

Modelos de Regressão Linear Mistos para dados discretos: Uma abordagem utilizando MCMC através do Stan integrado ao R

Felipe Vieira - RA 160424 Guilherme Artoni - RA 160318

Degeneração macular relacionada à idade (DMRI)

- É uma doença atualmente sem cura;
- Leva a perda progressiva da visão central;
- Visão turva é o principal sintoma;
- É muito comum em pessoas com mais de 55 anos.

Dados

- Grupo de Estudo de Terapias Farmacológicas para Degeneração Macular;
- O objetivo é avaliar se um novo medicamento para DMRI tem poder competitivo com o principal existente no mercado;
- Ensaios clínicos aleatorizados;
- Realizados em diferentes centros de estudos;
- 240 pacientes;
- Foi medida a qualidade da visão de todos os pacientes no início e após 4, 12, 24 e 52 semanas.

Análise Descritiva

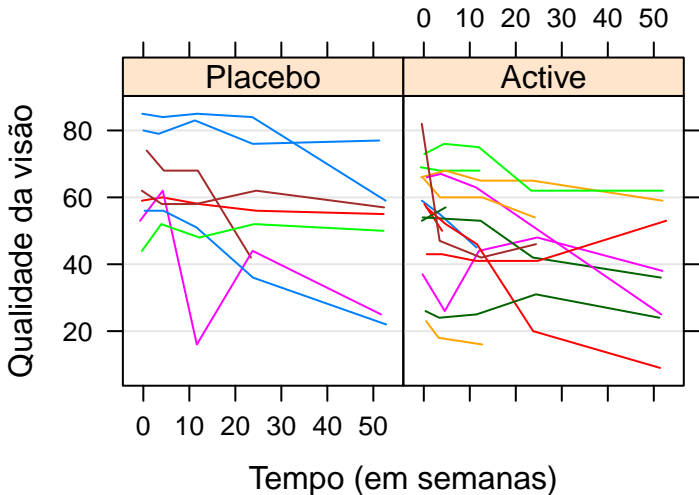


Figura 1: Gráfico de perifs de alguns indivíduos selecionados aleatoriamente

Análise Descritiva

Tabela 1: Número de observações, médias e medianas de cada semana observada.

	P:n	A:n	P:Média	A:Média	P:Mediana	A:Mediana
Início	119	121	55.34	54.58	56.0	57.0
4smn	117	114	53.97	50.91	54.0	52.0
12smn	117	110	52.87	48.67	53.0	49.5
24smn	112	102	49.33	45.46	50.5	45.0
52smn	105	90	44.44	39.10	44.0	37.0

Análise Descritiva

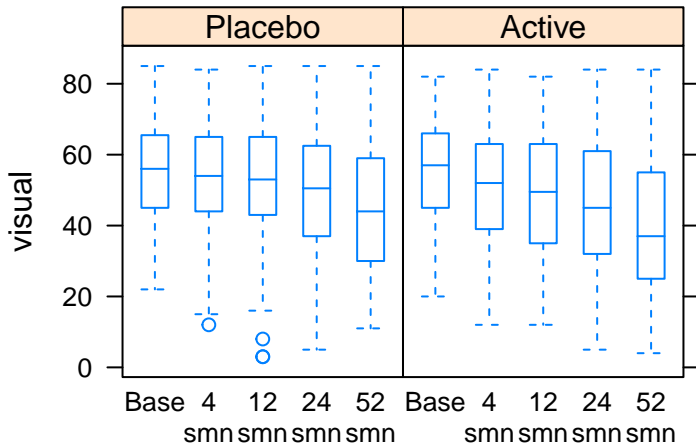


Figura 2: Boxplots das medidas de qualidade da visão de cada tempo observado.

Análise Descritiva

- Matriz de Variâncias e Covariâncias**

##		Início	4smn	12smn	24smn	52smn
##	Início	220.31	206.71	196.24	193.31	152.71
##	4smn	206.71	246.22	224.79	221.27	179.23
##	12smn	196.24	224.79	286.21	257.77	222.68
##	24smn	193.31	221.27	257.77	334.45	285.23
##	52smn	152.71	179.23	222.68	285.23	347.43

Análise Descritiva

- **Matriz de Correlações**

##	Início	4smn	12smn	24smn	52smn
## Início	1.00	0.89	0.78	0.71	0.55
## 4smn	0.89	1.00	0.85	0.77	0.61
## 12smn	0.78	0.85	1.00	0.83	0.71
## 24smn	0.71	0.77	0.83	1.00	0.84
## 52smn	0.55	0.61	0.71	0.84	1.00

Modelo Misto

$$\mathbf{Y}_{j(k_j \times 1)} = \mathbf{X}_{j(k_j \times p)}\boldsymbol{\beta}_{(p \times 1)} + \mathbf{Z}_{j(k_j \times q)}\mathbf{b}_{j(q \times 1)} + \boldsymbol{\xi}_{j(k_j \times 1)}$$

Onde $j = 1, 2, \dots, n$ é o indivíduo

- $\mathbf{Y}_j = (y_{j1}, \dots, y_{jk_j})$: vetor resposta, no qual k_j é o número de avaliações realizadas no indivíduo j .
- \mathbf{X}_j : matriz de planejamento associada aos efeitos fixos para o indivíduo j .
- $\boldsymbol{\beta}$: vetor de efeitos fixos
- \mathbf{Z}_j : matriz de planejamento associada aos efeitos aleatórios para o indivíduo j .
- \mathbf{b}_j : vetor de efeitos aleatórios associado ao indivíduo j .
- $\boldsymbol{\xi}_j$: vetor de erros associado ao indivíduo j .

Modelo Proposto

$$Y_{it} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2it} + \beta_3 x_{3i} + \beta_4 x_{2it} x_{3i} + b_{0i} + \xi_{it},$$

- Y_{it} é a qualidade da visão do paciente i ($i = 1, \dots, 240$) no tempo t ($t = 1, 2, 3, 4$, correspondendo aos valores 4º, 12º, 24º e 52º semana, respectivamente);
- x_{1i} é o valor inicial da qualidade da visão;
- x_{2it} é o tempo t de medição no paciente i ;
- x_{3i} é o indicador do tratamento, 0 se placebo e 1 caso contrário;
- $x_{2it} x_{3i}$ é a interação entre as duas covariáveis.

Modelo Proposto

$$Y_{it} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2it} + \beta_3 x_{3i} + \beta_4 x_{2it} x_{3i} + b_{0i} + \xi_{it},$$

- β_0 é o intercepto geral;
- β_1 é o incremento no valor esperado de Y_{it} quando variado em uma unidade o valor inicial da qualidade da visão;
- β_2 é o incremento no valor esperado de Y_{it} , quando acrescido o tempo em uma semana entre as que foram observadas;
- β_3 é o efeito geral no valor esperado de Y_{it} causado pelo tratamento;
- β_4 é o incremento sobre o valor esperado de Y_{it} , gerado pela variação do tempo em uma semana entre as que foram observadas sobre o paciente i que estava sob tratamento.

Modelo Proposto

$$Y_{it} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2it} + \beta_3 x_{3i} + \beta_4 x_{2it} x_{3i} + b_{0i} + \xi_{it},$$

- b_{0i} é o efeito aleatório específico para cada paciente. Tal que $b_{0i} \stackrel{iid}{\sim} \mathcal{N}(0, \tau) \forall i$;
- ξ_{it} é o erro aleatório. Tal que $\xi_{it} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2) \forall i$ e t ;
- b_{0i} representa uma variação específica do β_0 para cada paciente.

Modelo Proposto

- Forma matricial do modelo para o indivíduo i

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{pmatrix} = \begin{pmatrix} 1 & x_{1i} & 4 & x_{3i} & 4x_{3i} \\ 1 & x_{1i} & 12 & x_{3i} & 12x_{3i} \\ 1 & x_{1i} & 24 & x_{3i} & 24x_{3i} \\ 1 & x_{1i} & 52 & x_{3i} & 52x_{3i} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} b_{0i} + \begin{pmatrix} \xi_{i1} \\ \xi_{i2} \\ \xi_{i3} \\ \xi_{i4} \end{pmatrix}$$

Modelo Proposto

- Estrutura Uniforme da matriz de Variâncias e Covariâncias

$$\begin{pmatrix} \sigma^2 + \tau & \tau & \tau & \tau \\ \tau & \sigma^2 + \tau & \tau & \tau \\ \tau & \tau & \sigma^2 + \tau & \tau \\ \tau & \tau & \tau & \sigma^2 + \tau \end{pmatrix}$$

Inferência

Cadeias de Markov (CM)

- Sequência de variáveis aleatórias discretas
- Espaço de estados denotado por $S = \{s_1, s_2, \dots, s_k\}$

Se vale ,

$$p(X_t | X_{t-1}, X_{t-2}, \dots, X_0) = p(X_t | X_{t-1}),$$

para $t = 1, 2, \dots$, a sequência de variáveis aleatórias é uma CM.

A probabilidade que o processo mude de $X_t = s_i$ para $X_{t+1} = s_j$ é dada pela matriz de transição, $P = \{p_{ij}\}$.

