

Universidade Federal do Paraná
Departamento de Estatística

Seminário

Fundamentos sobre modelos de transição e uma aplicação
para o caso de variáveis binárias

Prof.: IDEMAURO ANTONIO RODRIGUES DE LARA

CURITIBA-PR

AGOSTO/2008

OBJETIVO DE UM MODELO DE TRANSIÇÃO

No contexto de um estudo longitudinal, os modelos de transição permitem:

- **Descrever a distribuição das transições de respostas no tempo, bem como avaliar os efeitos ou fatores envolvidos nas transições (ou mudanças).**

EXEMPLO-APLICAÇÃO

- Os dados referem-se a um ensaio clínico, envolvendo doença respiratória, extraído do trabalho de Koch et al. (1990);
- envolveu 111 pacientes com problemas respiratórios de dois centros médicos (56 do Centro 1 e 55 do Centro 2);
- o objetivo principal é avaliar o efeito de uma droga ativa sob a condição respiratória;
- os indivíduos foram avaliados quanto ao estado de saúde (respiração) e qualificados em bom (1) ou ruim (0);
- o experimento envolveu cinco tempos de avaliação: início (*baseline*) e quatro tempos igualmente espaçados;
- as variáveis explicativas associadas a cada paciente foram: centro clínico; tratamento (A= ativo, P=placebo); sexo e idade (em anos).

QUESTÕES DE INTERESSE

1. Qual é a probabilidade de pacientes com condição respiratória ruim na primeira visita passarem para o estado de saúde bom na segunda visita?
2. Como essa probabilidade é influenciada pelo tratamento?
3. Essa probabilidade se altera com o tempo?
4. Em que grupo, ativo ou placebo, há uma maior probabilidade de passar para o estado de saúde bom?

Há variáveis que afetam as probabilidades condicionais:

$$\underbrace{P(Y_{it} = y_{it} \mid Y_{i(t-1)}, \mathbf{x})}_{\text{sexo, tratamento, idade, centro}}$$

⇒ **Modelos de transição: MLGs + processos estocásticos**

TÓPICOS DE PROCESSOS ESTOCÁSTICOS

Cadeias de Markov Discretas

$\tau = \text{tempo} \implies \tau = \{0, 1, \dots, T\}$ (discreto)

$S = \text{espaço de estados} \implies S = \{1, 2, \dots, k\}$ (categorias de resposta)

- Propriedade markoviana

$$P[X(t+1) = b \mid X(0) = x_0, X(1) = x_1, \dots, X(t-1) = x_{(t-1)}, X(t) = a] = \\ = P[X(t+1) = b \mid X(t) = a].$$

Essa propriedade define as **probabilidades de transição** $\pi_{ab}(t) \Rightarrow \mathbf{P}(t)_{k \times k}$.

- $\pi_{ab}(t, t+1) \geq 0$
- $\sum_{b=1}^k \pi_{ab}(t, t+1) = 1, \quad \forall t \in \tau \text{ e } a, b \in S.$

Matricialmente:

$$\mathbf{P}(t, t + 1) = \begin{pmatrix} \pi_{11}(t, t + 1) & \pi_{12}(t, t + 1) & \dots & \pi_{1k}(t, t + 1) \\ \pi_{21}(t, t + 1) & \pi_{22}(t, t + 1) & \dots & \pi_{2k}(t, t + 1) \\ \vdots & \vdots & \dots & \vdots \\ \pi_{k1}(t, t + 1) & \pi_{k2}(t, t + 1) & \dots & \pi_{kk}(t, t + 1) \end{pmatrix} .$$

Em particular para dados binários:

$$\underbrace{P(1) = \begin{pmatrix} \pi_{00}(1) & \pi_{01}(1) \\ \pi_{10}(1) & \pi_{11}(1) \end{pmatrix} \dots P(T) = \begin{pmatrix} \pi_{00}(T) & \pi_{01}(T) \\ \pi_{10}(T) & \pi_{11}(T) \end{pmatrix}}_{\text{processo não estacionário=T matrizes}}$$

ESTIMAÇÃO DAS PROBABILIDADES

- $n_a(0)$ unidades no estado a em $t = 0$, com probabilidades de ocorrência π_a , para todo $a \in S$.

Uma amostra aleatória, $N = \sum_{a=1}^k n_a(0)$

$\mathbf{n}(0) = (n_1(0), \dots, n_k(0))' \sim M(N, \pi_1, \dots, \pi_k)$, ou seja:

$$f(\mathbf{n}(0)) = \frac{N!}{\prod_{a=1}^k n_a(0)!} \prod_{a=1}^k \pi_a^{n_a(0)}.$$

- $n_{ab}(t)$ o número de unidades no estado a em $(t - 1)$ e em b no tempo t , sendo $n_{a.}(t) = \sum_{b=1}^k n_{ab}(t) = n_a(t - 1)$. Fixado um estado a :

$\mathbf{n}_a(t) = (n_{a1}(t), \dots, n_{ak}(t)) \sim M(n_a(t), \pi_{a1}(t - 1, t), \dots, \pi_{ak}(t - 1, t))$, ou seja:

$$f(\mathbf{n}_a(t)) = \frac{n_a(t - 1)!}{\prod_{b=1}^k n_{ab}(t)!} \prod_{b=1}^k \pi_{ab}(t - 1, t)^{n_{ab}(t)}.$$

Se todas as transições são independentes, então a função de verossimilhança de π_a e $\pi_{ab}(t)$ pode ser expressa por:

$$L \propto \prod_{a=1}^k \pi_a^{n_a(0)} \prod_{t=1}^T \prod_{a=1}^k \prod_{b=1}^k \pi_{ab}(t-1, t)^{n_{ab}(t)},$$

a qual leva às estimativas de máxima verossimilhança para os parâmetros:

$$\hat{\pi}_a = \frac{n_a(0)}{N} \quad e \quad \hat{\pi}_{ab}(t-1, t) = \frac{n_{ab}(t)}{n_a(t-1)}.$$

Quando o processo é estacionário, tem-se que o conjunto formado pelos elementos $n_{ab} = \sum_{t=1}^T n_{ab}(t)$ são estatísticas suficientes para as probabilidades de transição e, nesse caso, são estimadas por:

$$\hat{\pi}_{ab} = \frac{n_{ab}}{n_a} = \frac{\sum_{t=1}^T n_{ab}(t)}{\sum_{t=1}^T n_a(t-1)}.$$

TESTE PARA A HIPÓTESE DE ESTACIONARIEDADE

- A hipótese principal é $H_0 : \pi_{ab}(t) = \pi_{ab}$, para todo $t \in \tau$ e $a, b \in S$.
Matricialmente $H_0 : \mathbf{P}(t) = \mathbf{P}$

Tabela 1: Tabela de contingência $T \times k$ para a apresentação das probabilidades de transição, para um estado fixo a

	b			
t	1	2	...	k
1	$\hat{\pi}_{a1}(1)$	$\hat{\pi}_{a2}(1)$...	$\hat{\pi}_{ak}(1)$
2	$\hat{\pi}_{a1}(2)$	$\hat{\pi}_{a2}(2)$...	$\hat{\pi}_{ak}(2)$
⋮	⋮	⋮	⋮	⋮
T	$\hat{\pi}_{a1}(T)$	$\hat{\pi}_{a2}(T)$...	$\hat{\pi}_{ak}(T)$

Essa hipótese pode ser verificada através das estatísticas:

$$\xi = \sum_{t=1}^T \sum_{a=1}^k \sum_{b=1}^k \frac{n_a(t-1) [\hat{\pi}_{ab}(t) - \hat{\pi}_{ab}]^2}{\hat{\pi}_{ab}}, \quad (1)$$

ou, pelo teste da razão de verossimilhanças:

$$\lambda = -2 \sum_{t=1}^T \sum_{a=1}^k \sum_{b=1}^k n_{ab}(t) \log \left(\frac{\hat{\pi}_{ab}}{\hat{\pi}_{ab}(t)} \right), \quad (2)$$

sendo o teste (2) equivalente ao teste dado por (1), ambos assintoticamente com distribuição $\chi_{k(k-1)(T-1)}^2$.

CADEIAS DE MARKOV DE ALCANCE MAIOR DO QUE 1

A propriedade markoviana, para uma cadeia de q -ésima ordem é definida pela expressão:

$$P[X(t) = a_{q+1} \mid X(t-1) = a_q, \dots, X(t-q) = a_1, \dots, X(1) = x_1, X(0) = x_0] = P[X(t) = a_{q+1} \mid X(t-1) = a_q, \dots, X(t-q) = a_1].$$

Por exemplo, uma cadeia de Markov de ordem 2, com dados binários, é caracterizada por matrizes de dimensões 4×2 :

$$\mathbf{P}(t) = \begin{pmatrix} \pi_{000}(t) & \pi_{001}(t) \\ \pi_{010}(t) & \pi_{011}(t) \\ \pi_{100}(t) & \pi_{101}(t) \\ \pi_{110}(t) & \pi_{111}(t) \end{pmatrix},$$

para $t = 2, 3, 4, \dots, T$.

MODELOS DE TRANSIÇÃO DE MARKOV

Quanto à estimação, destacam-se dois procedimentos:

1. O primeiro procedimento, baseado na análise de tabelas de contingência, é aplicado aos casos em que a amostra é homogênea;
2. O segundo procedimento é obtido a partir da definição de MLGs, incorporando a estrutura estocástica de uma cadeia de Markov.

TABELAS DE CONTINGÊNCIA E PROBABILIDADES DE TRANSIÇÃO

Anderson e Goodman (1957), