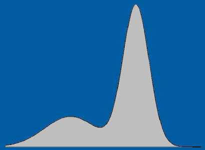


Misturas de distribuição

Misturas finitas de distribuição

- São adequados na para modelar dados com presença de subgrupos não observáveis;

Misturas finitas de distribuição

- São adequados na para modelar dados com presença de subgrupos não observáveis;
- Podem ser utilizados para agrupamento de observações.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Misturas finitas de distribuição

Definamos uma mistura finita de distribuições da seguinte maneira. Seja Z uma variável aleatória discreta com

$$P(Z = j) = p_j, \; j = 1, 2, \ldots, G.$$

Misturas finitas de distribuição

Definamos uma mistura finita de distribuições da seguinte maneira. Seja Z uma variável aleatória discreta com

$$P(Z = j) = p_j,\ j = 1, 2, \ldots, G.$$

Seja $Y|Z = j, \boldsymbol{\theta}_j$ uma variável aleatória contínua com função densidade de probabilidade $g(y|Z = j, \boldsymbol{\theta}_j)$.

Misturas finitas de distribuição

Definamos uma mistura finita de distribuições da seguinte maneira. Seja Z uma variável aleatória discreta com

$$P(Z = j) = p_j, \; j = 1, 2, \ldots, G.$$

Seja $Y|Z = j, \boldsymbol{\theta}_j$ uma variável aleatória contínua com função densidade de probabilidade $g(y|Z = j, \boldsymbol{\theta}_j)$. É possível mostrar que a densidade marginal de $Y|\boldsymbol{\theta}$ é dada por

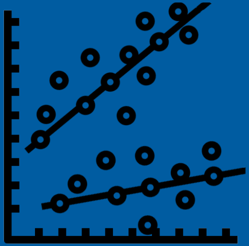
$$f(y|\boldsymbol{\theta}) = \sum_{j=1}^{G} p_j g(y|Z = j, \boldsymbol{\theta}_j),$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Misturas finitas de distribuição

Dizemos que $Y|\boldsymbol{\theta}$ é uma mistura finita de distribuições com G componentes em que cada componente tem densidade g.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics



Mistura de regressões t assimétricas

t de Student assimétrica

Dizemos que uma variável aleatória Y tem distribuição t de Student assimétrica quando

$$Y = \mu + X/\sqrt{U}.$$

Em que $X \sim \mathrm{SN}(0, \sigma^2, \lambda)$ (Azzalini, 1985) e $U \sim \mathrm{Gama}(\nu/2, \nu/2)$. Escrevemos $Y \sim \mathrm{ST}(\mu, \sigma^2, \lambda, \nu)$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

t de Student assimétrica

Dizemos que uma variável aleatória Y tem distribuição t de Student assimétrica quando

$$Y = \mu + X/\sqrt{U}.$$

Em que $X \sim \mathrm{SN}(0, \sigma^2, \lambda)$ (Azzalini, 1985) e $U \sim \mathrm{Gama}(\nu/2, \nu/2)$. Escrevemos $Y \sim \mathrm{ST}(\mu, \sigma^2, \lambda, \nu)$. A v.a. Y tem densidade dada por

$$\mathrm{St}(y|\mu, \sigma^2, \lambda, \nu) = 2\mathrm{t}(y|\mu, \sigma^2, \nu)\mathrm{T}_{\nu+1}\left(z\lambda\sqrt{\frac{\nu+1}{z^2+\nu}}\right),$$

com $z = (y-\mu)/\sigma$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

t de Student assimétrica

Dizemos que uma variável aleatória Y tem distribuição t de Student assimétrica quando

$$Y = \mu + X/\sqrt{U}.$$

Em que $X \sim \mathrm{SN}(0, \sigma^2, \lambda)$ (Azzalini, 1985) e $U \sim \mathrm{Gama}(\nu/2, \nu/2)$. Escrevemos $Y \sim \mathrm{ST}(\mu, \sigma^2, \lambda, \nu)$. A v.a. Y tem densidade dada por

$$\mathrm{St}(y|\mu, \sigma^2, \lambda, \nu) = 2\mathrm{t}(y|\mu, \sigma^2, \nu)\mathrm{T}_{\nu+1}\left(z\lambda\sqrt{\frac{\nu+1}{z^2+\nu}}\right),$$

com $z = (y-\mu)/\sigma$.

$$\mathbb{E}(Y) = \mu + \Delta\sqrt{\frac{\nu}{\pi}}\frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma(\nu/2)}$$

t de Student assimétrica

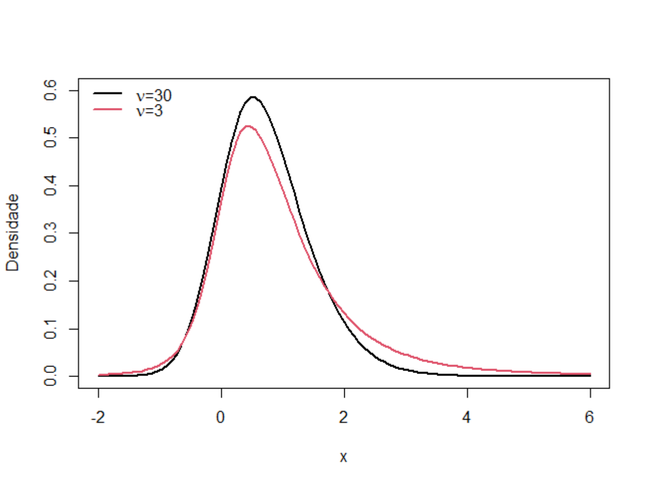


Figura 1: Função densidade da t assimétrica.

Representação hierárquica

A distribuição t assimétrica admite a seguinte representação hierárquica:

$$\begin{aligned} Y|T = t, U = u &\sim \mathrm{N}(\mu + \Delta t, u^{-1}\tau^2); \\ T|U = u &\sim \mathrm{N}_{(0,\infty)}(0, u^{-1}); \\ U &\sim \mathrm{Gama}(\nu/2, \nu/2). \end{aligned} \tag{1}$$

com $\tau^2 = \sigma^2(1 - \delta^2)$, $\Delta = \sigma\delta$ e $\delta = \lambda/\sqrt{1 + \lambda^2}$.

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \varepsilon_i, \; j = 1, \ldots, G.$$

PIGES UFSCAR/USP
Joint Graduate Program in Statistics

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \varepsilon_i, \;\; j = 1, \ldots, G.$$

- G é o número de componentes na mistura;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \varepsilon_i, \; j = 1, \ldots, G.$$

- G é o número de componentes na mistura;
- $\varepsilon_i | Z_i = j \sim \mathrm{ST}(b\Delta_j, \sigma_j^2, \lambda_j, \nu)$;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \varepsilon_i, \; j = 1, \ldots, G.$$

- G é o número de componentes na mistura;
- $\varepsilon_i | Z_i = j \sim \mathrm{ST}(b\Delta_j, \sigma_j^2, \lambda_j, \nu)$;
- Graus de liberdade (ν) iguais entre as componentes;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \varepsilon_i, \; j = 1, \ldots, G.$$

- G é o número de componentes na mistura;
- $\varepsilon_i | Z_i = j \sim \mathrm{ST}(b\Delta_j, \sigma_j^2, \lambda_j, \nu)$;
- Graus de liberdade (ν) iguais entre as componentes;
- b é tal que $\mathbb{E}[\varepsilon_i | Z_i = j] = 0$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Mistura de regressões t assimétricas

Seja Y_i a resposta do i-ésimo indivíduo, a este indivíduo considere uma variável latente discreta Z_i tal que, dado $Z_i = j$, Y_i dependa do vetor de covariáveis $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ conforme

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \varepsilon_i, \; j = 1, \ldots, G.$$

- G é o número de componentes na mistura;
- $\varepsilon_i | Z_i = j \sim \mathrm{ST}(b\Delta_j, \sigma_j^2, \lambda_j, \nu)$;
- Graus de liberdade (ν) iguais entre as componentes;
- b é tal que $\mathbb{E}[\varepsilon_i | Z_i = j] = 0$.

$$b = -\sqrt{\frac{\nu}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}.$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Verossimilhança

Como temos independência entre indivíduos, a função de verossimilhança será então

$$\mathrm{L}(\boldsymbol{\theta}'|\boldsymbol{y},\boldsymbol{X}) \equiv f(\boldsymbol{y}|\boldsymbol{X},\boldsymbol{\theta}') = \prod_{i=1}^{n}\sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Verossimilhança

Como temos independência entre indivíduos, a função de verossimilhança será então

$$\mathrm{L}(\boldsymbol{\theta}'|\boldsymbol{y}, \boldsymbol{X}) \equiv f(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\theta}') = \prod_{i=1}^{n} \sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

Consideraremos a reparametrização

$$\begin{cases} \tau_j^2 = \sigma_j^2(1 - \delta_j^2); \\ \Delta_j = \sigma_j \delta_j, \end{cases}$$

com $\delta_j = \lambda_j / \sqrt{1 + \lambda_j^2}$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Verossimilhança

Como temos independência entre indivíduos, a função de verossimilhança será então

$$\mathrm{L}(\boldsymbol{\theta}'|\boldsymbol{y},\boldsymbol{X}) \equiv f(\boldsymbol{y}|\boldsymbol{X},\boldsymbol{\theta}') = \prod_{i=1}^{n}\sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

Consideraremos a reparametrização

$$\begin{cases} \tau_j^2 = \sigma_j^2(1-\delta_j^2); \\ \Delta_j = \sigma_j\delta_j, \end{cases}$$

com $\delta_j = \lambda_j/\sqrt{1+\lambda_j^2}$. Desta forma,

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X}) \equiv f(\boldsymbol{y}|\boldsymbol{X},\boldsymbol{\theta}) = \prod_{i=1}^{n}\sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j + b\Delta_j, \tau_j^2 + \Delta_j^2, \Delta_j/\tau_j, \nu).$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Verossimilhança

O processo inferencial será feito para $\boldsymbol{\theta}$, após isto recuperamos os valores de interesse $\boldsymbol{\theta}'$ através da transformação inversa:

$$\begin{cases} \sigma_j^2 = \tau_j^2 + \Delta_j^2; \\ \lambda_j = \Delta_j / \tau_j. \end{cases}$$

Label Switching

O modelo apresenta um problema de identificabilidade conhecido como *Label Switching*.

Label Switching

O modelo apresenta um problema de identificabilidade conhecido como *Label Switching*.
Por exemplo, considere $n = 1$ e $G = 2$,

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{X}) = p_1 \mathrm{St}(y_1|\boldsymbol{x}_1^\top \boldsymbol{\beta}_1 + b\Delta_1, \tau_1^2 + \Delta_1^2, \Delta_1/\tau_1, \nu) + p_2 \mathrm{St}(y_1|\boldsymbol{x}_1^\top \boldsymbol{\beta}_2 + b\Delta_2, \tau_2^2 + \Delta_2^2, \Delta_2/\tau_2, \nu).$$

Label Switching

O modelo apresenta um problema de identificabilidade conhecido como *Label Switching*.
Por exemplo, considere $n = 1$ e $G = 2$,

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X}) = p_1 \mathrm{St}(y_1|\boldsymbol{x}_1^\top\boldsymbol{\beta}_1 + b\Delta_1, \tau_1^2 + \Delta_1^2, \Delta_1/\tau_1, \nu) + p_2 \mathrm{St}(y_1|\boldsymbol{x}_1^\top\boldsymbol{\beta}_2 + b\Delta_2, \tau_2^2 + \Delta_2^2, \Delta_2/\tau_2, \nu).$$

Dada uma permutação κ de $1, 2, \ldots, G$ qualquer,

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X}) = \prod_{i=1}^{n}\sum_{j=1}^{G} p_{\kappa(j)} \mathrm{St}(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_{\kappa(j)} + b\Delta_{\kappa(j)}, \tau_{\kappa(j)}^2 + \Delta_{\kappa(j)}^2, \Delta_{\kappa(j)}/\tau_{\kappa(j)}, \nu).$$

Label Switching

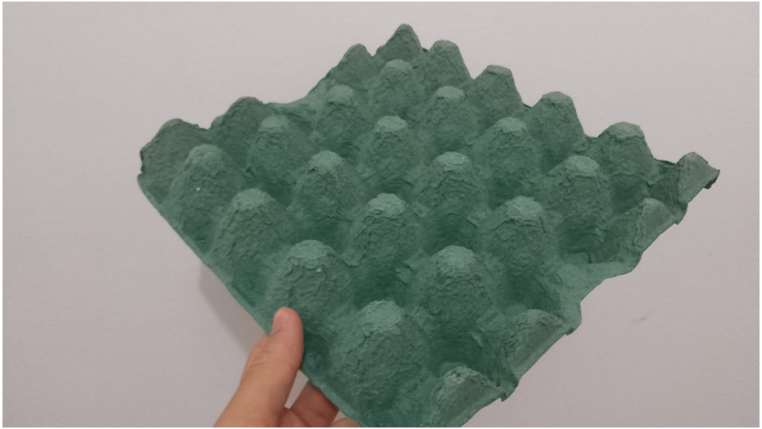
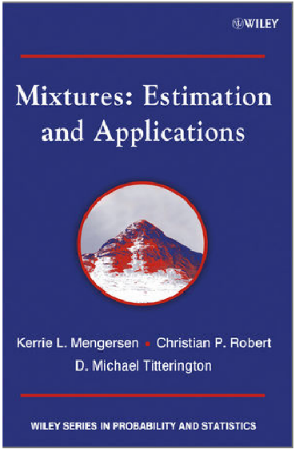


Figura 2: Cartela de ovo.

Label Switching



10

Dealing with label switching under model uncertainty

Sylvia Frühwirth-Schnatter

10.1 Introduction

This chapter considers identification of a finite mixture distribution with K components,

$$p(\mathbf{y}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K, \boldsymbol{\eta}) = \sum_{k=1}^{K} \eta_k f_{\mathcal{T}}(\mathbf{y}|\boldsymbol{\theta}_k), \qquad (10.1)$$

where $\mathbf{y}$ is the realisation of a univariate or multivariate, discrete- or continuous-valued random variable and the component densities $f_{\mathcal{T}}(\mathbf{y}|\boldsymbol{\theta}_k)$ arise from the same distribution family $\mathcal{T}(\boldsymbol{\theta})$ indexed by a parameter $\boldsymbol{\theta}$ taking values in Θ. Given a sample $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$, identification of (10.1) concerns estimating K, the component-specific parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K$ as well as the weight distribution $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_K)$. A comprehensive review of finite mixture models and their estimation is provided by the monographs of Titterington *et al.* (1985), McLachlan and Peel (2000) and Frühwirth-Schnatter (2006). This chapter is based on Bayesian inference using MCMC methods, which date back to Diebolt and Robert (1994); see Frühwirth-Schnatter (2006, Chapter 3) and Marin *et al.* (2005) for a review.

Figura 3: Solução para o *Label Switching*.

Distribuições a *priori*

As especificações a *priori* são

- Consideramos o número de componentes G conhecido;
- Independência a *priori* entre $\boldsymbol{p}$, $\boldsymbol{\beta}$, $\boldsymbol{\tau^2}$, $\boldsymbol{\Delta}$ e ν;
- Independência a *priori* entre componentes para os parâmetros específicos de cada componente, $\boldsymbol{\beta}$, $\boldsymbol{\tau^2}$ e $\boldsymbol{\Delta}$ (por exemplo $\boldsymbol{\beta_1} \perp \boldsymbol{\beta_2}$).

Distribuições a *priori*

As especificações são as seguintes:

$$
\begin{aligned}
\boldsymbol{p} &\sim \text{Dirichlet}(\xi_1, \xi_2, \ldots, \xi_G);\\
\boldsymbol{\beta}_j &\sim \text{N}_{k+1}(0, c^2\mathbb{I}_{k+1});\\
\Delta_j &\sim \text{N}(\eta, \omega^2);\\
\tau_j^2 &\sim \text{InvGama}(r, s);\\
\nu|\alpha &\sim \exp(\alpha), \alpha \sim \text{Unif}(0.02, 0.5),
\end{aligned}
$$

em que k é o número de covariáveis que compõem $\boldsymbol{X}$ e $\mathbb{I}_{k+1}$ denota a matriz identidade de ordem $k+1$.

Distribuições a *priori*

Journal of Applied Statistics >
Volume 42, 2015 - Issue 12

Submit an article | Journal homepage

Enter keywords, authors, DOI, etc

545
Views
17
CrossRef citations to date
0
Altmetric

Original Articles

Bayesian analysis of censored linear regression models with scale mixtures of normal distributions

Aldo M. Garay, Heleno Bolfarine, Victor H. Lachos & Celso R.B. Cabral

Pages 2694-2714 | Received 17 Sep 2013, Accepted 04 May 2015, Published online: 09 Aug 2015

Cite this article https://doi.org/10.1080/02664763.2015.1048671 Check for updates

Full Article | Figures & data | References | Citations | Metrics | Reprints & Permissions | Read this article

Abstract

Related

Figura 4: Motivação para a *priori* de ν

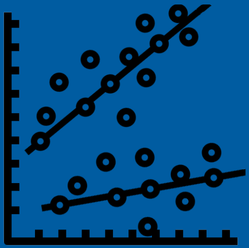
PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Distribuições a *priori*

A *priori* para $\boldsymbol{\theta}$ terá a seguinte estrutura:

$$\begin{aligned}
\pi(\boldsymbol{\theta}) &= \pi(\boldsymbol{p}, \boldsymbol{\beta}, \boldsymbol{\tau^2}, \boldsymbol{\Delta}, \nu, \alpha) \\
&= \pi(\boldsymbol{p})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\tau^2})\pi(\boldsymbol{\Delta})\pi(\nu, \alpha) \\
&= \pi(\boldsymbol{p})\pi(\nu|\alpha)\pi(\alpha)\prod_{j=1}^{G}\pi(\boldsymbol{\beta}_j)\pi(\tau_j^2)\pi(\Delta_j).
\end{aligned}$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics



Estimação, ajuste e seleção de modelos

Distribuição a *posteriori*

Em uma abordagem Bayesiana, o objeto principal é a distribuição a *posteriori*, que neste modelo deve ser tal que:

$$
\begin{aligned}
\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X}) &\propto \mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X})\pi(\boldsymbol{\theta}) \\
&= \prod_{i=1}^{n}\sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j + b\Delta_j, \tau_j^2 + \Delta_j^2, \Delta_j/\tau_j, \nu) \\
&\times \pi(\boldsymbol{p})\pi(\nu|\alpha)\pi(\alpha)\prod_{j=1}^{G}\pi(\boldsymbol{\beta}_j)\pi(\tau_j^2)\pi(\Delta_j).
\end{aligned}
$$

O que fazer?

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{u}) = \prod_{j=1}^{G} p_j^{m_j} \prod_{i=1}^{n} [\phi(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \Delta_j(b+t_i), u_i^{-1}\tau_j^2)\mathrm{Nt}(t_i|0, u_i^{-1})\mathrm{Gm}(u_i|\nu/2, \nu/2)].^{\mathbb{1}(Z_i=j)}$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) = \prod_{j=1}^{G} p_j^{m_j} \prod_{i=1}^{n} [\phi(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j+\Delta_j(b+t_i), u_i^{-1}\tau_j^2)\mathrm{Nt}(t_i|0,u_i^{-1})\mathrm{Gm}(u_i|\nu/2,\nu/2)]^{\mathbb{1}(Z_i=j)}$$

- Desejamos então encontrar $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) \propto \mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})\pi(\boldsymbol{\theta})$;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) = \prod_{j=1}^{G} p_j^{m_j} \prod_{i=1}^{n} [\phi(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \Delta_j(b+t_i), u_i^{-1}\tau_j^2)\mathrm{Nt}(t_i|0, u_i^{-1})\mathrm{Gm}(u_i|\nu/2, \nu/2)]^{\mathbb{1}(Z_i=j)}$$

- Desejamos então encontrar $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) \propto \mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})\pi(\boldsymbol{\theta})$;
- É possível encontrar as condicionais completas de $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})$;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) = \prod_{j=1}^{G} p_j^{m_j} \prod_{i=1}^{n} [\phi(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j+\Delta_j(b+t_i), u_i^{-1}\tau_j^2)\mathrm{Nt}(t_i|0,u_i^{-1})\mathrm{Gm}(u_i|\nu/2,\nu/2)]^{\mathbb{1}(Z_i=j)}$$

- Desejamos então encontrar $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) \propto \mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})\pi(\boldsymbol{\theta})$;
- É possível encontrar as condicionais completas de $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})$;
- Utilizaremos o algoritmo *Gibbs sampler*, para obtermos amostras da distribuição a *posteriori*;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) = \prod_{j=1}^{G} p_j^{m_j} \prod_{i=1}^{n} [\phi(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \Delta_j(b+t_i), u_i^{-1}\tau_j^2)\mathrm{Nt}(t_i|0,u_i^{-1})\mathrm{Gm}(u_i|\nu/2,\nu/2)]^{\mathbb{1}(Z_i=j)}.$$

- Desejamos então encontrar $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) \propto \mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})\pi(\boldsymbol{\theta})$;
- É possível encontrar as condicionais completas de $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})$;
- Utilizaremos o algoritmo *Gibbs sampler*, para obtermos amostras da distribuição a *posteriori*;
- Com o *Gibbs sampler* construímos uma Cadeia de Markov cuja distribuição estacionária é a *posteriori*;

O que fazer?

- Consideraremos que são observadas as variáveis Z, T e U mostradas na representação hierárquica.

$$\mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) = \prod_{j=1}^{G} p_j^{m_j} \prod_{i=1}^{n} [\phi(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + \Delta_j(b+t_i), u_i^{-1}\tau_j^2)\mathrm{Nt}(t_i|0, u_i^{-1})\mathrm{Gm}(u_i|\nu/2,\nu/2)].^{\mathbb{1}(Z_i=j)}$$

- Desejamos então encontrar $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u}) \propto \mathrm{L}(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})\pi(\boldsymbol{\theta})$;
- É possível encontrar as condicionais completas de $\pi(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{z},\boldsymbol{t},\boldsymbol{u})$;
- Utilizaremos o algoritmo *Gibbs sampler*, para obtermos amostras da distribuição a *posteriori*;
- Com o *Gibbs sampler* construímos uma Cadeia de Markov cuja distribuição estacionária é a *posteriori*;
- Quantidades de interesse são calculadas com base na cadeia obtida, técnica denominada MCMC.

Condicionais completas

São as distribuições de um parâmetro do modelo dados os demais parâmetros e a amostra, por exemplo, a condicional completa de $\boldsymbol{\beta_1}$ é $\boldsymbol{\beta}_1|\boldsymbol{\beta}_{-1}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{u}, \ldots$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Condicionais completas

São as distribuições de um parâmetro do modelo dados os demais parâmetros e a amostra, por exemplo, a condicional completa de $\boldsymbol{\beta_1}$ é $\beta_1|\boldsymbol{\beta}_{-1}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{u}, \ldots$.

$$
\begin{aligned}
\boldsymbol{p}|\boldsymbol{y}, \boldsymbol{z}, \ldots &\sim \text{Dirichlet}(\xi_1 + m_1, \xi_2 + m_2, \ldots, \xi_G + m_G) \\
\boldsymbol{\beta}_j|\boldsymbol{\beta}_{-j}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{u}, \ldots &\sim \text{N}_k(\Sigma_j^{-1}\boldsymbol{w}_j, \tau_j^2\Sigma_j^{-1}) \\
\tau_j^2|\boldsymbol{\tau^2}_{-j}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{u}, \ldots &\sim \text{InvGama}\left(\frac{m_j}{2} + r, \frac{S_3 + s}{2}\right) \\
\Delta_j|\boldsymbol{\Delta}_{-j}, , \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}, \boldsymbol{u}, \ldots &\sim \text{N}\left(\frac{\omega S_1 + \eta\tau_j^2}{\omega S_2 + \tau_j^2}, \frac{\omega^2\tau_j^2}{\omega S_2 + \tau_j^2}\right) \\
\alpha|\boldsymbol{y}, \ldots &\sim \text{Gama}_{(0.02,0.5)}(2, \nu), \\
\mathbb{P}(Z_i = j|\boldsymbol{z}_{-i}, \boldsymbol{y}, \ldots) &\propto p_j\text{St}(y_i|\boldsymbol{x}_i^\top\boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu)
\end{aligned}
$$

Condicionais completas

$$t_i | \boldsymbol{t}_{-i}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{u}, \ldots \sim \mathrm{N}_{(0,\infty)}\left(\frac{(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_j - b\Delta_j)\Delta_j}{\Delta_j^2 + \tau_j^2}, \frac{\tau_j^2}{u_i(\Delta_j^2 + \tau_j^2)}\right),$$

$$u_i | \boldsymbol{u}_{-i}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{t}, \ldots \sim \mathrm{Gama}\left(\frac{\nu}{2} + 1, \frac{(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}_j - (b + t_i)\Delta_j)^2}{2\tau_j^2} + \frac{t_i^2 + \nu}{2}\right)$$

Condicionais completas

$$t_i|\boldsymbol{t}_{-i},\boldsymbol{y},\boldsymbol{z},\boldsymbol{u},\ldots \sim \mathrm{N}_{(0,\infty)}\left(\frac{(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_j-b\Delta_j)\Delta_j}{\Delta_j^2+\tau_j^2},\frac{\tau_j^2}{u_i(\Delta_j^2+\tau_j^2)}\right),$$

$$u_i|\boldsymbol{u}_{-i},\boldsymbol{y},\boldsymbol{z},\boldsymbol{t},\ldots \sim \mathrm{Gama}\left(\frac{\nu}{2}+1,\frac{(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_j-(b+t_i)\Delta_j)^2}{2\tau_j^2}+\frac{t_i^2+\nu}{2}\right)$$

Definindo $\boldsymbol{w}_j=\boldsymbol{X}_j^\top\boldsymbol{U}_j(\boldsymbol{y}_j-(b+\boldsymbol{t}_j)\Delta_j)$ e $\Sigma_j=\boldsymbol{X}_j^\top\boldsymbol{U}_j\boldsymbol{X}_j+\tau_j^2c^{-2}\mathbb{I}_{k+1}$, com $\boldsymbol{U}_j=\mathrm{diag}(\boldsymbol{u}_j)$, $\boldsymbol{X}_j$ como a matriz composta das i-ésimas linhas de $\boldsymbol{X}$ tais que $z_i=j$, $\boldsymbol{y}_j,\boldsymbol{t}_j,\boldsymbol{u}_j$ denotando o vetor composto pelos i-ésimos elementos dos respectivos vetores $\boldsymbol{y},\boldsymbol{t},\boldsymbol{u}$ tais que $z_i=j$,

- $S_1=\sum_{I_j}u_i(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_j)(b+t_i)$;
- $S_2=\sum_{I_j}u_i(b+t_i)^2$;
- $S_3=\sum_{I_j}u_i(y_i-\boldsymbol{x}_i^\top\boldsymbol{\beta}_j-(b+t_i)\Delta_j)^2$, com $I_j=\{i|z_i=j\}$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Condicionais completas

A exceção é condicional completa de ν, cuja densidade é

$$\pi(\nu|\boldsymbol{y},\dots) \propto e^{-\alpha\nu} \prod_{i=1}^{n} \sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

Que não é o núcleo de uma distribuição conhecida, utilizaremos o algoritmo de Metropolis-Hastings nesse passo.

Condicionais completas

A exceção é condicional completa de ν, cuja densidade é

$$\pi(\nu|\boldsymbol{y},\ldots) \propto e^{-\alpha\nu} \prod_{i=1}^{n} \sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

Que não é o núcleo de uma distribuição conhecida, utilizaremos o algoritmo de Metropolis-Hastings nesse passo.

1. Geramos uma observação proposta $\nu^* \sim \mathrm{LN}(\log(\nu^{(q-1)}), \phi_\nu^2)$;

Condicionais completas

A exceção é condicional completa de ν, cuja densidade é

$$\pi(\nu|\boldsymbol{y},\dots) \propto e^{-\alpha\nu} \prod_{i=1}^{n} \sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

Que não é o núcleo de uma distribuição conhecida, utilizaremos o algoritmo de Metropolis-Hastings nesse passo.

1. Geramos uma observação proposta $\nu^* \sim \mathrm{LN}(\log(\nu^{(q-1)}), \phi_\nu^2)$;
2. Calculamos $\mathrm{aceit}(\nu^{(q-1)}, \nu^*) = \min\left\{1, \frac{\pi(\nu^*|\boldsymbol{y},\dots)\nu^*}{\pi(\nu^{(q-1)}|\boldsymbol{y},\dots)\nu^{(q-1)}}\right\}$;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Condicionais completas

A exceção é condicional completa de ν, cuja densidade é

$$\pi(\nu|\boldsymbol{y},\dots) \propto e^{-\alpha\nu} \prod_{i=1}^{n} \sum_{j=1}^{G} p_j \mathrm{St}(y_i|\boldsymbol{x}_i^\top \boldsymbol{\beta}_j + b\Delta_j, \sigma_j^2, \lambda_j, \nu).$$

Que não é o núcleo de uma distribuição conhecida, utilizaremos o algoritmo de Metropolis-Hastings nesse passo.

1. Geramos uma observação proposta $\nu^* \sim \mathrm{LN}(\log(\nu^{(q-1)}), \phi_\nu^2)$;
2. Calculamos $\mathrm{aceit}(\nu^{(q-1)}, \nu^*) = \min\left\{1, \frac{\pi(\nu^*|\boldsymbol{y},\dots)\nu^*}{\pi(\nu^{(q-1)}|\boldsymbol{y},\dots)\nu^{(q-1)}}\right\}$;
3. Com probabilidade $\mathrm{aceit}(\nu^{(q-1)}, \nu^*)$, fazemos $\nu^{(q)} = \nu^*$, caso contrário $\nu^{(q)} = \nu^{(q-1)}$.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Solucionando *Label Switching*

- Obtemos as amostras;

Solucionando *Label Switching*

- Obtemos as amostras;
- Descartamos uma certa quantidade de forma a obter convergência, ficamos com Q amostras;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Solucionando *Label Switching*

- Obtemos as amostras;
- Descartamos uma certa quantidade de forma a obter convergência, ficamos com Q amostras;

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} \\ \Delta_1^{(2)} & \Delta_2^{(2)} \\ \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix}.$$

Solucionando *Label Switching*

- Obtemos as amostras;
- Descartamos uma certa quantidade de forma a obter convergência, ficamos com Q amostras;

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} \\ \Delta_1^{(2)} & \Delta_2^{(2)} \\ \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix}.$$

A solução proposta por Frühwirth-Schnatter (2011) é:

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} & \Delta_1^{(2)} & \Delta_2^{(2)} & \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix}$$

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Solucionando *Label Switching*

- Obtemos as amostras;
- Descartamos uma certa quantidade de forma a obter convergência, ficamos com Q amostras;

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} \\ \Delta_1^{(2)} & \Delta_2^{(2)} \\ \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix}.$$

A solução proposta por Frühwirth-Schnatter (2011) é:

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} & \Delta_1^{(2)} & \Delta_2^{(2)} & \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix} \xrightarrow{k-means} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Solucionando *Label Switching*

- Obtemos as amostras;
- Descartamos uma certa quantidade de forma a obter convergência, ficamos com Q amostras;

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} \\ \Delta_1^{(2)} & \Delta_2^{(2)} \\ \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix}.$$

A solução proposta por Frühwirth-Schnatter (2011) é:

$$\begin{bmatrix} \Delta_1^{(1)} & \Delta_2^{(1)} & \Delta_1^{(2)} & \Delta_2^{(2)} & \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix} \xrightarrow{k-means} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} \Delta_2^{(2)} & \Delta_1^{(2)} \\ \Delta_1^{(3)} & \Delta_2^{(3)} \end{bmatrix}$$

Solucionando *Label Switching*

- Ao final do processo temos Q_κ amostras;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Solucionando *Label Switching*

- Ao final do processo temos Q_κ amostras;
- Uma baixa fração Q_κ/Q indica que o número G fixado sobrestima o número de componentes que gerou os dados;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Solucionando *Label Switching*

- Ao final do processo temos Q_κ amostras;
- Uma baixa fração Q_κ/Q indica que o número G fixado sobrestima o número de componentes que gerou os dados;
- Uma desvantagem deste método é que perdemos o controle sobre o número de amostras geradas.

Seleção de modelos

- Foi escolhido o critério de informação WAIC;

Seleção de modelos

- Foi escolhido o critério de informação WAIC;
- É invariante ao *Label Switching*, podendo ser calculado com as amostras sem a correção.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Seleção de modelos

- Foi escolhido o critério de informação WAIC;
- É invariante ao *Label Switching*, podendo ser calculado com as amostras sem a correção.

$$\text{WAIC} = -\frac{4}{Q}\sum_{q=1}^{Q}\sum_{i=1}^{n}\log f(y_i|\boldsymbol{\theta}^{(q)}) + 2\sum_{i=1}^{n}\log\left(\frac{1}{Q}\sum_{q=1}^{Q} f(y_i|\boldsymbol{\theta}^{(q)})\right).$$

Análise de resíduos

- Utilizamos os resíduos quantílicos, que possuem distribuição normal caso o modelo esteja bem especificado;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Análise de resíduos

- Utilizamos os resíduos quantílicos, que possuem distribuição normal caso o modelo esteja bem especificado;
- Podemos obter uma amostra dos resíduos quantilicos utilizando as amostras sem a correção de *Label Switching*.

Gerando dados

Gerando dados

- Simulados utilizando a representação hierárquica do modelo fixando $G = 3$;

Gerando dados

- Simulados utilizando a representação hierárquica do modelo fixando $G = 3$;
- Considerados 3 tipos de dados:

Gerando dados

- Simulados utilizando a representação hierárquica do modelo fixando $G = 3$;
- Considerados 3 tipos de dados:
  1. Balanceados com componentes separadas;

Gerando dados

- Simulados utilizando a representação hierárquica do modelo fixando $G = 3$;
- Considerados 3 tipos de dados:
  1. Balanceados com componentes separadas;
  2. Desbalanceados com componentes separadas;

Gerando dados

- Simulados utilizando a representação hierárquica do modelo fixando $G = 3$;
- Considerados 3 tipos de dados:
  1. Balanceados com componentes separadas;
  2. Desbalanceados com componentes separadas;
  3. Desbalanceados com componentes próximas.

Gerando dados

- Simulados utilizando a representação hierárquica do modelo fixando $G = 3$;
- Considerados 3 tipos de dados:
    1. Balanceados com componentes separadas;
    2. Desbalanceados com componentes separadas;
    3. Desbalanceados com componentes próximas.
- Dois tamanhos amostrais $n = 100$ e $n = 1000$;
- Matriz modelo $\boldsymbol{X}$: uns na primeira coluna e a segunda n amostras de uma normal padrão.

Gerando dados

- Parâmetros comuns entre os tipos de dados:

$$\boldsymbol{\sigma^2} = \begin{bmatrix} 1 & 4 & 9 \end{bmatrix}^\top, \quad \boldsymbol{\lambda} = \begin{bmatrix} 10 & -8 & -8 \end{bmatrix}^\top \text{ e } \quad \nu = 3;$$

Gerando dados

- Parâmetros comuns entre os tipos de dados:

$$\boldsymbol{\sigma^2} = \begin{bmatrix} 1 & 4 & 9 \end{bmatrix}^\top, \quad \boldsymbol{\lambda} = \begin{bmatrix} 10 & -8 & -8 \end{bmatrix}^\top \text{ e } \quad \nu = 3;$$

- Parâmetros exclusivos:
  - Tipo 1:

$$\boldsymbol{p} = \begin{bmatrix} 0,3 & 0,4 & 0,3 \end{bmatrix}^\top, \boldsymbol{\beta} = \begin{bmatrix} 30 & 15 & -30 \\ 4 & 4 & -4 \end{bmatrix}$$

  - Tipo 2:

$$\boldsymbol{p} = \begin{bmatrix} 0,05 & 0,15 & 0,8 \end{bmatrix}^\top, \boldsymbol{\beta} = \begin{bmatrix} 30 & 15 & -30 \\ 4 & 4 & -4 \end{bmatrix}$$

  - Tipo 3:

$$\boldsymbol{p} = \begin{bmatrix} 0,8 & 0,15 & 0,05 \end{bmatrix}^\top, \boldsymbol{\beta} = \begin{bmatrix} 0 & -10 & -25 \\ 4 & -5 & -4 \end{bmatrix}$$

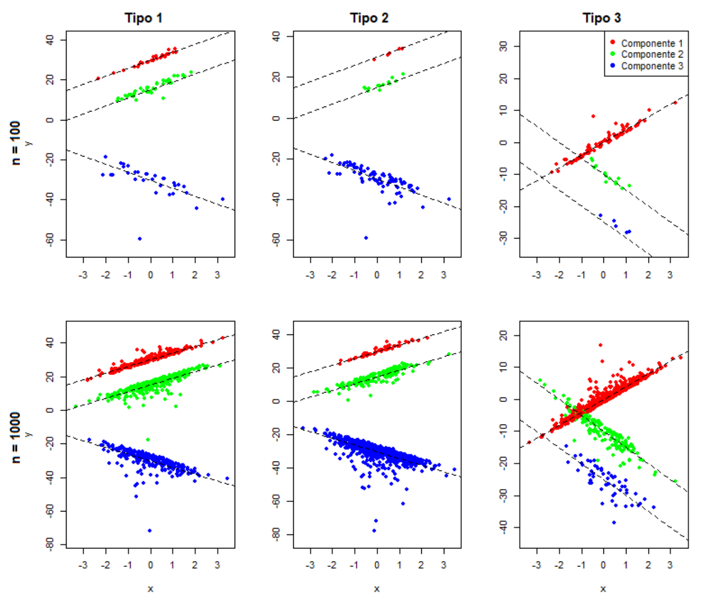


Figura 5: Conjuntos de dados simulados.

Ajuste

- O modelo foi ajustado em cada conjunto de dados fixando $G = 2, 3, 4$;

Ajuste

- O modelo foi ajustado em cada conjunto de dados fixando $G = 2, 3, 4$;
- Fixamos os hiperparâmetros em:

$$
\begin{aligned}
\xi_1 = \xi_2 = \cdots = \xi_G = 1 \\
c = 10 \\
\eta = 0;\ \omega = 10 \\
r = s = 0,1.
\end{aligned}
$$

Ajuste

- O modelo foi ajustado em cada conjunto de dados fixando $G = 2, 3, 4$;
- Fixamos os hiperparâmetros em:

$$\xi_1 = \xi_2 = \cdots = \xi_G = 1$$
$$c = 10$$
$$\eta = 0; \ \omega = 10$$
$$r = s = 0,1.$$

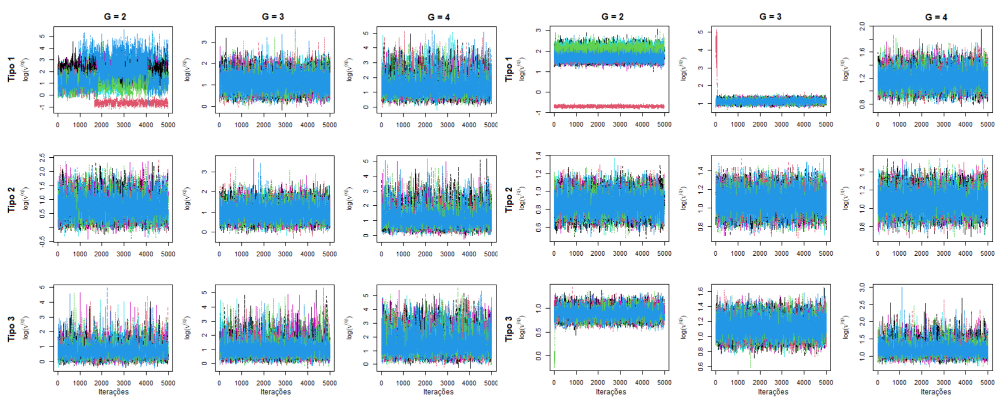
- Com o *Gibbs sampler* obtivemos $Q = 5.000$ amostras após os descartes;

Ajuste

- O modelo foi ajustado em cada conjunto de dados fixando $G = 2, 3, 4$;
- Fixamos os hiperparâmetros em:

$$
\begin{aligned}
\xi_1 = \xi_2 = \cdots = \xi_G = 1 \\
c = 10 \\
\eta = 0;\ \omega = 10 \\
r = s = 0,1.
\end{aligned}
$$

- Com o *Gibbs sampler* obtivemos $Q = 5.000$ amostras após os descartes;
- O procedimento foi replicado 10 vezes.



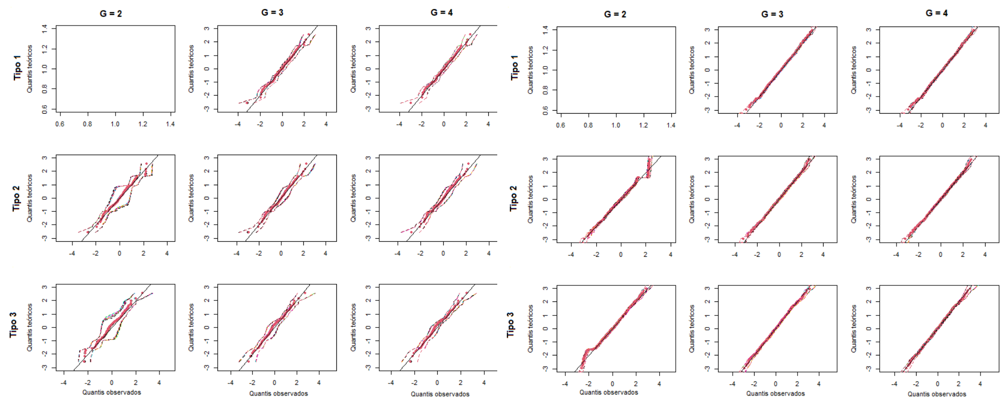
(a) $n = 100$ (b) $n = 1000$

Figura 6: Logaritmo natural das amostras da *posteriori* de ν ao longo das iterações do GS considerando os 3 tipos de dados e número de componentes, para $n = 100$ e $n = 1000$.

Tipo	n	$G = 2$			$G = 3$		
		Taxa	WAIC2/WAIC3	WAIC2/WAIC4	Taxa	WAIC3/WAIC2	WAIC3/WAIC4
1	100	-	-	-	0,9937	-	0,1
	1000	-	-	-	1	-	1
2	100	0,9982	0	0	0,1018	1	0,4
	1000	1	0	0	0,9495	1	0,9
3	100	0,9872	0	0	0,5953	1	0
	1000	1	0	0	0,6539	1	1

Tipo	n	$G = 4$		
		Taxa	WAIC4/WAIC2	WAIC4/WAIC3
1	100	0,1391	-	0,9
	1000	0,1319	-	0
2	100	0,0200	1	0,6
	1000	0,1001	1	0,1
3	100	0,1087	1	1
	1000	0,0957	1	0

Tabela 1: Taxa média de permutações válidas e critério de seleção de modelos para os dados simulados.



(a) $n = 100$ (b) $n = 1000$

Figura 7: Gráfico quantil-quantil dos resíduos quantílicos considerando diversos números de componentes, com $n = 100$ e $n = 1000$ para cada réplica.

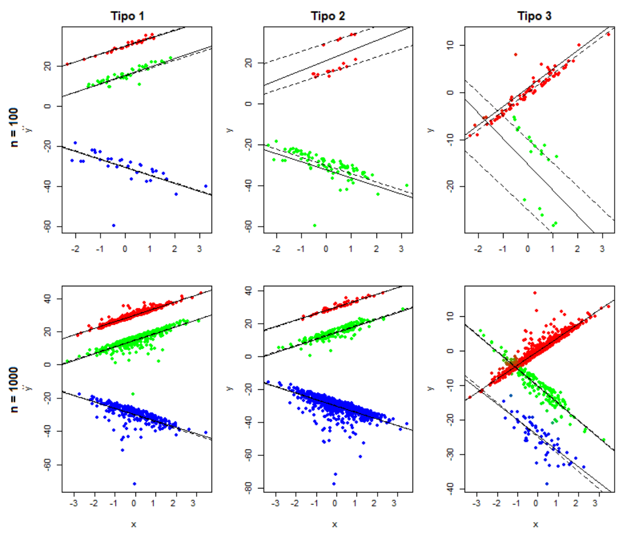


Figura 8: Resultado de classificação e retas de regressão.

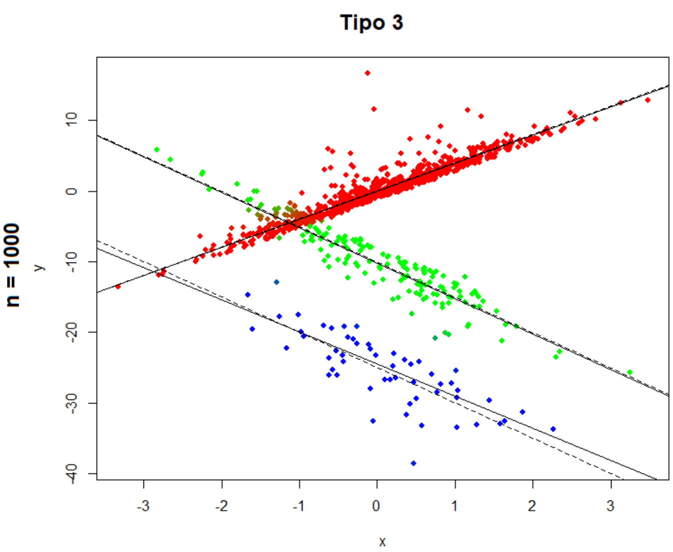


Figura 9: Resultado de classificação e retas de regressão.

Tipo	n	σ_1^2	σ_2^2	σ_3^2	λ_1	λ_2	λ_3	ν
1	100	$0,5788$	$2,8488$	$15,9764$	$2,4419$	$-4,2165$	$-11,8060$	$3,3156$
	1000	$1,3424$	$4,5225$	$8,7590$	$9,1936$	$-10,7483$	$-7,1787$	$3,0757$
2	100	$158,0027$	$35,5182$	-	$9,2582$	$-6,2435$	-	$2,1873$
	1000	$1,1787$	$5,0340$	$9,5194$	$4,7982$	$-7,6128$	$-8,8242$	$2,9235$
3	100	$2,3204$	$22,0667$	-	$4,9606$	$-6,7849$	-	$2,1900$
	1000	$1,0772$	$4,8147$	$7,8710$	$9,2715$	$-5,2249$	$-0,3573$	$3,0855$
Fixado		1	4	9	10	-8	-8	3

Tabela 2: Estimativas pela média a *posteriori* para os dados simulados.

Percepção de tons

Dados

- Disponíveis no R pelo pacote `mixtools`;

Dados

- Disponíveis no R pelo pacote `mixtools`;
- Uma frequência fundamental e um sobretom desta frequência foram tocados para um músico treinado;

Dados

- Disponíveis no R pelo pacote `mixtools`;
- Uma frequência fundamental e um sobretom desta frequência foram tocados para um músico treinado;
- O musico então era instruído a "afinar" o sobretom para a oitava acima da frequência fundamental;

Dados

- Disponíveis no R pelo pacote `mixtools`;
- Uma frequência fundamental e um sobretom desta frequência foram tocados para um músico treinado;
- O musico então era instruído a "afinar" o sobretom para a oitava acima da frequência fundamental;
- Os dados são compostos por 150 replicações deste experimento com o mesmo músico;

Dados

- Disponíveis no R pelo pacote `mixtools`;
- Uma frequência fundamental e um sobretom desta frequência foram tocados para um músico treinado;
- O musico então era instruído a “afinar” o sobretom para a oitava acima da frequência fundamental;
- Os dados são compostos por 150 replicações deste experimento com o mesmo músico;
- Anteriormente analisados por Sabillón et al. (2023) utilizando um modelo de misturas de regressões normais.

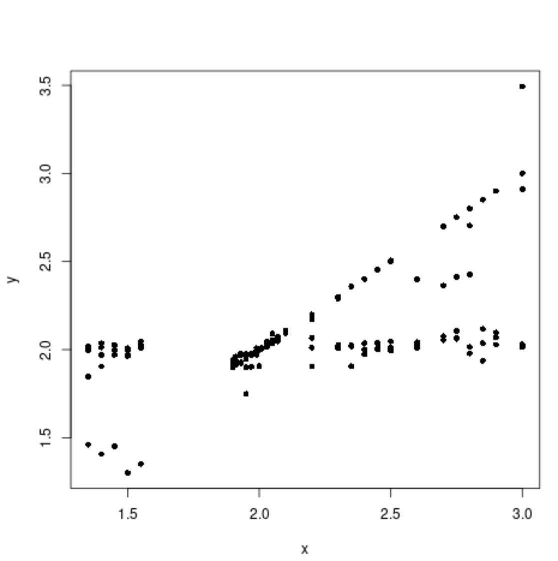


Figura 10: Diagrama de dispersão para os dados do estudo de percepção de tons.

Ajuste

- O modelo foi ajustado fixando $G = 2, 3, 4$;

Ajuste

- O modelo foi ajustado fixando $G = 2, 3, 4$;
- Fixamos os hiperparâmetros em:

$$
\begin{gathered}
\xi_1 = \xi_2 = \cdots = \xi_G = 1 \\
c = 10 \\
\eta = 0;\ \omega = 10 \\
r = s = 0,1.
\end{gathered}
$$

Ajuste

- O modelo foi ajustado fixando $G = 2, 3, 4$;
- Fixamos os hiperparâmetros em:

$$\xi_1 = \xi_2 = \cdots = \xi_G = 1$$
$$c = 10$$
$$\eta = 0; \ \omega = 10$$
$$r = s = 0, 1.$$

- Com o *Gibbs sampler* obtivemos $Q = 5.000$ amostras após os descartes;

Ajuste

- O modelo foi ajustado fixando $G = 2, 3, 4$;
- Fixamos os hiperparâmetros em:

$$\xi_1 = \xi_2 = \cdots = \xi_G = 1$$
$$c = 10$$
$$\eta = 0;\ \omega = 10$$
$$r = s = 0,1.$$

- Com o *Gibbs sampler* obtivemos $Q = 5.000$ amostras após os descartes;
- O procedimento foi replicado 10 vezes.

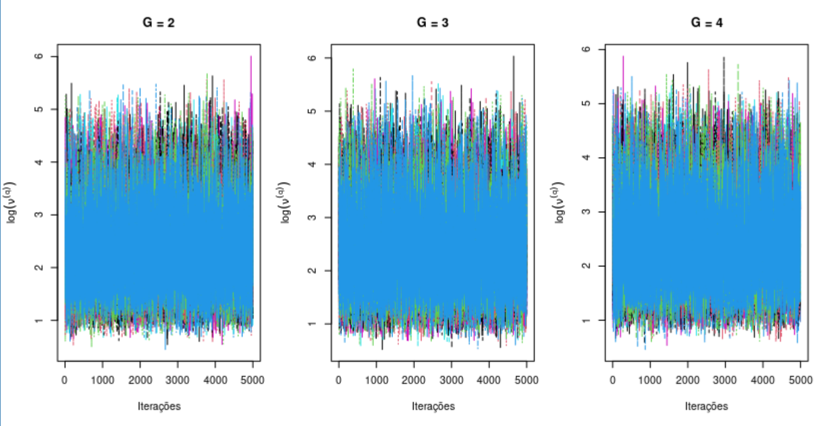


Figura 11: Logaritmo natural das amostras da *posteriori* de ν ao longo das iterações do GS para os dados do estudo de percepção de tons, considerando $G = 2, 3$ e 4.

Nº Componentes	Taxa	WAIC2/WAIC3	WAIC2/WAIC4	WAIC3/WAIC4
2	1	1	1	-
3	0,1204	-	-	1
4	0,0161	-	-	-

Tabela 3: Taxa média de permutações válidas e proporção de vezes que o WAIC de um modelo é menor que o WAIC de outro modelo dentre as replicações.

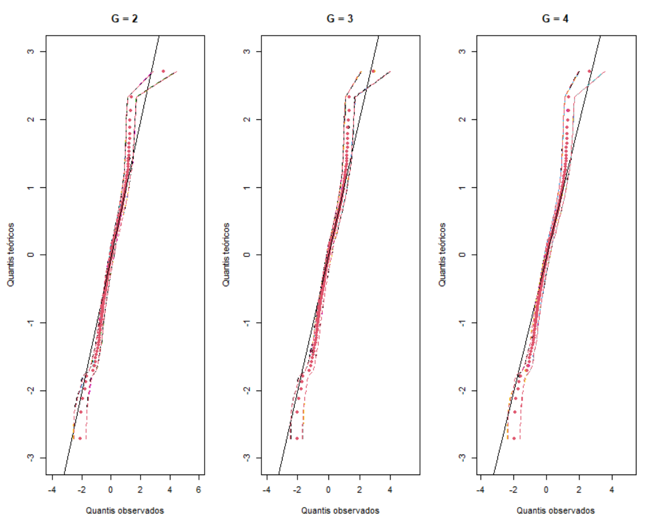


Figura 12: Gráfico quantil-quantil para os dados do estudo de percepção de tons, em cada replicação.

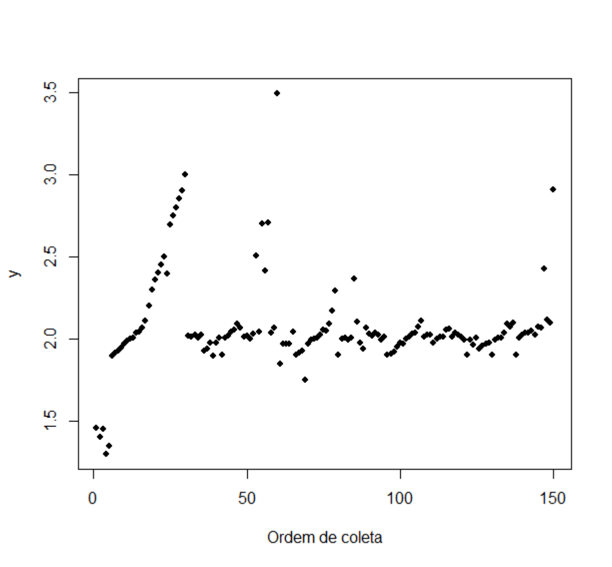


Figura 13: Valor da variável resposta por ordem de coleta.

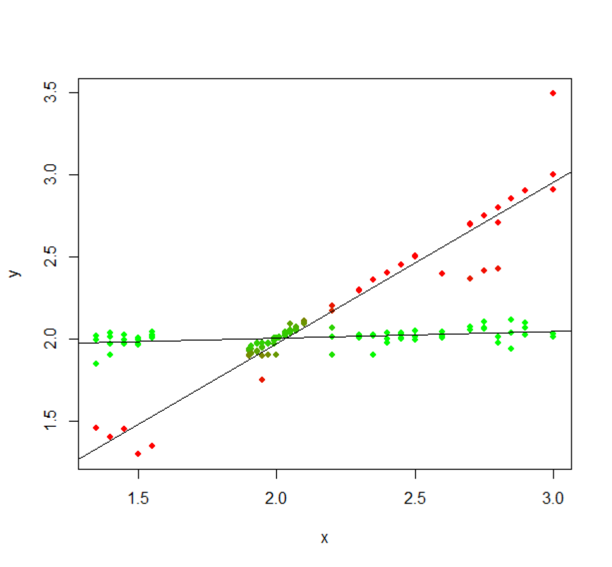


Figura 14: Resultados de classificação e retas de regressão estimadas para os dados do experimento de percepção de tons.

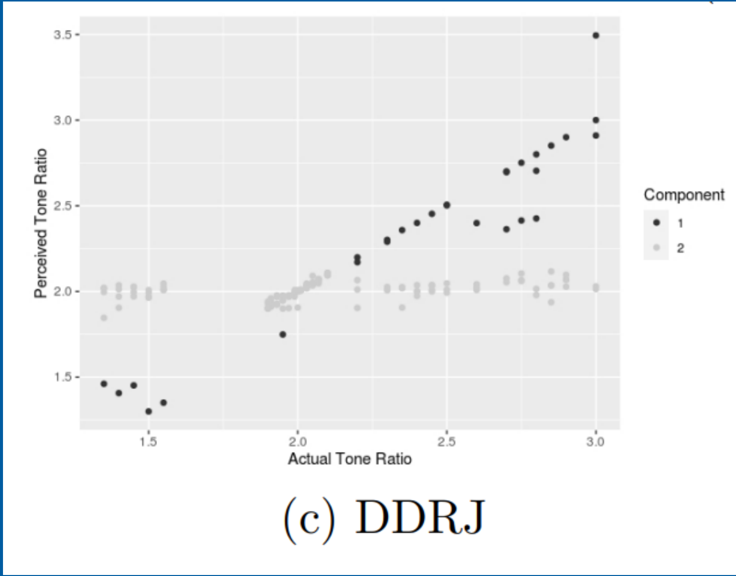


Figura 15: Resultados de classificação de Sabillón et al. (2023).

Considerações finais

- Do estudo de simulação:

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;
  - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;

PIGES UFSCAR/USP
Joint Graduate Program in Statistics

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;
  - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;
- Do estudo de percepção de tons:

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;
  - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;
- Do estudo de percepção de tons:
  - Classificação similar a Sabillón et al. (2023);

Considerações finais

- Do estudo de simulação:
    - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
    - O tamanho amostral influencia na determinação do número de componentes;
    - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;
- Do estudo de percepção de tons:
    - Classificação similar a Sabillón et al. (2023);
    - O modelo não é adequado;

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;
  - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;
- Do estudo de percepção de tons:
  - Classificação similar a Sabillón et al. (2023);
  - O modelo não é adequado;
- Futuramente:
  - Simulação com mais conjuntos de dados para verificar o vício dos estimadores;

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;
  - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;
- Do estudo de percepção de tons:
  - Classificação similar a Sabillón et al. (2023);
  - O modelo não é adequado;
- Futuramente:
  - Simulação com mais conjuntos de dados para verificar o vício dos estimadores;
  - Obter outros conjuntos de dados;

Considerações finais

- Do estudo de simulação:
  - O modelo é capaz de recuperar o valor dos parâmetros fixados de forma consistente quando $n = 1000$;
  - O tamanho amostral influencia na determinação do número de componentes;
  - A taxa de permutações válidas indica de forma mais precisa a sobrestimação do número de componentes;
- Do estudo de percepção de tons:
  - Classificação similar a Sabillón et al. (2023);
  - O modelo não é adequado;
- Futuramente:
  - Simulação com mais conjuntos de dados para verificar o vício dos estimadores;
  - Obter outros conjuntos de dados;
  - Estimação do número de componentes em conjunto dos parâmetros.

Referências I

Azzalini, A. (1985). “A class of distributions which includes the normal ones.” *Scandinavian Journal of Statistics*, 12, 171–178.

Frühwirth-Schnatter, S. (2011). “Dealing with label switching under model uncertainty.” *Mixtures: estimation and applications*, 213–239.

Sabillón, G. A., Cotrim, L. G. F., e Zuanetti, D. A. (2023). “A data-driven reversible jump for estimating a finite mixture of regression models.” *TEST*, 32, 1, 350–369.

PIPGES UFSCAR/USP
Joint Graduate Program in Statistics

Obrigado.

github.com/Bola382/Stocastic