A restrição natural sobre a matriz de transição é que a soma das linhas seja 1, $\sum_j p_{ij} = 1$, para todo i .

Uma CM admite uma distribuição de equilíbrio π se existir π tal que $\pi = \pi P$.

Inferência

- **Irredutibilidade**

Uma CM, cujas as variáveis aleatórias são discretas, é dita irredutível se for uma cadeia em que partindo-se de um estado qualquer, pode-se atingir qualquer estado, inclusive o inicial, em um número finito de transições.

- **Reversibilidade**

Uma CM é dita reversível se a probabilidade de estar em um estado s_i e mover-se para s_j é igual à probabilidade de estar no estado s_j e mover-se para o estado s_i , ou seja,

$$\pi(i)p(i, j) = \pi(j)p(j, i)$$

A reversibilidade de uma CM está associada ao fato da mesma admitir uma distribuição de equilíbrio.

Inferência

- **Amostrador de Gibbs**

O Amostrador de Gibbs (AG) é um método para aproximar uma distribuição multivariada tomando somente amostras de distribuições univariadas.

Suponha que a distribuição a posteriori tenha função de densidade $\pi(\theta_1, \theta_2, \dots, \theta_p)$ para os p parâmetros $\theta_1, \theta_2, \dots, \theta_p$ e seja $\pi(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_p)$ a função densidade condicional para θ_i dado os valores dos outros parâmetros.

Inferência

- **Amostrador de Gibbs**

Tomando arbitrariamente valores iniciais $\{\theta_1(0), \theta_2(0), \dots, \theta_p(0)\}$ para os p parâmetros e em seguida mudá-los um a um selecionando novos valores como segue:

$\theta_1(1)$ é escolhido de $\pi(\theta_1|\theta_2(0), \theta_3(0), \dots, \theta_p(0))$

$\theta_2(1)$ é escolhido de $\pi(\theta_2|\theta_1(1), \theta_3(0), \dots, \theta_p(0))$

$\theta_3(1)$ é escolhido de $\pi(\theta_3|\theta_1(1), \theta_2(1), \theta_4(0) \dots, \theta_p(0))$

\vdots

$\theta_p(1)$ é escolhido de $\pi(\theta_p|\theta_1(1), \theta_2(1), \dots, \theta_{p-1}(1))$

Inferência

- **Metropolis - Hastings**

Deve-se construir um núcleo de transição $p(\theta, \phi)$ de forma que π seja a distribuição de equilíbrio da cadeia. Uma forma simples de fazer isso é através de cadeias onde o núcleo p satisfaça

$$\pi(\theta)p(\theta, \phi) = \pi(\phi)p(\phi, \theta), \quad \forall \quad (\theta, \phi)$$

Embora não seja necessária, ela é suficiente para que π seja a distribuição de equilíbrio da cadeia.

Inferência

- **Metropolis - Hastings**

O núcleo $p(\theta, \phi)$ é constituído de 2 elementos: um núcleo de transição arbitrário $q(\theta, \phi)$ e uma probabilidade $\alpha(\theta, \phi)$ de forma que,

$$p(\theta, \phi) = q(\theta, \phi)\alpha(\theta, \phi), \quad \text{se } \theta \neq \phi$$

Portanto, o núcleo de transição define uma densidade $p(\theta, .)$ para todos os valores diferentes de θ . Consequentemente, resta uma probabilidade positiva da cadeia ficar em θ dada por

$$p(\theta, \theta) = 1 - \int q(\theta, \phi)\alpha(\theta, \phi)d\phi$$

A expressão mais comum para a probabilidade de aceitação é

$$\alpha(\theta, \phi) = \min\left\{1, \frac{\pi(\phi)q(\phi, \theta)}{\pi(\theta)q(\theta, \phi)}\right\}$$

Inferência

- **Metropolis - Hastings**

Em termos práticos a simulação de uma amostra de π usando a CM:

- i) inicialize o contador de iterações da cadeia $j = 1$ e tome um valor arbitrário para $\theta(0)$;
- ii) mova a cadeia para um novo valor ϕ gerado da densidade $q(\theta(j - 1), \cdot)$;
- iii) calcule a probabilidade de aceitação do movimento $\alpha(\theta(j - 1), \phi)$. Se o movimento for aceito, $\theta(j) = \phi$, caso contrário $\theta(j) = \theta(j - 1)$ e a cadeia não se move;
- iv) mude o contador de j para $j + 1$ e retorne ao item ii) até a convergência.

Inferência

- **Metropolis - Hastings**

A etapa iii) é realizada após a geração de uma quantidade uniforme $u \sim U(0, 1)$ independente de todas as outras variáveis. Se $u \leq \alpha$, o movimento é aceito e se $u > \alpha$ o movimento não é permitido.

Inferência

- **Metropolis-within-Gibbs**

O AG pode ser visto como um caso particular do Metropolis-Hastings (MH) onde as distribuições geradoras de candidatos são as condicionais da distribuição de interesse e o candidato é aceito com probabilidade 1.

Considere a situação em que para um vetor aleatório com K elementos, sendo que alguns deles tem densidades condicionais bem definidas e fáceis de gerar, enquanto que os demais não tem esta facilidade.

Neste caso podemos fazer uso combinado do AG para as variáveis fáceis de gerar e do MH para as demais.

Resultados

$$Y_{it} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2it} + \beta_3 X_{3i} + \beta_4 X_{2it} X_{3i} + b_{0i} + \xi_{it},$$

Tabela 2: Resultados do ajuste do modelo.

	Média	DP da média	DP	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
beta[1]	9.27	0.07	2.68	4.12	7.44	9.24	11.09	14.52	1683.11	1
beta[2]	0.83	0.00	0.04	0.74	0.80	0.83	0.86	0.91	1677.70	1
beta[3]	-0.21	0.00	0.02	-0.26	-0.23	-0.21	-0.20	-0.17	6273.53	1
beta[4]	-2.42	0.03	1.53	-5.40	-3.45	-2.39	-1.38	0.64	1968.00	1
beta[5]	-0.05	0.00	0.03	-0.12	-0.07	-0.05	-0.03	0.02	6350.55	1
sigma_e	8.65	0.00	0.25	8.19	8.48	8.64	8.82	9.15	5098.08	1
sigma_p	9.01	0.01	0.53	8.02	8.65	8.99	9.36	10.11	4597.83	1

Resultados

Estimativa da Matriz de Variâncias e Covariâncias Condicionais:

$$\begin{pmatrix} 74.82 & 0 & 0 & 0 \\ 0 & 74.82 & 0 & 0 \\ 0 & 0 & 74.82 & 0 \\ 0 & 0 & 0 & 74.82 \end{pmatrix}$$

Resultados

**Estimativa da Matriz de Variâncias e Covariâncias
Marginais:**

$$\begin{pmatrix} 156 & 81.18 & 81.18 & 81.18 \\ 81.18 & 156 & 81.18 & 81.18 \\ 81.18 & 81.18 & 156 & 81.18 \\ 81.18 & 81.18 & 81.18 & 156 \end{pmatrix}$$

Resultados

Estimativa da Matriz de Correlações:

$$\begin{pmatrix} 1 & 0.52 & 0.52 & 0.52 \\ 0.52 & 1 & 0.52 & 0.52 \\ 0.52 & 0.52 & 1 & 0.52 \\ 0.52 & 0.52 & 0.52 & 1 \end{pmatrix}$$

Resultados

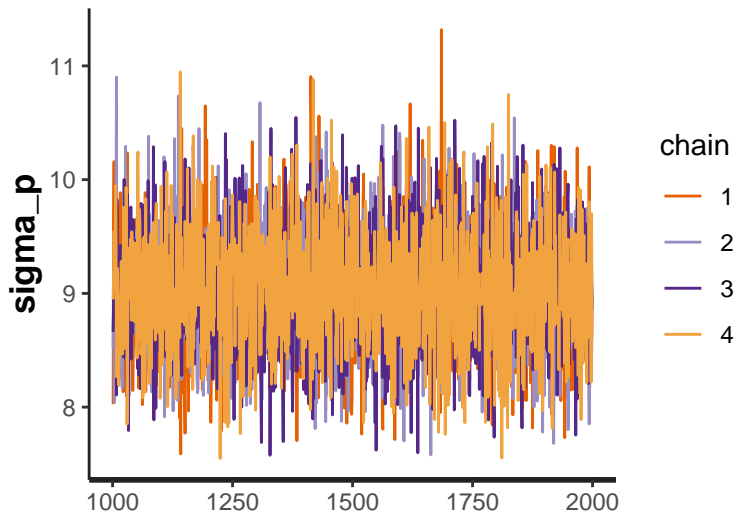


Figura 3: Gráfico mostrando o resultado sobre convergência da CM, após descartado o warm-up.

Considerações Finais

- A partir da análise descritiva e ajuste do modelo verificamos que os pacientes perderam parte da acuidade visual durante o tempo do experimento o que vai de acordo com a característica da doença;
- Essa perda foi mais grave entre os pacientes que estavam sob o tratamento do *interferon*— α ;
- Podemos ver que a metodologia de inferência proposta se adequou a estrutura do modelo, visto que obtivemos a convergência das sequência de forma consistente.