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

Guilherme Artoni - RA 160318 Felipe Vieira - RA 160424 Student - RAXXXXXX
Malba Tahan - RAXXXXXX

The Stan project develops a probabilistic programming language that implements full Bayesian statistical inference via Markov Chain Monte Carlo, rough Bayesian inference via 'variational' approximation, and (optionally penalized) maximum likelihood estimation via optimization. In all three cases, automatic differentiation is used to quickly and accurately evaluate gradients without burdening the user with the need to derive the partial derivatives.

1 Introdução

A degeneração macular relacionada à idade (DMRI) é uma doença atualmente sem cura que ocorre em uma parte da retina chamada mácula e que leva a perda progressiva da visão central. A DMRI é uma alteração muito comum em pessoas com mais de 55 anos, sendo a causa mais frequente de baixa acuidade visual nessa faixa etária. Com o intuito de avaliar se um novo medicamento para DMRI tem poder competitivo com o principal existente atualmente, iremos comparar por meio de modelos de regressão linear a qualidade da acuidade visual de pacientes com DMRI sob ambos tratamentos durante aproximadamente dois meses.

MOTIVAÇÃO SOBRE O MODELO:

- Modelo com efeitos fixos e aleatórios.
- Considerar variabilidade de perfis individuais.
- Estudar e estimar componentes da variância descompactando e adicionando complexidade aos erros.
- Resolver problemas de interdependência dos dados.

Importante em diversas disciplinas

- ANOVA com efeitos mistos (Estatística, Econometria)
- Modelos lineares hierárquicos (Educação)
- Modelos de efeitos contextuais (Sociologia)

MOTIVAÇÃO SOBRE OS DADOS/EXPERIMENTO

- Dados disponibilizados pelo Grupo de Estudo de Terapias Farmacológicas para Degeneração Macular (GETFDM) em 1997.
- Trata-se de informações sobre ensaios clínicos aleatorizados realizados em diferentes centros de estudos.
- O objetivo era comparar um tratamento experimental chamado *interferon-*α e o placebo para pacientes diagnosticados com Degeneração Macular Relacionada a Idade (DMRI).
- Os dados mostrados são em relação ao placebo e a maior dose administrada do interferon-α.

MOTIVAÇÃO SOBRE A DOENÇA

- Pacientes com DMRI gradativamente perdem a visão.
- Durante os ensaios, a qualidade da visão de cada um dos 240 pacientes foi medida no início e após 4, 12, 24 e 52 semanas.
- A qualidade da visão foi medida através da quantidade de letras que os pacientes foram capazes de ler em gráficos de visão padronizados.
- Segue que temos dados longitudinais para cada paciente em forma de medidas da qualidade de sua visão.

MOTIVAÇÃO SOBRE MODELOS MISTOS

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

Onde j = 1,2,...,n é o individuo

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

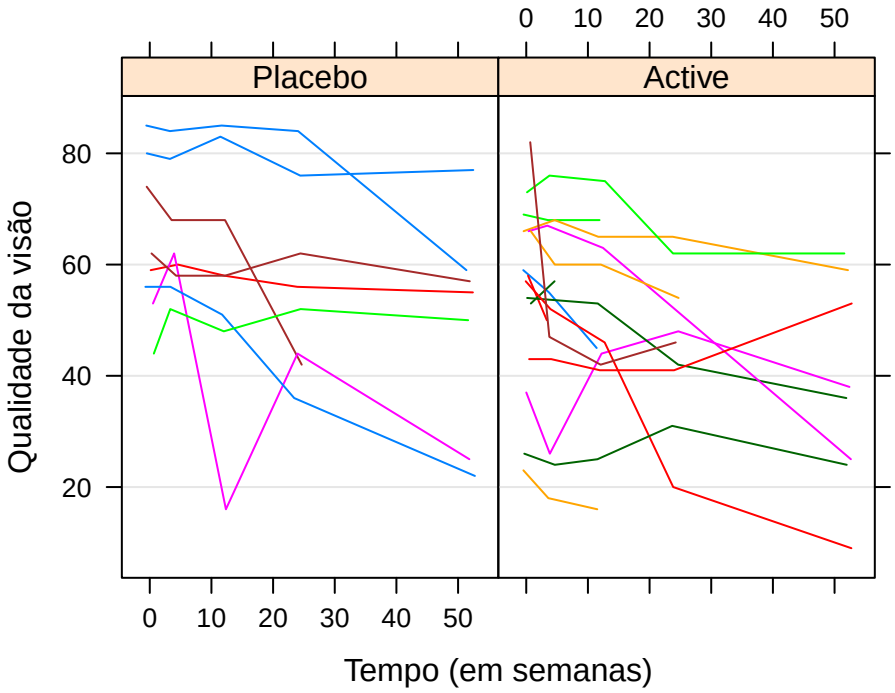
$$\boldsymbol{b}_j \sim \boldsymbol{\mathcal{N}}_k(\boldsymbol{0}, \boldsymbol{\mathcal{D}}) \;\; e \;\; \boldsymbol{\xi}_j \sim \boldsymbol{\mathcal{N}}_{n_i}(\boldsymbol{0}, \boldsymbol{\Sigma_j})$$

Contagem dos dados não faltantes		
Tempo	Placebo	Active
Início	119	121
4º semana	117	114
12º semana	117	110
24º semana	112	102
52º semana	105	90

Tabela 1: Contagem das vezes que os pacientes foi fazer a medição da qualidade da visão.

2 Metodologia

2.1 Análise Descritiva



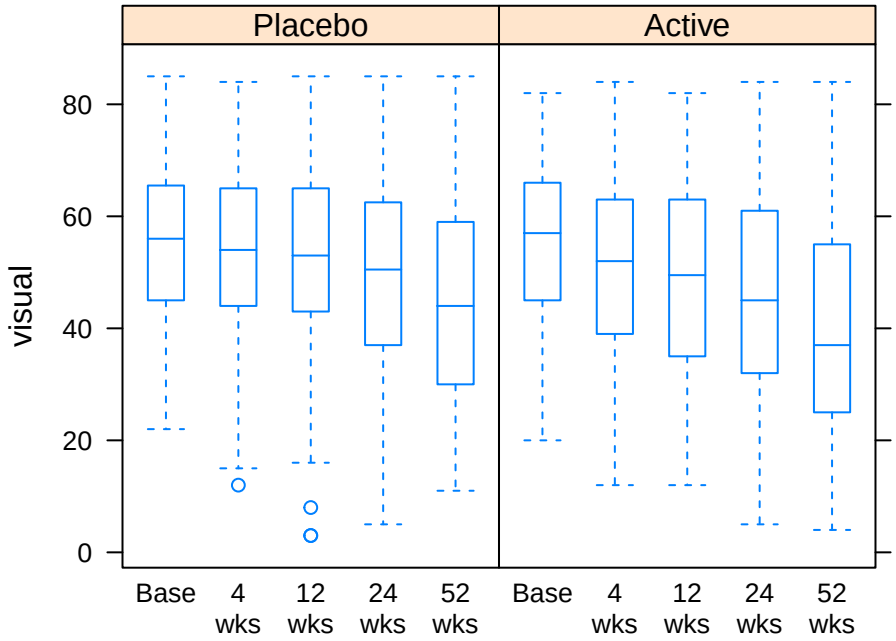
O experimento iniciou-se com 240 pacientes sendo que 119 receberam um placebo e 121 receberam a droga. Com o passar das semanas a quantidade de pessoas no estudo foram diminuindo, isso ocorreu por conta de efeitos colaterais sentidos em ambos os tratamentos. Por conta disso, ao fim do experimento obtivemos alguns dados faltantes conforme mostra a tabela 1.

Podemos observar um decrescimento ao longo do tempo das médias e medianas da medida de qualidade da visão conforme mostra a figura Y (BOXPLOT). Nota-se também um aumento da variabilidade dos dados coletados nas últimas semanas, o aumento no número de dados faltantes pode ser uma possível causa para o crescimento dessa variabilidade. Há também um forte indício de simetria nas distribuições de ambos tratamentos por conta de que médias e medianas apresentaram valores próximos conforme indicado na tabela 2.

Médias e medianas amostrais das medidas de qualidade da visão

Tempo	Placebo			Active		
	Núm. de Indiv.	Média	Mediana	Núm. de Indiv.	Média	Mediana
Início	119	55,34	56,0	121	54,58	57,0
4º semana	117	53,97	54,0	114	50,91	52,0
12º semana	117	52,87	53,0	110	48,67	49,5
24º semana	112	49,33	50,5	102	45,46	45,0
52º semana	105	44,44	44,0	90	39,10	37,0

Tabela 2: Médias e medianas amostrais das medidas de qualidade da visão.



A partir da matriz de variância e covariância $\boldsymbol{\Sigma}$ observa-se que há um aumento da variabilidade dos dados coletados nas últimas semanas, em concordância com as informações contidas nos boxplots. Considerando as correlações $\mathbf{D}$, estas sugerem uma forte ou moderada correlação entre os tratamentos e uma diminuição entre as últimas medidas tomadas, possivelmente consequência dos dados faltantes.

- **Matriz de variancias e covariancias e matriz de correlações amostrais:**

$$\boldsymbol{\Sigma} = \begin{pmatrix} 220.31 & 206.71 & 196.24 & 193.31 & 152.71 \\ 206.71 & 246.22 & 224.79 & 221.27 & 179.23 \\ 196.24 & 224.79 & 286.21 & 257.77 & 222.68 \\ 193.31 & 221.27 & 257.77 & 334.45 & 285.23 \\ 152.71 & 179.23 & 222.68 & 285.23 & 347.43 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} 1.00 & 0.89 & 0.78 & 0.71 & 0.55 \\ 0.89 & 1.00 & 0.85 & 0.77 & 0.61 \\ 0.78 & 0.85 & 1.00 & 0.83 & 0.71 \\ 0.71 & 0.77 & 0.83 & 1.00 & 0.84 \\ 0.55 & 0.61 & 0.71 & 0.84 & 1.00 \end{pmatrix}$$

2.2 Modelagem

3 MODELO NORMAL INDEPENDENTE HOMOCEDASTICO

Vamos considerar o seguinte modelo:

Ajuste do Modelo, estimativas pontuais e intervalares

Parâmetros	Estimativa	EP	Estat-t	IC(95%)	p-valor
β_1	0.83	0.03	29.21	[0,77 ; 0,89]	$<$ 2.2e-16
β_{01}	8.08	1.94	4.16	[4,26 ; 11,89]	3.6e-05
β_{02}	7.08	1.94	3.65	[3,27 ; 10,89]	$<$ 0.001
β_{03}	3.63	1.95	1.86	[-0,20 ; 7,46]	0.063
β_{04}	-1.75	1.99	-0.88	[-5,65 ; 2,16]	0.380
β_{21}	-2.35	1.63	-1.44	[-5,55 ; 0,84]	0.149
β_{22}	-3.71	1.64	-2.26	[-6,93 ; -0,48]	0.024
β_{23}	-3.45	1.69	-2.04	[-6,77 ; -0,12]	0.042
β_{24}	-4.47	1.78	-2.52	[-7,96 ; -0,98]	0.012
σ	12.38				

Tabela 3: Estimativas pontuais dos parâmetros.

Análise de Variância

FV	GL	SQ	QM	Estat-F	p-valor
Est. Inic.	1	2165776	2165776	14138.99	$<$ 2.2e-16
Tempo	4	14434	3608	23.56	$<$ 2.2e-16
Tratamento	4	2703	676	4.41	0.002
Resíduos	858	131426	153		

Tabela 4: Análise de Variância com teste - F sequencial.

$$Y_{it} = \beta_{0t} + \beta_1 x_{1i} + \beta_{2t} x_{2i} + \xi_{it}, \qquad (3)$$

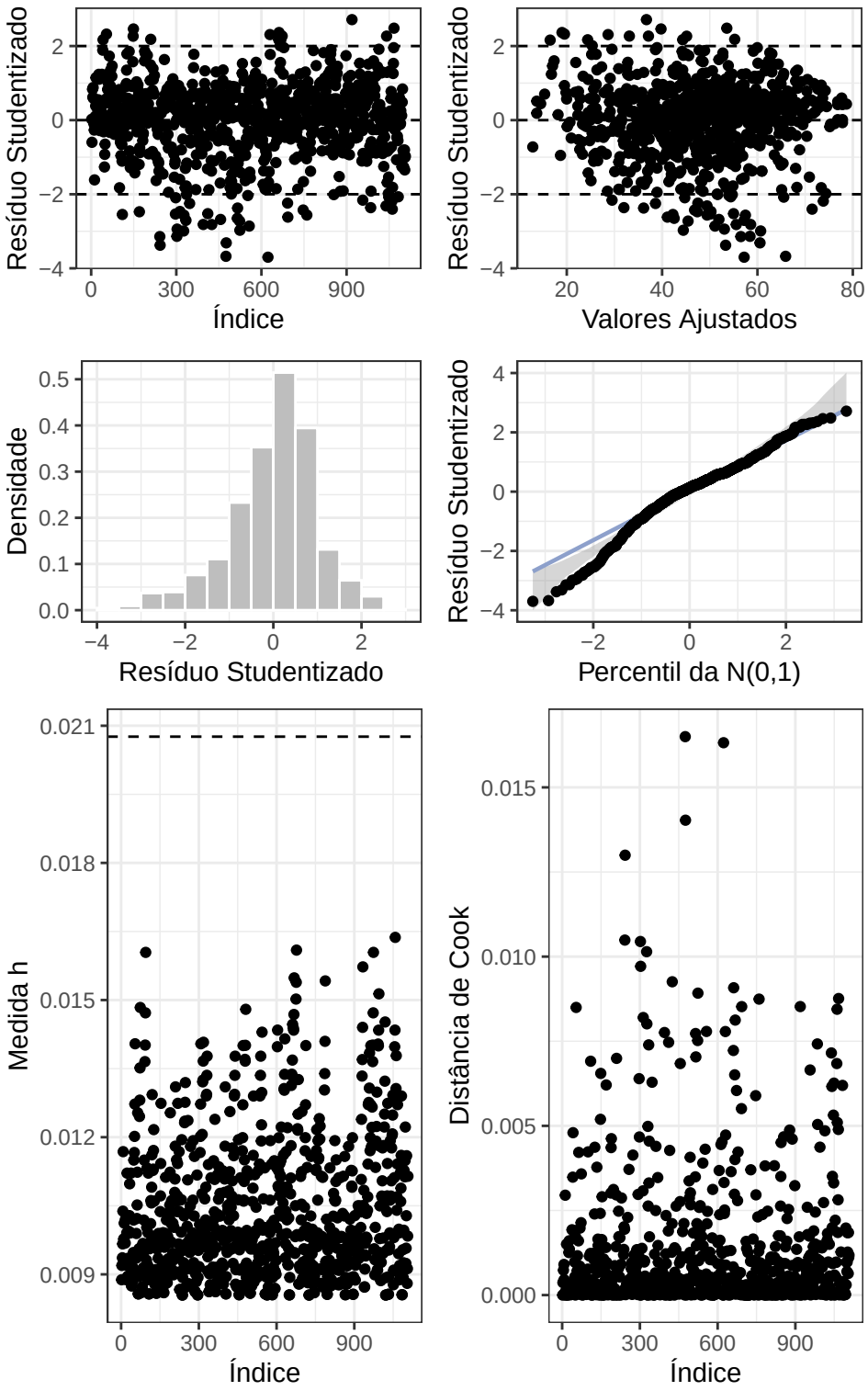
- Y_{it} é o valor da qualidade da visão medido no paciente i (i = 1, ..., 240) no tempo t (t = 1, 2, 3, 4, correspondendo ao valores 4, 12, 24 e 52 semanas, respectivamente).
- x_{1i} é valor medido inicialmente da qualidade da visão.
- x_{2i} é o indicador do tratamento (0 se placebo e 1 caso contrário).

Interpretação dos parâmetros

- β_{0t} representa o intercepto específico para cada tempo de medição, ou seja o valor esperado de Y_{it} quando as covariáveis x_{1i} e x_{2i} são simultaneamente iguais a zero.
- β_1 é o incremento positivo ou negativo no valor esperado de Y_{it} devido a variação em uma unidade da qualidade da visão inicial, fixada a covariável x_{2i}.
- β_{2t} é o incremento positivo ou negativo no valor esperado de Y_{it} específico para cada tempo de medição, fixada a covariável x_{1i}, ou seja é o efeito causado pelo tratamento proposto ao longo das semanas.
- ξ_{it} representa o erro aleatório, tal que $\xi_{it} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ $\forall$ i e t.
- **Observação**: Como atualmente ainda não há uma cura para DMRI, é pouco provável que um indivíduo que chegue com $x_{1i} = 0$ no início do estudo, ou seja com sua qualidade de visão nula venha a ter uma medição posterior positiva.
- Junto a isso desconsideraremos β_0, um intercepto geral no ajuste do modelo.

Dadas as estimativas dos efeitos do tratamento serem negativas isso favorece os efeitos do placebo, ou seja o placebo tem melhores resultados que o tratamento proposto. R^2: 0,9432 e R^2 ajustado : 0,9426

- Visto que há significância no teste-F isso indica que existe efeito no tratamento com a variação do tempo.



- Como o modelo proposto não leva em consideração as correlações entre as observações da qualidade visual de cada indivíduo, nem a heterocedasticidade que há entre as diferentes medidas ao longo do tempo, ele não deve ser considerado

como base para inferências.

4 INFERÊNCIA BAYESIANA

Segundo Manly (1997) a ideia básica por trás da Inferência Bayesiana é mudar as probabilidades para os parâmetros tomando valores númericos específicos para novas probabilidades como um resultado da coleta de mais dados, com essa mudança sendo alcançada através do Teorema de Bayes. Como um exemplo da abordagem Bayesiana, suponha que temos interesse no valor de um parâmetro θ de uma determinada população, e que antes de qualquer informação ser observada é de alguma forma possível afirmar que θ deve assumir um dos valores entre θ_1, θ_1, ..., θ_n e que a probabilidade de o valor ser θ_i é $\pi(\theta_i)$. Suponha também que alguns dados novos são coletados e a probabilidade de observar estes dados é $\pi(dados|\theta_i)$ se de fato $\theta = \theta_i$. Então o Teorema de Bayes afirma que a probabilidade de θ ser igual a θ_i, dado novas observações, é

$$\pi(\theta_i|dados) = \frac{\pi(dados|\theta_i)p(\theta_i)}{\sum_{j=1}^{n} \pi(data|\theta_j)\pi(\theta_j)}, \quad (??)$$

onde $\pi(\theta_i|dados)$ é a distribuição a posteriori de θ. Contudo frequentemente lidamos com situações em que vários parâmetros estão envolvidos, de tal forma que no geral

$$\pi(\theta_1, \theta_2, \ldots, \theta_p|dados) \propto \pi(dados|\theta_1, \theta_2, \ldots, \theta_p)\pi(\theta_1, \theta_2, \ldots, \theta_p), \quad (??)$$

ou seja, a distribuição a posteriori de vários parâmetros dado um conjunto de dados é proporcional a probabilidade dos dados quando conhecidos os parâmetros multiplicada pela probabilidade a priori dos parâmetros.

5 MONTE CARLO MARKOV CHAIN

5.1 INTRODUÇÃO ÀS CADEIAS DE MARKOV

Considere uma sequência de variáveis aleatórias discretas $\{X_0, X_1, X_2, \ldots\}$ e espaço de estados denotado por S=$\{s_1, s_2, \ldots, s_k\}$. A sequência de variáveis aleatórias $\{X_0, X_1, X_2, \ldots\}$ é uam Cadeia de Markov (CM), se

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

para t= $1, 2, \ldots$, ou seja, dada X_{t-1}, a distribuição de X_t independe de suas predecessoras, $X_{t-2}, X_{t-3}, \ldots$. Em um instante t qualquer, a probabilidade que o processo mude de um estado $X_t = s_i$ para um estado $X_{t+1} = s_j$ 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.

5.2 O AMOSTRADOR DE GIBBS

O amostrador de Gibbs é um método para aproximar uma distribuição multivariada tomando somente amostras de distribuições univariadas. O benefício deste método com Inferência Bayesiana é que torna relativamente fácil amostrar de uma distribuição a posteriori multivariada até mesmo quando o número de parâmetros envolvidos é muito grande.

Suponha que a distribuição a posteriori tenha função de densidade $\pi(\theta_1, \theta_2, \ldots, \theta_p)$ para os p parâmetros $\theta_1, \theta_2, \ldots, \theta_p$ e seja $\pi(\theta_i|\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_p)$ a função densidade condicional para θ_i dado os valores dos outros parâmetros. O problema então é gerar um grande número de amostras aleatórias da distribuição a posteriori com o objetivo de aproximar a própria distribuição e

distribuições de várias funções dos parâmetros. Isto é feito tomando arbitrariamente valores iniciais $\{\theta_1(0), \theta_2(0), \ldots, \theta_p(0)\}$ para os p parâmetros e em seguida mudá-los um a um selecionando novos valores como segue:

$$\begin{aligned}
\theta_1(1) \quad &\text{é escolhido de} \quad \pi(\theta_1|\theta_2(0), \theta_3(0), \ldots, \theta_p(0)) \\
\theta_2(1) \quad &\text{é escolhido de} \quad \pi(\theta_2|\theta_1(1), \theta_3(0), \ldots, \theta_p(0)) \\
\theta_3(1) \quad &\text{é escolhido de} \quad \pi(\theta_3|\theta_1(1), \theta_2(1), \theta_4(0) \ldots, \theta_p(0)) \\
&\vdots \\
\theta_p(1) \quad &\text{é escolhido de} \quad \pi(\theta_p|\theta_1(1), \theta_2(1), \ldots, \theta_{p-1}(1))
\end{aligned}$$

Nesse ponto todos os valores iniciais foram substituídos, o que representa um ciclo completo do algoritmo. O processo então é repetido muitas vezes produzindo a sequência $\{\theta_1(1), \theta_2(1), \ldots, \theta_p(1)\}, \{\theta_1(2), \theta_2(2), \ldots, \theta_p(2)\}, \ldots, \{\theta_1(N), \theta_2(N), \ldots, \theta_p(N)\}$ a qual é chamada de Cadeia de Markov, pois em cada etapa do algoritmo a mudança é feita dependendo apenas do valor atual de θ.

