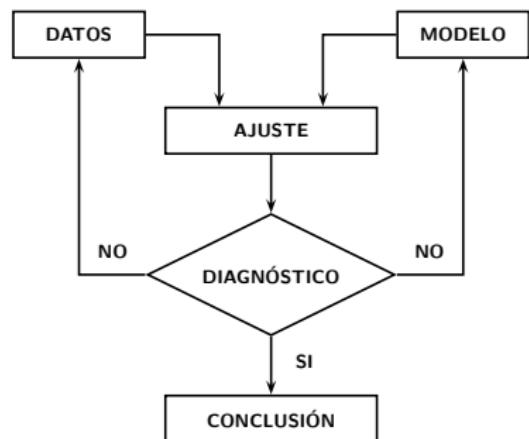


- ▶ D.Sc. Statistics, Universidade de São Paulo, Brasil.
(supervisor: Gilberto A. Paula).
- ▶ Áreas de investigación:
Modelos para datos longitudinales,
Diagnóstico de influencia,
Funciones de inferencia.
- ▶ Red de Colaboración: Gilberto A. Paula, Cibele Russo (Brasil), Manuel Galea (Chile), Federico Crudu (Italia).
- ▶ Trabajos relevantes: Annals of the Institute of Statistical Mathematics, Biometrical Journal, Computational Statistics & Data Analysis, Economics Letters, Signal Image and Video Processing, Spatial Statistics, Statistical Papers, Statistics and Computing.
- ▶ Proyectos FONDECYT, de cooperación internacional (PROSUL, CNPq).
- ▶ Creador de paquetes contribuídos a R (fastmatrix, heavy, L1pack, SpatialPack).
- ▶ Editor de la revista *Chilean Journal of Statistics*.

³Página web: <http://fosorios.mat.utfsm.cl/>



Esquema de Modelación Estadística



Recolección de datos: **Muestreo**.

Análisis exploratorio de datos.

Análisis Multivariado.

Técnicas de Regresión.

Series de Tiempo, entre (muchas) otras.

Inferencia Estadística.

Bondad de ajuste, técnicas gráficas.

Análisis de Sensibilidad.

Comunique sus resultados!



Estimación del SSIM

Trabajo:

Osorio, F., Vallejos, R., Barraza, W., Ojeda, S., Landi, M.A. (2021+). Statistical estimation of the structural similarity index for image quality assessment. *Signal, Image and Video Processing* (por aparecer).

Considere dos imágenes $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^N$, el coeficiente de similaridad estructural (SSIM) es dado por:

$$\text{SSIM}(\mathbf{x}, \mathbf{y}; \boldsymbol{\theta}) = l(\mathbf{x}, \mathbf{y})^\alpha \cdot c(\mathbf{x}, \mathbf{y})^\beta \cdot s(\mathbf{x}, \mathbf{y})^\gamma,$$

donde

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\bar{x}\bar{y} + c_1}{\bar{x}^2 + \bar{y}^2 + c_1}, \quad c(\mathbf{x}, \mathbf{y}) = \frac{2s_x s_y + c_2}{s_x^2 + s_y^2 + c_2},$$

$$s(\mathbf{x}, \mathbf{y}) = \frac{s_{xy} + c_3}{s_x s_y + c_3}.$$

Objetivo:

Basado en imágenes observadas \mathbf{x} e \mathbf{y} , estimar $\boldsymbol{\theta} = (\alpha, \beta, \gamma)^\top$ y probar la hipótesis:

$$H_0 : \alpha = \beta = \gamma = 1.$$



Estimación del SSIM

Se consideró un **modelo no-lineal heteroscedástico⁴** bajo el supuesto:

$$Z_i \sim N(\phi f_i(\boldsymbol{\theta}), f_i^2(\boldsymbol{\theta})g^2(\phi)), \quad i = 1, \dots, n,$$

donde

$$f_i(\boldsymbol{\theta}) = \text{SSIM}(\mathbf{x}_i, \mathbf{y}_i; \boldsymbol{\theta}), \quad g^2(\phi) = \phi^2(\phi^2 - 1),$$

corresponden a **funciones de media y de varianza** y $Z_i = 1/\text{RMSE}(\mathbf{x}_i, \mathbf{y}_i)$.

Resultados:

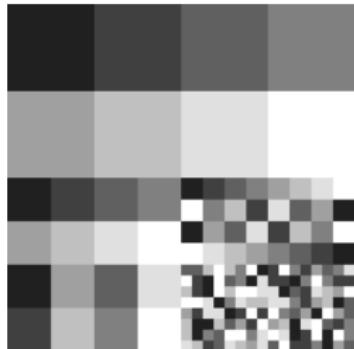
- ▶ **Algoritmo de estimación:** Híbrido entre método secante multivariado (BFGS) con **pseudo-verosimilitud** (método de Brent).⁵
- ▶ **Test de hipótesis** usando el **estadístico gradiente** (Terrell, 2000).
- ▶ **Matriz de información de Fisher** y método eficiente para evaluar **función score**.
- ▶ **Experimento numérico** con **datos sintéticos** (desde base de datos USC-SIPI) y **datos desde la constelación de satélites ICEYE SAR**.

⁴Inspirado en el contexto de **funciones de producción** (Econometría).

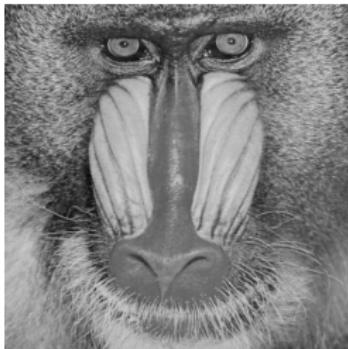
⁵Código C y R disponible en github.com/faosorios/SSIM.



Estimación del SSIM: estudio de simulación Monte Carlo⁸



(a) texmos2.S512



(b) Baboon



(c) Lenna

Imágenes de referencia⁶ ($\mathbf{x} \in \mathbb{R}_+^N$) fueron contaminadas con ruido multiplicativo⁷ usando una distribución Gamma(L, L), es decir,

$$\mathbf{y} = \mathbf{x} \cdot \mathbf{w}, \quad w_t \sim \text{Gamma}(L, L), \quad t = 1, \dots, N,$$

para $L = 1, 2, 4, 8, 16$ y 32 looks. Para cada look 1,000 imágenes fueron construidas.

⁶Extraídas desde la base de imágenes USC-SIPI, URL: <http://sipi.usc.edu/database>

⁷Disponible en la función `imnoise` desde el paquete R `SpatialPack`.

⁸Se ajustaron 144 000 modelos (tiempo: 34 hrs, 40 min, 16 seg).



Estimación del SSIM: contaminación de Lenna



(a) $L = 1$



(b) $L = 2$



(c) $L = 4$



(d) $L = 8$



(e) $L = 16$



(f) $L = 32$



Estimación del SSIM: imágenes de Radar de Apertura Sintética (SAR)⁹



(a) Copeland



(b) Dam



(c) Corpus Christi



(d) Mississippi

⁹ Constelación de satélites ICEYE SAR: <https://www.iceye.com/downloads/datasets>



Estimación del SSIM: resultados con imágenes SAR

Copeland (4096 × 2560)				Corpus Christi (4096 × 2560)		
Filter	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
Lee	1.891	1.932	3.136	1.610	1.639	2.705
Enhanced Lee	1.691	1.722	2.987	1.600	1.626	2.512
Kuan	1.691	1.722	2.987	1.600	1.626	2.512

Dam (1200 × 1200)				Mississippi (4096 × 4096)		
Filter	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$
Lee	1.000	1.000	1.000	1.591	1.623	2.670
Enhanced Lee	1.000	1.000	1.000	1.468	1.491	2.553
Kuan	1.000	1.000	1.000	1.468	1.491	2.553



Estimación del SSIM: resultados con imágenes SAR¹⁰

Copeland (4096 × 2560)		Corpus Christi (4096 × 2560)		
Filter	SSIM		SSIM	
	Under H_0	Under H_1	Under H_0	Under H_1
Lee	0.680	0.326	0.828	0.614
Enhanced Lee	0.652	0.306	0.818	0.615
Kuan	0.652	0.306	0.818	0.615

Dam (1200 × 1200)		Mississippi (4096 × 4096)		
Filter	SSIM		SSIM	
	Under H_0	Under H_1	Under H_0	Under H_1
Lee	0.999	0.999	0.890	0.740
Enhanced Lee	0.999	0.999	0.884	0.737
Kuan	0.999	0.999	0.884	0.737

¹⁰ Recuerde que $H_0 : \alpha = \beta = \gamma = 1$.



Influencia local para el análisis de acuerdo/concordancia

Suponga que deseamos evaluar el grado de acuerdo entre dos instrumentos de medición y sea $(X_{11}, X_{12}), \dots, (X_{n1}, X_{n2})$ observaciones bivariadas con vector de medias μ y matriz de covarianza Σ .

Considere

$$\rho_c = \frac{2\sigma_{12}}{\sigma_{11} + \sigma_{22} + (\mu_1 - \mu_2)^2}, \quad \psi_c = P(|X_{i1} - X_{i2}| \leq c), \quad c > 0.$$

el coeficiente de concordancia (Lin, 1989) y la probabilidad de acuerdo (Stevens et al., 2017), respectivamente. Bajo normalidad,

$$\psi_c = \Phi\left(\frac{c - \mu_D}{\sigma_D}\right) - \Phi\left(\frac{-c - \mu_D}{\sigma_D}\right),$$

con $\mu_D = \mu_1 - \mu_2$, $\sigma_D^2 = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$.



Influencia local para el análisis de acuerdo/concordancia

Trabajo:

Leal, C., Galea, M., Osorio, F. (2019).

Assessment of local influence for the analysis of agreement.

Biometrical Journal **61**, 955-972.

Objetivo:

Considerar funciones de influencia $f(\omega)$ y estudiar el comportamiento del modelo perturbado

$$\mathcal{P}_\omega = \{p(\mathbf{x}; \boldsymbol{\theta}, \omega) : \boldsymbol{\theta} \in \Theta, \omega \in \Omega\}.$$

con ω_0 vector de perturbación nula, tal que $\mathcal{P}_{\omega_0} = \mathcal{P}$, con

$$\mathcal{P} = \{\mathbf{N}_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})^{\otimes n} : \boldsymbol{\mu} \in \mathbb{R}^2, \boldsymbol{\Sigma} \geq 0\}.$$

Llevar a cabo influencia local de primer y segundo orden sobre $\rho_c(\omega)$ y $\psi_c(\omega)$.



Influencia local para el análisis de acuerdo/concordancia

Se obtuvo las **curvaturas normal** (Cook, 1986) y **conformal** (Poon y Poon, 1999)

$$C_{f,h} = \frac{\mathbf{h}^\top \mathbf{H}_f \mathbf{h}}{(1 + \nabla_f^\top \nabla_f) \mathbf{h}^\top (\mathbf{I} + \nabla_f \nabla_f^\top) \mathbf{h}},$$

$$B_{f,h} = \frac{\mathbf{h}^\top \mathbf{H}_f \mathbf{h}}{\|\mathbf{H}_f\|_M \mathbf{h}^\top (\mathbf{I} + \nabla_f \nabla_f^\top) \mathbf{h}},$$

y **medidas de influencia de primer y segundo orden** (Zhu et al., 2007)

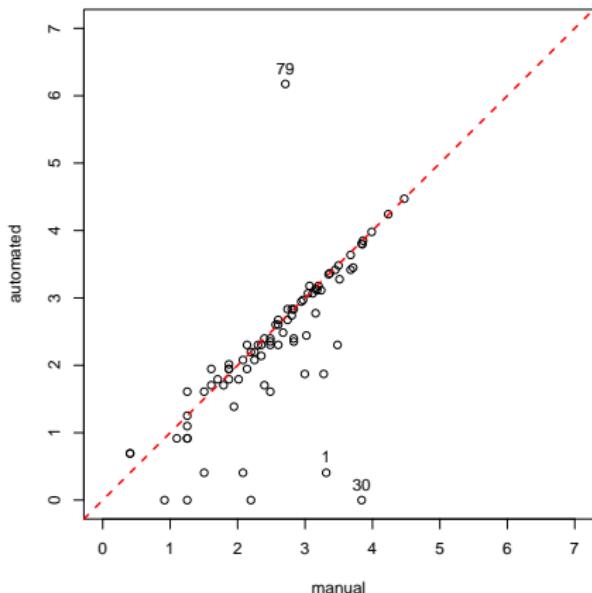
$$\text{FI}_{f,h} = \frac{\mathbf{h}^\top \nabla_f \nabla_f^\top \mathbf{h}}{\mathbf{h}^\top \mathbf{G}(\boldsymbol{\omega}_0) \mathbf{h}}, \quad \text{SI}_{f,h} = \frac{\mathbf{h}^\top \tilde{\mathbf{H}}_f \mathbf{h}}{\mathbf{h}^\top \mathbf{G}(\boldsymbol{\omega}_0) \mathbf{h}},$$

donde $\nabla_f = \partial f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} \big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$ y $\mathbf{H}_f = \partial^2 f(\boldsymbol{\omega}) / \partial \boldsymbol{\omega} \partial \boldsymbol{\omega}^\top \big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}$, con $\mathbf{G}(\boldsymbol{\omega})$ matriz de información de Fisher con relación a $\boldsymbol{\omega}$,

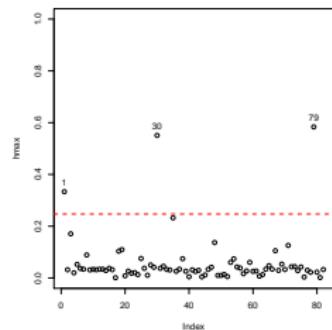
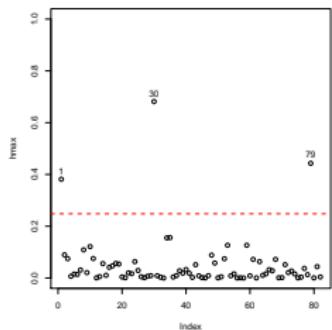
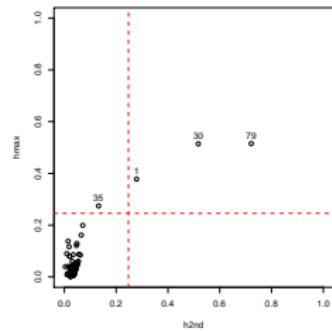
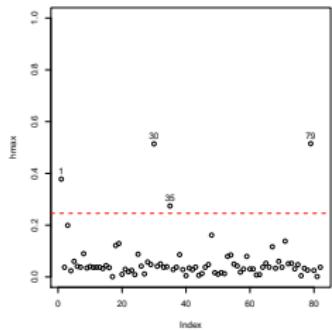
$$g_{ij}(\boldsymbol{\omega}) = \mathbb{E}_{\boldsymbol{\omega}} \left\{ \frac{\partial \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_i} \frac{\partial \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \omega_j} \right\}, \quad i, j = 1, \dots, n.$$



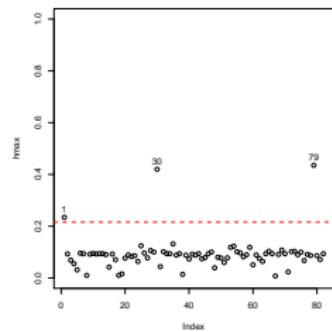
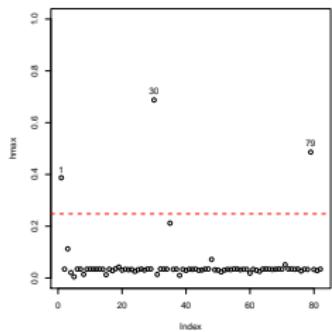
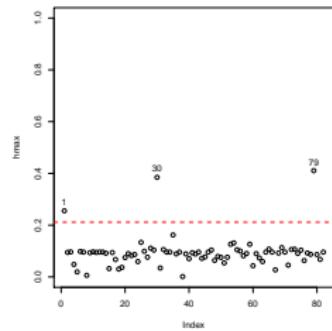
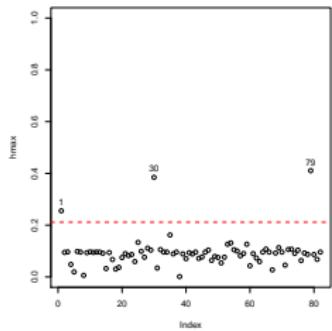
Datos de desorden transitorio de sueño (Svetnik et al., 2007)



Datos de desorden transitorio de sueño, objetivo: $\rho_c(\omega)$



Datos de desorden transitorio de sueño, objetivo: $\psi_c(\omega)$



Influencia local para el análisis de acuerdo/concordancia

Resultados:

- ▶ Influencia local de primer y segundo orden aplicado a funciones objetivo ρ_c y ψ_c .
- ▶ Experimentos numéricos: estudio de simulación Monte Carlo y aplicación a datos de la realidad sugieren que influencia de primer orden es más potente para la detección de observaciones atípicas.
- ▶ Se determinó que el esquema de perturbación de ponderación de casos corresponde a una perturbación apropiada.
- ▶ Código en R para la estimación y el diagnóstico de influencia se encuentra disponible en [github](#).¹¹
- ▶ Resultado lateral: distribución asintótica para la probabilidad de concordancia, ψ_c .

¹¹ URL: github.com/faosorios/CCC



Un método robusto para determinar la probabilidad de acuerdo

Trabajo:

Osorio, F., Galea, M.

Agreement assessment between two measurement systems using robust P-splines.

En preparación.

Objetivo:

Basado en las ideas de Bland and Altman (1999), se desea modelar la relación entre

$$Y_i = X_{i1} - X_{i2}, \quad z_i = (X_{i1} + X_{i2})/2, \quad i = 1, \dots, n,$$

usando un *procedimiento robusto* (Osorio, 2016) para *P-splines*,

$$Y_i = g(z_i) + \epsilon_i, \quad i = 1, \dots, n,$$

donde g es una función suave definida en $[a, b]$ y $\{\epsilon_i\}$ son disturbios aleatorios.



Un método robusto para determinar la probabilidad de acuerdo

Usaremos la notación $\text{SMN}(\mu, \phi; H)$ para indicar la clase de **distribuciones de mezcla de escala** (Andrews y Mallows, 1974), con posición $\mu \in \mathbb{R}$, escala $\phi > 0$ y variable mezcladora U con distribución H .

Una variable aleatoria $Y \sim \text{SMN}(\mu, \phi; H)$ tiene la siguiente representación:

$$Y|U \sim N(\mu, \phi/u) \quad \text{and} \quad U \sim H(\nu).$$

La función de densidad de Y asume la forma

$$f(y) = (2\pi\phi)^{-1/2} \int_0^\infty u^{1/2} \exp\left\{-\frac{1}{2}uD^2(\theta)\right\} dH(u),$$

donde $D^2(\theta) = (Y - \mu)^2/\phi$.

Observación:

Esta clase incluye distribuciones simétricas con **colas más pesadas que la normal**.



Un método robusto para determinar la probabilidad de acuerdo

Algunos ejemplos de distribuciones en la clase de SMN:

- **Student-*t*:** $Y \sim t(\mu, \phi; \nu)$, con densidad

$$f(y) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu)^{1/2}} \phi^{-1/2} \left(1 + D^2(\theta)/\nu\right)^{-\frac{\nu+1}{2}}, \quad \nu > 0.$$

y podemos escribir $Y|U \sim N(\mu, \phi/u)$ y $U \sim \text{Gamma}(\nu/2, \nu/2)$.

- **Slash:** $Y \sim \text{Slash}(\mu, \phi, \nu)$ en cuyo caso la densidad de Y es dada por

$$f(y) = \nu(2\pi\phi)^{-1/2} \int_0^1 u^{\nu-1/2} \exp\left(-\frac{1}{2}uD^2(\theta)\right) du, \quad \nu > 0.$$

Tenemos que $Y|U \sim N(\mu, \phi/u)$ y $U \sim \text{Beta}(\nu, 1)$.



Un método robusto para determinar la probabilidad de acuerdo

Asumiendo el modelo,

$$Y_i \stackrel{\text{ind}}{\sim} \text{SMN}(g(z_i), \phi, H), \quad i = 1, \dots, n. \quad (1)$$

lleva a la probabilidad de acuerdo,

$$\psi_i = P(|Y_i| \leq c) = F_H\left(\frac{c - g(z_i)}{\sqrt{\phi}}\right) - F_H\left(\frac{-c - g(z_i)}{\sqrt{\phi}}\right),$$

donde $F_H(\cdot)$ denota la función de distribución acumulada de una distribución de mezcla de escala normal estándar $\text{SMN}(0, 1; H)$.

Observación

El modelo puede ser escrito como:

$$Y_i \stackrel{\text{ind}}{\sim} \text{SMN}(\mathbf{b}_i^\top \mathbf{a}, \phi; H), \quad i = 1, \dots, n,$$

donde $\mathbf{b} = (B_1(z), \dots, B_p(z))^\top$, $\mathbf{a} = (a_1, \dots, a_p)^\top$, y p es el número de funciones base. En este caso $B_j(z)$ ($j = 1, \dots, p$) corresponde a B-splines.



Un método robusto para determinar la probabilidad de acuerdo

La esperanza condicional en la etapa-E del [algoritmo EM penalizado](#) adopta la forma:

$$Q_\lambda(\boldsymbol{\theta}; \widehat{\boldsymbol{\theta}}) = -\frac{n}{2} \log \phi - \frac{1}{2\phi} \{ (\mathbf{Y} - \mathbf{B}\mathbf{a})^\top \text{diag}(\widehat{\mathbf{U}})(\mathbf{Y} - \mathbf{B}\mathbf{a}) + \lambda \mathbf{a}^\top \mathbf{K}^\top \mathbf{K} \mathbf{a} \},$$

donde $\text{diag}(\widehat{\mathbf{U}}) = \text{diag}(\widehat{U}_1, \dots, \widehat{U}_n)$ con $\widehat{U}_i = E(U_i | Y_i, \boldsymbol{\theta}^{(k)})$ que debe ser evaluado para una elección específica de H .

Tenemos que $\widehat{\boldsymbol{\theta}} = (\widehat{\mathbf{a}}^\top, \widehat{\phi})^\top$ tiene distribución asintótica

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N_{p+1}(\mathbf{0}, \mathbf{V}),$$

en la práctica \mathbf{V} puede ser estimada usando el [método de Oakes \(1999\)](#) como

$$\frac{\partial^2 \ell_\lambda(\boldsymbol{\theta}; \mathbf{Y})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}} = \left\{ \frac{\partial^2 Q(\boldsymbol{\theta}; \widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} + \frac{\partial^2 Q(\boldsymbol{\theta}; \widehat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \widehat{\boldsymbol{\theta}}^\top} - \frac{\partial^2 J_\lambda(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right\} \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}},$$

donde $J_\lambda(\boldsymbol{\theta}) = \frac{\lambda}{2\phi} \|\mathbf{K}\mathbf{a}\|^2$.



Un método robusto para determinar la probabilidad de acuerdo

Resultado 1:

Para el modelo dado en la Ecuación (1) la matriz de información observada asociada con la función $Q_\lambda(\theta; \hat{\theta})$ evaluada en $\theta = \hat{\theta}$ asume la forma:

$$\begin{aligned} \frac{\partial^2 \ell_\lambda(\theta; \mathbf{Y})}{\partial \theta \partial \theta^\top} \Big|_{\theta=\hat{\theta}} &= -\frac{1}{\hat{\phi}} \begin{pmatrix} \mathbf{B}^\top \text{diag}(\hat{\mathbf{U}}) \mathbf{B} + \lambda \mathbf{K}^\top \mathbf{K} & \mathbf{0} \\ \mathbf{0} & n/(2\hat{\phi}) \end{pmatrix} \\ &\quad + \frac{1}{\hat{\phi}} \begin{pmatrix} (\partial \hat{\mathbf{U}} / \partial \hat{\mathbf{a}}^\top)^\top \text{diag}(\mathbf{e}) \mathbf{B} & (\partial \hat{\mathbf{U}} / \partial \hat{\mathbf{a}}^\top)^\top \text{diag}(\mathbf{e}) \mathbf{e} / (2\hat{\phi}) \\ (\partial \hat{\mathbf{U}} / \partial \hat{\phi})^\top \text{diag}(\mathbf{e}) \mathbf{B} & (\partial \hat{\mathbf{U}} / \partial \hat{\phi})^\top \text{diag}(\mathbf{e}) \mathbf{e} / (2\hat{\phi}) \end{pmatrix}, \end{aligned}$$

donde $\mathbf{e} = \mathbf{Y} - \hat{\mathbf{g}}(\mathbf{z})$.

Observación:

En general, la función de pesos \hat{U}_i es dada por

$$\hat{U}_i = \mathbb{E}(U_i | Y_i, \hat{\theta}) = \frac{\int_0^\infty u_i^{3/2} \exp(-\frac{1}{2} u_i D_i^2(\theta)) dH(u_i)}{\int_0^\infty u_i^{1/2} \exp(-\frac{1}{2} u_i D_i^2(\theta)) dH(u_i)} \Big|_{\theta=\hat{\theta}}$$



Un método robusto para determinar la probabilidad de acuerdo

Resultado 2:

Para la probabilidad de acuerdo asumiendo P-splines bajo la clase de mezclas de escala normal, tenemos

$$\sqrt{n}(\widehat{\psi}_i - \psi_i) \xrightarrow{D} N_1\left(0, \left(\frac{\partial \psi_i}{\partial \theta}\right)^T V \left(\frac{\partial \psi_i}{\partial \theta}\right)\right),$$

donde $\partial \psi_i / \partial \theta = (\partial \psi_i / \partial \mathbf{a}^T, \partial \psi_i / \partial \phi)^T$, con

$$\frac{\partial \psi_i}{\partial \mathbf{a}} = \frac{1}{\sqrt{\phi}} \left\{ f\left(\frac{-c - g(z_i)}{\sqrt{\phi}}\right) - f\left(\frac{c - g(z_i)}{\sqrt{\phi}}\right) \right\} \mathbf{b}_i,$$

$$\frac{\partial \psi_i}{\partial \phi} = -\frac{1}{2\phi} \left\{ \left(\frac{c - g(z_i)}{\sqrt{\phi}}\right) f\left(\frac{c - g(z_i)}{\sqrt{\phi}}\right) + \left(\frac{-c - g(z_i)}{\sqrt{\phi}}\right) f\left(\frac{-c - g(z_i)}{\sqrt{\phi}}\right) \right\},$$

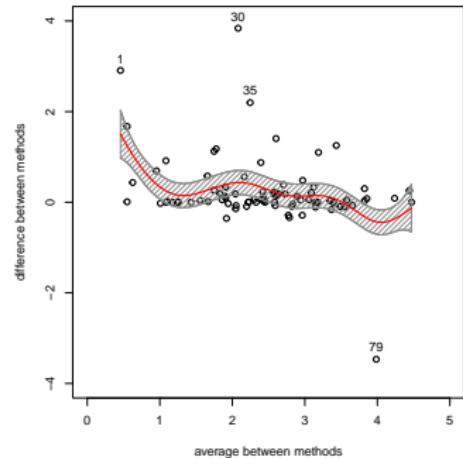
lo que debe ser evaluado en $\theta = \widehat{\theta}$.

Observación:

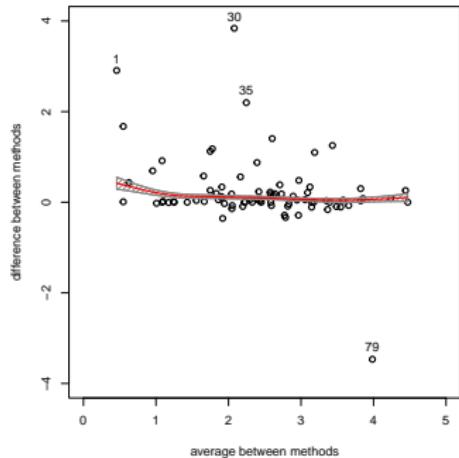
Esto nos permite construir bandas de confianza (BC) para la probabilidad de acuerdo.



Desorden transitorio de sueño: $\hat{g}(z)$ usando P-splines



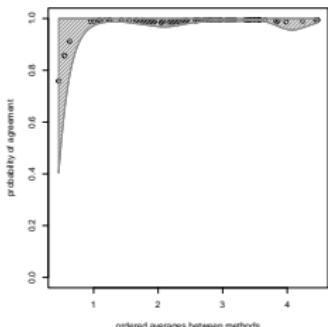
(a) normal



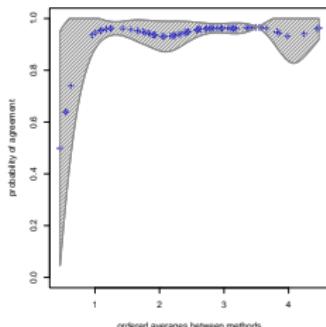
(b) Student- t , $\nu = 4$



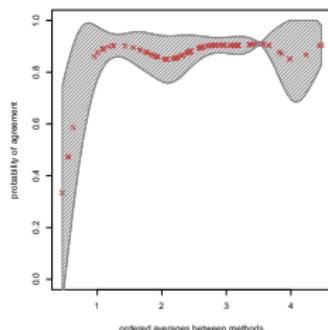
Desorden transitorio de sueño: PA estimada con BC bajo normalidad



(a) $c = 2.0$



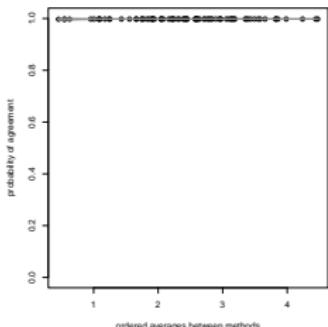
(b) $c = 1.5$



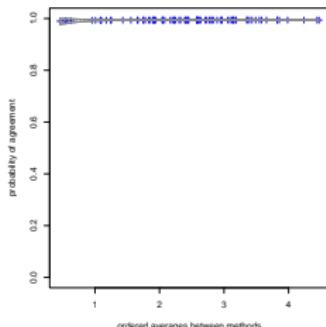
(c) $c = 1.2$



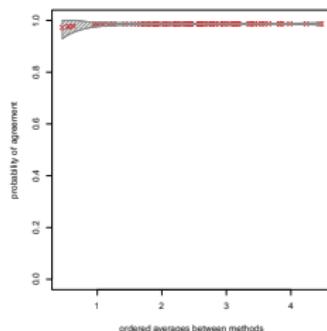
Desorden transitorio de sueño: PA estimada con BC bajo Student-*t*



(a) $c = 2.0$



(b) $c = 1.5$



(c) $c = 1.2$



Temas tentativos para tesis: Alfredo Alegría¹²



- ▶ Campos aleatorios para modelar datos espacio-temporales sobre grandes porciones del planeta.
- ▶ Funciones de covarianza cruzada flexibles y su aplicación en el análisis de datos espacio-temporales multivariados.
- ▶ Métodos de estimación y predicción para campos aleatorios no-gaussianos y/o para grandes conjuntos de datos.

¹²E-mail: alfredo.alegria@usm.cl



Temas tentativos para tesis: Ronny Vallejos¹³



- ▶ Desarrollo de medidas de concordancia para datos espacio-temporales.
- ▶ Extensión de la probabilidad de concordancia para datos espaciales.
- ▶ Estimación del coeficiente de similaridad entre imágenes.
- ▶ Tamaño muestral efectivo para datos espacio-temporales.

¹³E-mail: ronny.vallejos@usm.cl



Temas tentativos para tesis: Felipe Osorio¹⁴



- ▶ Influencia local para regresión LAD.
- ▶ Probabilidad de concordancia para varios instrumentos de medición.
- ▶ Diagnóstico de influencia en el contexto de estimación máximo Lq -verosímil.

¹⁴E-mail: felipe.osorios@usm.cl (no olvidar la 's' !)



Modelación Estadística: Estudiantes de pregrado

- ▶ Nicolás Alfaro.
- ▶ Edgard González.
- ▶ Pablo Huenchulao.
- ▶ Manuel Jara.
- ▶ Marcela Miranda.
- ▶ Eric Muñoz.
- ▶ Bastián Sepúlveda.
- ▶ Sebastián Vera.
- ▶ Gabriel Vidal (2021).
- ▶ Gabriel Molina (2020).
- ▶ Alberto Rubio (2019).
- ▶ Alexis Tapia (2019).
- ▶ Alonso Ogueda (2018).
- ▶ Ignacio Vásquez (2017).
- ▶ Francisco Alfaro (2017).
- ▶ Wilson Barraza (2017).
- ▶ Javier Pérez (2017).
- ▶ Ángelo Gárate (2016).
- ▶ Carlos Schwarzenberg (2016).
- ▶ Jean Paul Maidana (2015).
- ▶ Agustín Uribe (2015).
- ▶ Alfredo Alegría (2014).
- ▶ Claudio Henríquez (2014).
- ▶ Consuelo Moreno (2014).
- ▶ Jonathan Acosta (2013).
- ▶ Francisco Cuevas (2011).
- ▶ Danilo Pezo (2011).
- ▶ Jorge Littin (2006).



Modelación Estadística: Estudiantes de postgrado

- ▶ John Gómez.
- ▶ Fabian Ramírez.
- ▶ Alonso Ogueda (2021).
- ▶ Carlos Schwarzenberg (2021).
- ▶ Francisco Alfaro (2019).
- ▶ Javier Perez (2019).
- ▶ Sebastián Torres (2019).
- ▶ Ángelo Garate (2018).
- ▶ Jonathan Acosta (2017).
- ▶ Diego Mancilla (2014).
- ▶ Francisco Cuevas (2013).
- ▶ Danilo Pezo (2013).
- ▶ Paola Carvajal (2006).
- ▶ Ronny Vallejos (1999).



Algunos trabajos recientes

-  **Acosta, J., Alegría, A., Osorio, F., Vallejos, R. (2021).**
Assessing the effective sample size for large spatial datasets: A block likelihood approach.
Computational Statistics & Data Analysis 162, 107282.
-  **Alegría, A., Bissiri, P.G., Cleanthous, G., Porcu, E., White, P. (2021).**
Multivariate isotropic random fields on spheres: Nonparametric bayesian modeling and L^p fast approximations.
Electronic Journal of Statistics 15, 2360-2392.
-  **Alegría, A., Cuevas, F., Diggle, P., Porcu, E. (2021).**
The \mathcal{F} -family of covariance functions: A Matérn analogue for modeling random fields on spheres.
Spatial Statistics 43, 100512.
-  **Alegría, A., Emery, X., Porcu, E. (2021).**
Bivariate Matérn covariances with cross-dimple for modeling coregionalized variables.
Spatial Statistics 41, 100491.
-  **Emery, X., Alegría, A., Arroyo, D. (2021).**
Covariance models and simulation algorithm for stationary vector random fields on spheres crossed with Euclidean spaces.
SIAM Journal on Scientific Computing 43, A3114-A3134.



Algunos trabajos recientes

-  **Osorio, F., Vallejos, R., Barraza, W., Ojeda, S., Landi, M. (2021+).**
Statistical estimation of the structural similarity index for image quality assessment.
Signal, Image and Video Processing (to appear).
-  **Vallejos, R., Acosta, J. (2021+).**
Effective sample size for multivariate spatial processes with an application to soil contamination.
Natural Resource Modeling (to appear).
-  **Alegria, A. (2020).**
Cross-dimple in the cross-covariance functions of bivariate isotropic random fields on spheres.
Stat 9, e301.
-  **Alegria, A., Cuevas, F. (2020).**
Karhunen-Loeve expansions for axially symmetric gaussian processes: Modeling strategies and L^2 approximations.
Stochastic Environmental Research and Risk Assessment 34, 1953-1965.
-  **Alegria, A., Emery, X., Lantuéjoul, C. (2020).**
The turning arcs: A computationally efficient algorithm to simulate isotropic vector-valued gaussian random fields on the d-sphere.
Statistics and Computing 30, 1403-1418.

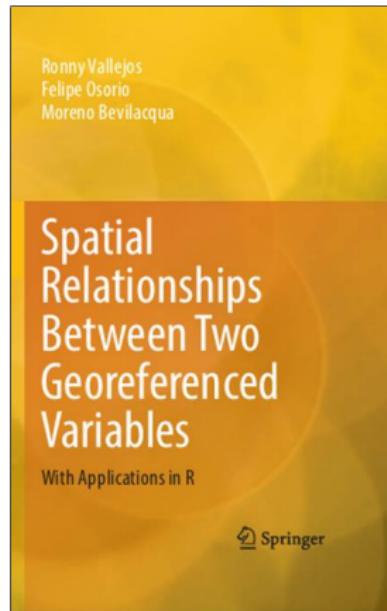


Algunos trabajos recientes

-  **Crudu, F., Osorio, F. (2020).**
Bilinear form test statistics for extremum estimation.
Economics Letters 187, 108885.
-  **Emery, X., Alegria, A. (2020).**
A spectral algorithm to simulate nonstationary random fields on spheres and multifractal star-shaped random sets.
Stochastic Environmental Research and Risk Assessment 34, 2301-2311.
-  **Vallejos, R., Pérez, J., Ellison, A., Richardson, A. (2020).**
A spatial concordance correlation coefficient with an application to image analysis.
Spatial Statistics 40, 100405.
-  **Vallejos, R., Osorio, F., Bevilacqua, M. (2020).**
Spatial Relationships Between Two Georeferenced Variables: With Applications in R.
Springer, Cham.



Asociación espacial entre dos procesos estocásticos (libro Springer)



- ▶ Ronny Vallejos, Felipe Osorio (USM) y Moreno Bevilacqua (UAI).
- ▶ Asociación entre dos procesos espaciales:
 - procedimientos de test de hipótesis.
 - coeficientes de asociación/codispersión.
 - asociación entre imágenes.
- ▶ Paquetes en R: [SpatialPack](#) y [GeoModels](#).
- ▶ Publicado por [Springer](#) (2020).



Cursos de especialidad: Modelación Estadística

Especialidad:

- ▶ Análisis de regresión.
- ▶ Análisis multivariado.
- ▶ Series de tiempo.

Magíster:

- ▶ Estadística espacial.
- ▶ Modelos lineales generalizados.
- ▶ Series de tiempo avanzadas.
- ▶ Simulación estocástica.

Otros cursos:

- ▶ Base de datos (INF)
- ▶ Física computacional (FIS)
- ▶ Inteligencia artificial (INF)
- ▶ Minería de datos (ELO)
- ▶ Proc. imágenes digitales (ELO)
- ▶ Teoría de información (ELO)



Comentarios/dudas?

