

Suavização de taxas com modelos espaciais

Fernando G. Moro

Universidade Federal do Paraná

22 de julho de 2014

1 Estatística espacial e dados de área;

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;
- 6 Modelo bayesiano hierarquico(CAR generalizado);

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;
- 6 Modelo bayesiano hierarquico(CAR generalizado);
- 7 Dados;

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;
- 6 Modelo bayesiano hierarquico(CAR generalizado);
- 7 Dados;
- 8 Resultados;

- 1 Estatística espacial e dados de área;
- 2 Vizinhaça e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;
- 6 Modelo bayesiano hierarquico(CAR generalizado);
- 7 Dados;
- 8 Resultados;
- 9 Bibliografia

- 1 Estatística espacial e dados de área;
- 2 Vizinhança e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;
- 6 Modelo bayesiano hierarquico(CAR generalizado);
- 7 Dados;
- 8 Resultados;
- 9 Bibliografia
- 10 Modelos de espaço e tempo para dados de área

- 1 Estatística espacial e dados de área;
- 2 Vizinhança e matriz de pesos;
- 3 Análise de cluster(k-médias)
- 4 Bayes Empírico
- 5 Modelo autoregressivo condicional;
- 6 Modelo bayesiano hierarquico(CAR generalizado);
- 7 Dados;
- 8 Resultados;
- 9 Bibliografia
- 10 Modelos de espaço e tempo para dados de área

Estatística espacial: É um ramo da estatística que analisa dados distribuídos em uma área. Onde geralmente procura-se identificar padrões neste espaço que possam explicar parte da variabilidade dos dados através de um modelo estatístico.

Dados de área: Ela é um caso particular da estatística espacial. Onde dentro da área abrangente existem pequenas subáreas, sendo uma observação em cada uma.

Vizinhança: É um agrupamento das subáreas de uma forma que seja condizente com o contexto do problema e/ou explique a dispersão dos dados no espaço.

Matriz de vizinhança: É uma matriz de n por n , onde n é o número de subáreas, e seja a forma mais simples dada por:

$$w_{ij} = \begin{cases} 1 & \text{se } i \text{ vizinho de } j \\ 0 & \text{se } i \text{ é não vizinho de } j \end{cases} \quad (1)$$

para $i, j = 1, 2, \dots, n$

Análise de cluster(K-médias)

Seja X_l e X_k dois vetores de observações com $l \neq k$, então a distância Euclidiana é da forma:

$$d(X_l, X_k) = [(X_l - X_k)'(X_l - X_k)]^{1/2} = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{1/2} \quad (2)$$

O método de k-médias aloca cada elemento amostral é alocado a um cluster cujo centróide é o mais próximo do vetor de valores observados para o respectivo elemento

Passos do algoritmo:

- Escolha do número de centróides

Análise de cluster(K-médias)

Seja X_l e X_k dois vetores de observações com $l \neq k$, então a distância Euclidiana é da forma:

$$d(X_l, X_k) = [(X_l - X_k)'(X_l - X_k)]^{1/2} = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{1/2} \quad (2)$$

O método de k-médias aloca cada elemento amostral é alocado a um cluster cujo centróide é o mais próximo do vetor de valores observados para o respectivo elemento

Passos do algoritmo:

- Escolha do número de centróides
- Cada elemento amostral é comparado com os centróides, segundo alguma medida de distância

Análise de cluster(K-médias)

Seja X_l e X_k dois vetores de observações com $l \neq k$, então a distância Euclidiana é da forma:

$$d(X_l, X_k) = [(X_l - X_k)'(X_l - X_k)]^{1/2} = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{1/2} \quad (2)$$

O método de k-médias aloca cada elemento amostral é alocado a um cluster cujo centróide é o mais próximo do vetor de valores observados para o respectivo elemento

Passos do algoritmo:

- Escolha do número de centróides
- Cada elemento amostral é comparado com os centróides, segundo alguma medida de distância
- Recalculo dos centróides

Análise de cluster(K-médias)

Seja X_l e X_k dois vetores de observações com $l \neq k$, então a distância Euclidiana é da forma:

$$d(X_l, X_k) = [(X_l - X_k)'(X_l - X_k)]^{1/2} = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{1/2} \quad (2)$$

O método de k-médias aloca cada elemento amostral é alocado a um cluster cujo centróide é o mais próximo do vetor de valores observados para o respectivo elemento

Passos do algoritmo:

- Escolha do número de centróides
- Cada elemento amostral é comparado com os centróides, segundo alguma medida de distância
- Recalculo dos centróides
- Passos 2 e 3 devem ser repetidos até que todos os elementos estejam bem alocados

Análise de cluster(K-médias)

Seja X_l e X_k dois vetores de observações com $l \neq k$, então a distância Euclidiana é da forma:

$$d(X_l, X_k) = [(X_l - X_k)'(X_l - X_k)]^{1/2} = \left[\sum_{i=1}^p (X_{il} - X_{ik})^2 \right]^{1/2} \quad (2)$$

O método de k-médias aloca cada elemento amostral é alocado a um cluster cujo centróide é o mais próximo do vetor de valores observados para o respectivo elemento

Passos do algoritmo:

- Escolha do número de centróides
- Cada elemento amostral é comparado com os centróides, segundo alguma medida de distância
- Recalculo dos centróides
- Passos 2 e 3 devem ser repetidos até que todos os elementos estejam bem alocados

Sejam λ_i e v_i a média e a variância da distribuição a priori de cada taxa θ_i , onde i é a contagem de subáreas, temos então:

$$\hat{\theta}_i = w_i \frac{y_i}{n_i} + (1 - w_i) \lambda_i \quad (3)$$

Onde:

$$w_i = \frac{v_i}{(v_i + \lambda_i / n_i)} \quad (4)$$

Obs: Aqui consideremos a distribuição à priori sendo uma gamma e a verossimilhança dada por uma poisson com os n_{is} como offset.

Modelo autoregressivo condicional

Os modelos autoregressivos espaciais ou modelos auto-normais, são modelos onde uma parte é das covariáveis e outra é uma autoregressão nos resíduos. Assim, temos que:

$$\mathbf{Z} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \quad (5)$$

Onde:

$$\mathbf{e} = \rho \mathbf{W} \mathbf{e} \quad (6)$$

Note que a autoregressão dos erros é ponderada pela matriz de vizinhança $\mathbf{W}$ e o parâmetro de dependência ρ , de forma que a matriz de vizinhança irá 'separar' as observações dos vizinhos na autoregressão.

Modelo autorregressivo condicional

Seja Z_i e Z_j observações de subáreas, de forma que $i \neq j$, então temos que:

$$Z_i|Z_j \sim N(\mu_i + \sum_j W_{ij}(Z_j - \mu_j), \sigma_i^2) \quad (7)$$

Seja $\mathbf{Z}$ o vetor de observações, então a distribuição marginal é dada por:

$$\mathbf{Z} \sim NM(\mathbf{X}\beta, (\mathbf{I} - \rho\mathbf{W})^{-1}\sigma^2\mathbf{I}) \quad (8)$$

Logo temos que a verossimilhança marginal é dada por:

$$-\frac{n}{2}\log(2\pi\sigma^2) + \frac{1}{2}|\mathbf{I} - \rho\mathbf{W}| - \frac{1}{2\sigma^2}(\mathbf{Z} - \mu)'(\mathbf{I} - \rho\mathbf{W})(\mathbf{Z} - \mu) \quad (9)$$

Modelo bayesiano hierarquico (CAR generalizado)

Agora temos a verossimilhança dada pela equação:

$$L(Z_i|\beta, \sigma^2) = \int_{\mathbf{e}} \prod_{i=1}^n f(Z_i|\beta, \sigma^2, \mathbf{e}) f(\mathbf{e}) d\mathbf{e} \quad (10)$$

No caso da adição de mais uma variável latente iid, a verossimilhança é dada por:

$$L(Z_i|\beta, \sigma^2, \nu^2) = \int_{\mathbf{e}} \int_{b_k} \prod_{k=1}^n \prod_{i=1}^n f(Z_i|\beta, \sigma^2, b_k, \mathbf{e}) f(b_k) f(\mathbf{e}) db_k d\mathbf{e} \quad (11)$$

Obs.: Note que $i=k$.

Temos também que:

$$\mathbf{e} \sim NM_n(\mathbf{0}, \Sigma_{\text{car}}) \quad (12)$$

$$b_k \sim N(0, \nu^2) \quad (13)$$

$$\Sigma_{\text{car}} = (\mathbf{I} - \mathbf{W})^{-1} \sigma^2 \mathbf{I} \quad (14)$$

Modelo bayesiano hierarquico (CAR generalizado):

Seja o estimador bayesiano do modelo dado por:

$$f(\beta, \sigma^2, v^2 | Z_i) = \frac{L(Z_i | \beta, \sigma^2, v^2) \pi(\beta) \pi(\sigma^2) \pi(v^2)}{C(Z_i)} \quad (15)$$

Onde a constante $C(Z_i)$:

$$C(Z_i) = \int_{\beta} \int_{\sigma^2} \int_{v^2} L(Z_i | \beta, \sigma^2, v^2) \pi(\beta) \pi(\sigma^2) \pi(v^2) dv^2 d\sigma^2 d\beta \quad (16)$$

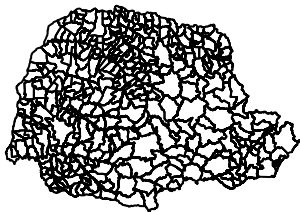
A idéia da modelagem é a suavização de taxas, no caso mortalidade infantil. As observações foram tiradas do DATASUS, referentes ao período de agosto de 2009 a agosto de 2010, sendo compostas por duas variáveis: a contagem de óbitos de crianças com menos de 1 ano de idade, e o número de nascidos vivos no período de referência. Este último entra como offset. Também foram utilizadas mais 12 variáveis para a análise de cluster.

Pacotes utilizados

```
> require(sp)
> require(maptools)
> require(spdep)
> require(RColorBrewer)
> require(tools)
> require(classInt)
> require(stringr)
> require(INLA)
> require(pscl)
```

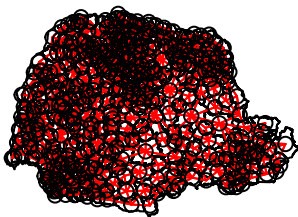
Carregando e plotando polígono:

```
> PRmun <- readShapePoly("C:/Users/Intel I3/Dropbox/FernandoMoro/PR/munpr.shp")  
> plot(PRmun)
```



Gravando e plotando o grafo de vizinhança

```
> viz<-poly2nb(PRmun)
> plot(PRmun)
> plot(viz,coordinates(PRmun),col='red',add=T)
```



Vendo comportamento da razão:

```
> summary(raz)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.00000 0.00000 0.01136 0.01271 0.01816 0.06122

> hist(raz)
```

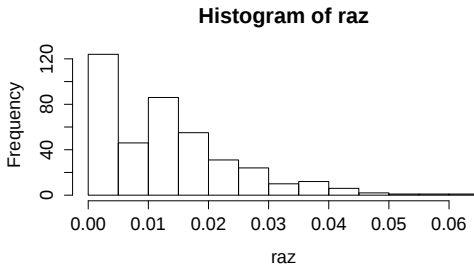


Gráfico das Taxas Bayesianas Empíricas Vs. Taxas Brutas

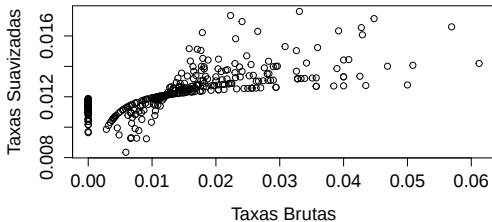


Gráfico legenda de cores

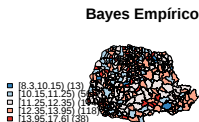


Figura: Legenda de cores para TBE



Figura: Legenda de cores para Taxa Bruta

Teste de Moran's I para resíduos do modelo

```
mdelo0.5<-zeroinfl(obt~ 1,dist='poisson',  
                  link='log',data=data,offset=log(nasc))  
moran.mc(mdelo0.5$residuals,  
listw=nb2listw(viz,style='B'),nsim=1000)
```

Monte-Carlo simulation of Moran's I

```
data:  mdelo0.5$residuals  
weights: nb2listw(viz, style = "B")  
number of simulations + 1: 1001
```

```
statistic = 0.0115, observed rank = 757, p-value = 0.2438  
alternative hypothesis: greater
```

Declaração do modelo

```
nb2INLA('viz1.adj',viz)
formula=obt~1+f(id,graph='viz1.adj',model='besag')
mdelo1<-inla(formula=formula,family='zeroinflated.poisson.1',data=data,
             control.predictor=list(compute=T),
             control.compute=list(dic=TRUE),E=nasc)
```

Sumário do modelo

Time used:

Pre-processing	Running inla	Post-processing	Total
7.0024	2.8225	0.2300	10.0550

Fixed effects:

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-4.3778	0.0274	-4.4319	-4.3777	-4.3242	-4.3776	0

Random effects:

Name	Model
id	Besags ICAR model

Model hyperparameters:

	mean	sd	0.025quant	0.5quant
zero-probability parameter for zero-inflated poisson_1	0.0071	0.0024	0.0032	0.0070
Precision for id	12.7714	3.9620	5.6575	12.7339
	0.975quant		mode	
zero-probability parameter for zero-inflated poisson_1	0.0112		0.0057	
Precision for id	20.1507		12.3677	

Expected number of effective parameters(std dev): 24.79(8.58)

Number of equivalent replicates : 16.10

Deviance Information Criterion: 1377.09

Effective number of parameters: 25.21

Marginal Likelihood: -1002.09

Posterior marginals for linear predictor and fitted values computed

Densidades a posteriori

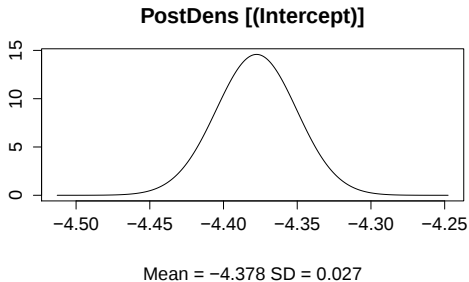


Figura: Gráfico posteriori de β

Densidades a posteriori

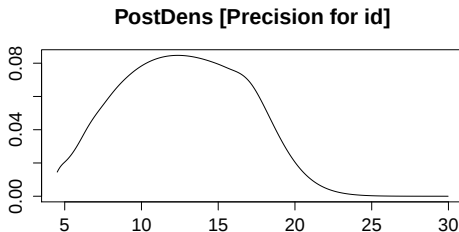


Figura: Gráfico posteriori de τ^2

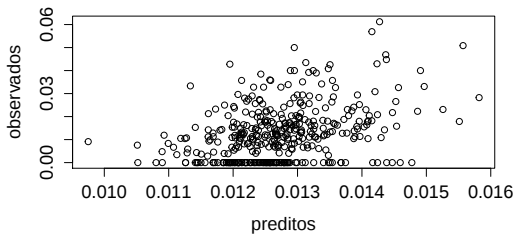


Figura: Preditos Vs. Observados

Comparando Bayes Empírico com preditos do modelo

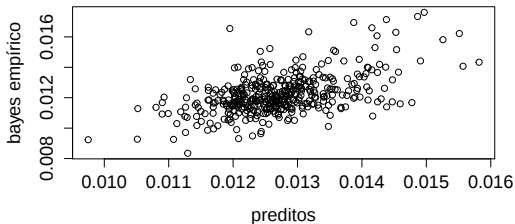


Figura: Bayes Empírico Vs. Preditos do modelo

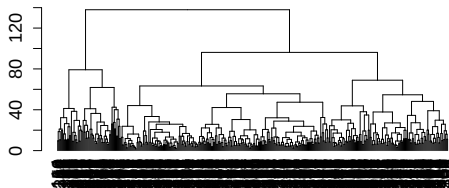


Figura: Gráfico Dendograma

Agrupando as subáreas de 1 a 30 grupos

```
cluster<-list()  
for(i in 1:30){  
  cluster[[i]]<-kmeans(mat1,i)$cluster  
}
```

Construindo matriz de pesos

```
mat<-list()
for(j in 1:30){
  mat[[j]]<-matrix(0,ncol=399,nrow=399)
}
for(h in 1:30){
  for(i in 1:399){
    for(j in 1:399){
      if(i != j){
        ifelse(cluster[[h]][i]==cluster[[h]][j],mat[[h]][i,j]<- 1,
        mat[[h]][i,j]<- 0)
      }
    }
  }
}
```

Passando para lista e calculando o Moran's I

```
pesos2<-list()
for(i in 1:30){
  pesos2[[i]]<-mat2listw(mat[[i]])
}
pval<-vector()
for(i in 1:30){
  pval[i]<-moran.mc(mdelo0.5$residuals,listw=pesos2[[i]],
  nsim=1000)$p.value
}
```

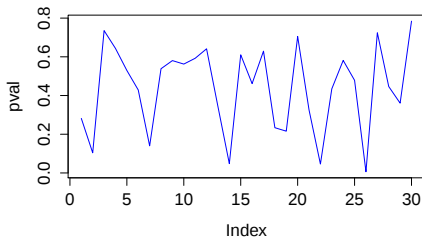


Figura: P-valor para cada grupo definido

Vendo estrutura de vizinhança

```
pval[26]
[1] 0.005994006
pesos2final<-pesos2[[26]]
summary(pesos2final)
Characteristics of weights list object:
Neighbour list object:
Number of regions: 399
Number of nonzero links: 7474
Percentage nonzero weights: 4.694694
Average number of links: 18.73183
Link number distribution:

 5  6  7  8  9 10 11 15 16 17 18 20 21 23 26 27 28 34
18 14  8 27 20 11 24 16 17 18 19 42 22 24 27 28 29 35
18 least connected regions:
38 84 97 114 137 147 162 178 206 242 258 279 288 294 298 300 331 385 with 5 lin
35 most connected regions:
6 14 27 29 31 34 47 52 54 60 78 80 93 100 104 112 115 130 163 174 177 179 181 1
```

Declarando o segundo modelo

```
nb2INLA('viz2.adj',pesos2final$neighbours)
formula2=obt~1+f(id,graph='viz2.adj',model='besag')
mdelo2<-inla(formula=formula2,family='zeroinflated.poisson.1',data=data,
             control.predictor=list(compute=T),
             control.compute=list(dic=TRUE),E=nasc)
```

Resultados

Saída do segundo modelo

Time used:

Pre-processing	Running inla	Post-processing	Total
0.5614	36.0822	0.2994	36.9430

Fixed effects:

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-4.3838	0.0294	-4.4419	-4.3837	-4.3264	-4.3835	0

Random effects:

Name	Model
id	Besags ICAR model

Model hyperparameters:

	mean	sd	0.025quant	0.5quant
zero-probability parameter for zero-inflated poisson_1	0.0164	0.0186	0.0016	0.0106
Precision for id	1.2153	0.0239	1.1689	1.2150
	0.975quant mode			
zero-probability parameter for zero-inflated poisson_1	0.0665	0.0043		
Precision for id	1.2630	1.2145		

Expected number of effective parameters(std dev): 45.93(0.781)

Number of equivalent replicates : 8.687

Deviance Information Criterion: 1361.99

Effective number of parameters: 45.18

Marginal Likelihood: -1244.08

Posterior marginals for linear predictor and fitted values computed

Densidades a posteriori

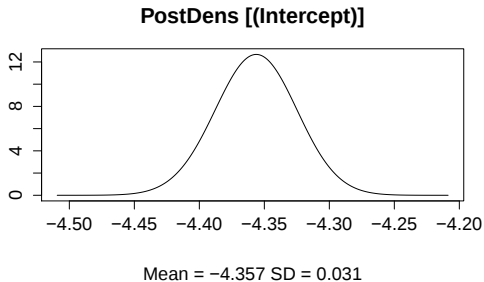


Figura: Gráfico posteriori de β

Densidades a posteriori

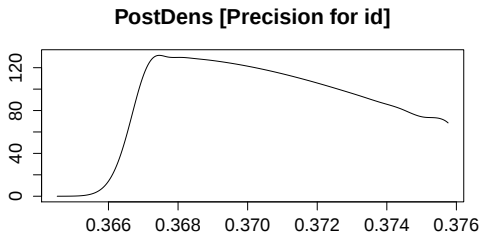


Figura: Gráfico posteriori de τ^2

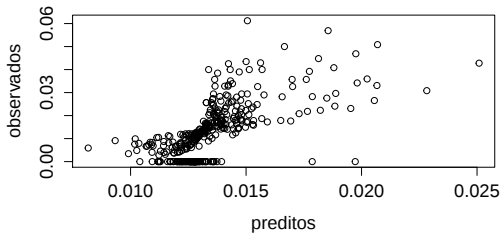


Figura: Preditos Vs. Observados

Comparação bayes emérico com os preditos do modelo

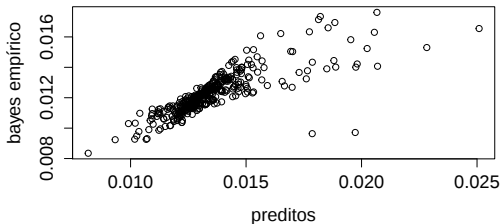


Figura: Bayes Empírico Vs. preditos do modelo

Comparando preditos dos modelos

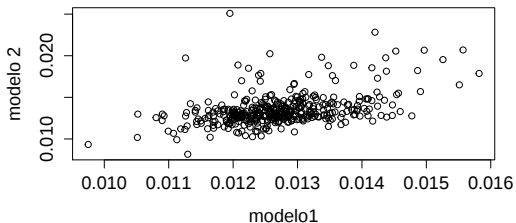


Figura: Modelo 1 Vs. Modelo 2

Modelo com duas variáveis latentes

```
formula3=obt~1+f(id,graph='viz2.adj',model='bym')  
mdelo3<-inla(formula=formula3,family='zeroinflated.poisson.1',data=data,  
             control.predictor=list(compute=T),  
             control.compute=list(dic=TRUE),E=nasc)
```

Resultados

Saída do terceiro modelo

Time used:

Pre-processing	Running inla	Post-processing	Total
0.6199	17.2624	0.3000	18.1823

Fixed effects:

	mean	sd	0.025quant	0.5quant	0.975quant	mode	kld
(Intercept)	-4.3796	0.0325	-4.4446	-4.3792	-4.3167	-4.3784	0

Random effects:

Name	Model
id	BYM model

Model hyperparameters:

	mean	sd	0.025quant
zero-probability parameter for zero-inflated poisson_1	0.0059	0.0036	0.0007
Precision for id (iid component)	20.6701	7.0114	10.0088
Precision for id (spatial component)	1894.4391	1849.3498	130.2254
	0.5quant	0.975quant	mode
zero-probability parameter for zero-inflated poisson_1	0.0054	0.0135	0.0023
Precision for id (iid component)	19.6580	37.2399	17.7832
Precision for id (spatial component)	1352.9846	6788.8534	355.8875

Expected number of effective parameters(std dev): 49.49(13.57)

Number of equivalent replicates : 8.063

Deviance Information Criterion: 1362.32

Effective number of parameters: 50.28

Marginal Likelihood: -879.99

Posterior marginals for linear predictor and fitted values computed

Densidades a posteriori

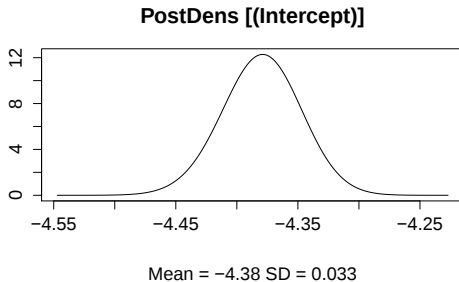


Figura: Gráfico posteriori de β

Densidades a posteriori

PostDens [Precision for id (spatial component)]

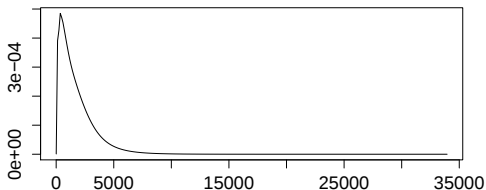


Figura: Gráfico posteriori de τ^2

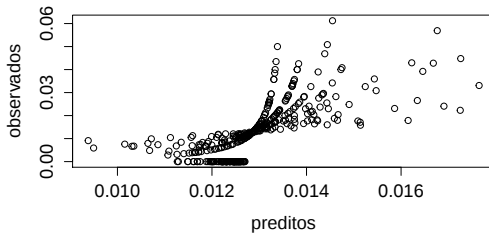


Figura: Preditos Vs. Observados

Comparação bayes emérico com os preditos do modelo

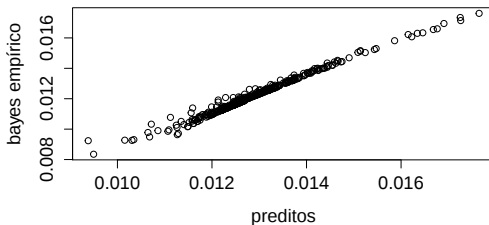


Figura: Bayes Empírico Vs. preditos do modelo

Comparando preditos dos modelos

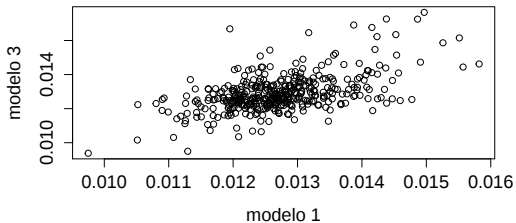


Figura: Modelo 3 Vs. Modelo 1

Comparando preditos dos modelos

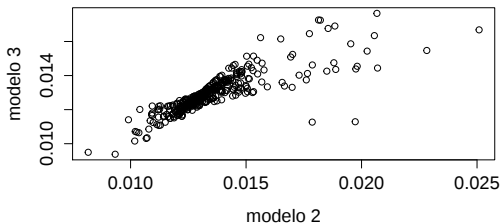


Figura: Modelo 3 Vs. Modelo 2

- 1 Schabenberger, O., Gotway, C.A.(2005) *Statistical Methods for Spatial Data Analysis*. Chapman Hall/CRC, New York.

- ① Schabenberger, O., Gotway, C.A.(2005) *Statistical Mhetods for Spatial Data Analysis*. Chapman Hall/CRC, New York.
- ② Bailey, T.C., Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman Scientific Technical, England.

- ① Schabenberger, O., Gotway, C.A.(2005) *Statistical Mhetods for Spatial Data Analysis*. Chapman Hall/CRC, New York.
- ② Bailey, T.C., Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman Scientific Technical, England.
- ③ Mingoti, S.A. (2013) *Análise de Dados Através de Métodos de Estatística Multivariada*. EDITORAufmg, Belo Horizonte.

- 1 Schabenberger, O., Gotway, C.A.(2005) *Statistical Mhetods for Spatial Data Analysis*. Chapman Hall/CRC, New York.
- 2 Bailey, T.C., Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman Scientific Technical, England.
- 3 Mingoti, S.A. (2013) *Análise de Dados Através de Métodos de Estatística Multivariada*. EDITORAufmg, Belo Horizonte.
- 4 Casella, G., Berger, R.L. (2010) *Inferência Estatística*. CENGAGE learning, São Paulo.

- 1 Schabenberger, O., Gotway, C.A. (2005) *Statistical Methods for Spatial Data Analysis*. Chapman Hall/CRC, New York.
- 2 Bailey, T.C., Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman Scientific Technical, England.
- 3 Mingoti, S.A. (2013) *Análise de Dados Através de Métodos de Estatística Multivariada*. EDITORAufmg, Belo Horizonte.
- 4 Casella, G., Berger, R.L. (2010) *Inferência Estatística*. CENGAGE learning, São Paulo.
- 5 Ribeiro Jr., P.J., Bonat, W.B., Teixeira Krainski, E., Marques Zeviani, W. (2012) *Métodos Computacionais em Inferência Estatística*. 20ª SINAPSE, Curitiba.

- 1 Schabenberger, O., Gotway, C.A. (2005) *Statistical Methods for Spatial Data Analysis*. Chapman Hall/CRC, New York.
- 2 Bailey, T.C., Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman Scientific Technical, England.
- 3 Mingoti, S.A. (2013) *Análise de Dados Através de Métodos de Estatística Multivariada*. EDITORAufmg, Belo Horizonte.
- 4 Casella, G., Berger, R.L. (2010) *Inferência Estatística*. CENGAGE learning, São Paulo.
- 5 Ribeiro Jr., P.J., Bonat, W.B., Teixeira Krainski, E., Marques Zeviani, W. (2012) *Métodos Computacionais em Inferência Estatística*. 20^a SINAPSE, Curitiba.
- 6 Ribeiro, Jr., P.J., Bonat, W. B., Marques Zeviani, W. (2013) *Likelihood analysis for a class of beta mixed models*. LEG/DEST, Curitiba.

- ❶ Schabenberger, O., Gotway, C.A. (2005) *Statistical Methods for Spatial Data Analysis*. Chapman Hall/CRC, New York.
- ❷ Bailey, T.C., Gatrell, A.C. (1995) *Interactive Spatial Data Analysis*. Longman Scientific Technical, England.
- ❸ Mingoti, S.A. (2013) *Análise de Dados Através de Métodos de Estatística Multivariada*. EDITORAufmg, Belo Horizonte.
- ❹ Casella, G., Berger, R.L. (2010) *Inferência Estatística*. CENGAGE learning, São Paulo.
- ❺ Ribeiro Jr., P.J., Bonat, W.B., Teixeira Krainski, E., Marques Zeviani, W. (2012) *Métodos Computacionais em Inferência Estatística*. 20^a SINAPSE, Curitiba.
- ❻ Ribeiro, Jr., P.J., Bonat, W. B., Marques Zeviani, W. (2013) *Likelihood analysis for a class of beta mixed models*. LEG/DEST, Curitiba.

Modelo espaço tempo para dados de áreas

Seja Y_{st} a variável resposta avaliada no tempo t e na área s , também seja o preditor linear eta_{st} , então temos o modelo geral dado por:

$$\eta_{ij} = \mu + \alpha_t + \gamma_t + \theta_s + \phi_s + \psi_{st} \quad (17)$$

Onde μ é o termo que depende das covariáveis, α_t e γ_t são as variáveis latentes temporais com e sem estrutura temporal à priori, o mesmo com θ_s e ϕ_s . O termo ψ_{st} é o processo bidimensional. Para γ_t e ϕ_s adota-se prioris sem as estruturas temporais e espaciais, já α_t e θ_s adota-se prioris com estrutura temporal e espacial, respectivamente. O parâmetro ψ_{st} adotata-se uma normal multivariada da seguinte forma:

$$\psi_{st} \sim NM(\mathbf{0}, \Sigma) \quad (18)$$

Onde a matriz Σ^{-1} é a matriz de precisão, definida da seguinte forma:

$$\Sigma^{-1} = \mathbf{F} \otimes (\mathbf{I} - \rho \mathbf{W}) \mathbf{M}^{-1} \quad (19)$$

Onde $\mathbf{F}$ é a estrutura temporal, $\mathbf{W}$ é a matriz de pesos de vizinhança, e $\mathbf{M} = \text{diag}(\sigma_s^2)$.