5.3 VANTAGENS E DESVANTAGENS DO MÉTODO AMOSTRADOR DE GIBBS

Dois fatores fazem este algoritmo útil. Primeiro que pode ser mostrado que $\{\theta_1(i), \theta_2(i), \ldots, \theta_p(i)\}$ segue a distribuição com densidade $\pi(\theta_1, \theta_2, \ldots, \theta_p)$ para valores grandes de i. Segundo que amostrar observações da distribuição condicional é frequentemente relativamente mais fácil, tornando o método de fácil implementação.

Complicações aparecem por que os conjuntos sucessivos de valores amostrais geradores podem ser correlacionados, porém pode ser resolvido tomando somente valores a partir da r-ésima etapa da sequência, com r grande o suficiente para garantir que os valores tenham correlações negligenciáveis. Paralelamente podem ser geradas várias sequências diferentes com valores iniciais escolhidos aleatoriamente e somente os conjuntos de valores finais $\{\theta_1(N), \theta_2(N), \ldots, \theta_p(N)\}$ serem mantidos e comparados.

6 MODELO MISTO

Vamos considerar agora o seguinte modelo:

$$Y_{it} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2it} + \beta_3 x_{3i} + \beta_4 x_{2it} x_{3i} + b_{0i} + \xi_{it}, \qquad (4)$$

- Y_{it} é a qualidade da visão do paciente i (i = 1, ..., 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.
- β_0 é o intercepto geral.
- β_1 é o incremento positivo ou negativo no valor esperado de Y_{it} quando variado em uma unidade o valor inicial da qualidade da visão.
- β_2 é o incremento positivo ou negativo na valor esperado de Y_{it}, quando acrescido o tempo em uma semana entre as que foram observadas.
- β_3 é o efeito geral positivo ou negativo no valor esperado de Y_{it} causado pelo tratamento.

- β_4 é o incremento positivo ou negativo 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.
- b_{0i} é o efeito aleatório específico para cada paciente. Tal que $b_{0i} \sim \mathcal{N}(0,\ \mathrm{d}_{11})\ \forall$ i.
- ξ_{it} é o erro aleatório. Tal que $\xi_{it} \overset{iid}{\sim} \mathcal{N}(0,\ \sigma^2)\ \forall$ i e t.
- b_{0i} representa uma variação especifica do β_0 para cada paciente.

Em notação matricial, o modelo para o sujeito i com o conjunto completo das quatro medidas da qualidade da visão é expresso por:

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

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\xi}_i$$

Com $\boldsymbol{\mathcal{D}} \equiv d_{11}$ e $\boldsymbol{\mathcal{R}}_{\mathrm{i}} \equiv \sigma^2\boldsymbol{I}_4$, no qual $\mathbf{I}_4$ é a matrix identidade 4 x 4.

Logo, a parte aleatória do modelo 4, considerando a matriz de variâncias e covariâncias marginal para o indivíduo i com as 4 observações é dada por:

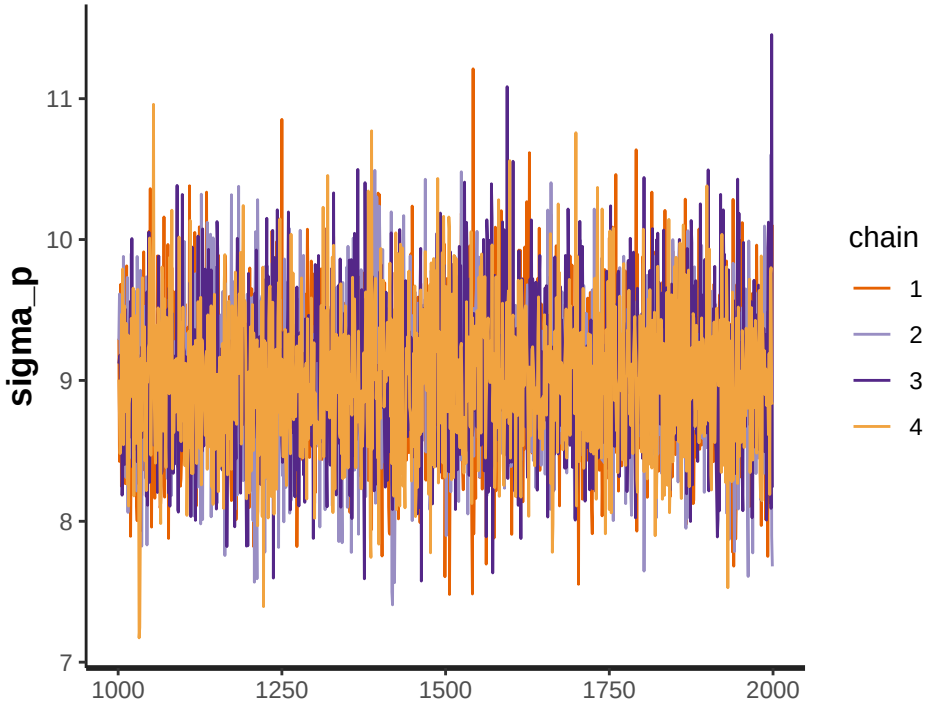
$$\boldsymbol{\mathcal{V}}_{\mathrm{i}} \equiv \boldsymbol{Z}_{\mathrm{i}}\boldsymbol{\mathcal{D}}\boldsymbol{Z}_{\mathrm{i}}^{'} + \boldsymbol{\sigma}^2\boldsymbol{I}_4$$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} d_{11} \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} \boldsymbol{\sigma^2} & 0 & 0 & 0 \\ 0 & \boldsymbol{\sigma^2} & 0 & 0 \\ 0 & 0 & \boldsymbol{\sigma^2} & 0 \\ 0 & 0 & 0 & \boldsymbol{\sigma^2} \end{pmatrix}$$

$$= \begin{pmatrix} \boldsymbol{\sigma^2} + d_{\mathbf{11}} & d_{\mathbf{11}} & d_{\mathbf{11}} & d_{\mathbf{11}} \\ d_{\mathbf{11}} & \boldsymbol{\sigma^2} + d_{\mathbf{11}} & d_{\mathbf{11}} & d_{\mathbf{11}} \\ d_{\mathbf{11}} & d_{\mathbf{11}} & \boldsymbol{\sigma^2} + d_{\mathbf{11}} & d_{\mathbf{11}} \\ d_{\mathbf{11}} & d_{\mathbf{11}} & d_{\mathbf{11}} & \boldsymbol{\sigma^2} + d_{\mathbf{11}} \end{pmatrix}$$

Pela estrutura da matriz de variâncias e covariâncias temos um coeficiente de correlação comum para todos os pares de variáveis $\varrho = d_{11}/(\sigma^2 + d_{11})$. Pelo fato do d_{11} ser não negativo, implica que ϱ é também não negativo.

	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
beta[1]	9.29	0.06	2.64	4.09	7.57	9.31	11.00	14.43	1661.61	1
beta[2]	0.83	0.00	0.04	0.74	0.80	0.83	0.85	0.92	1544.76	1
beta[3]	-0.21	0.00	0.02	-0.26	-0.23	-0.21	-0.20	-0.17	5356.73	1
beta[4]	-2.38	0.03	1.49	-5.22	-3.38	-2.42	-1.38	0.62	2028.87	1
beta[5]	-0.05	0.00	0.03	-0.12	-0.07	-0.05	-0.03	0.01	5522.73	1
sigma_e	8.65	0.00	0.25	8.19	8.48	8.64	8.81	9.14	4476.80	1
sigma_p	9.01	0.01	0.53	8.02	8.65	9.00	9.35	10.08	4779.63	1



Matrizes de Variâncias e Covariâncias Condicionais:

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

Matrizes de Variâncias e Covariâncias Marginais:

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

Matrizes de Correlações:

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

Referências

Manly, Bryan F. J. 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall.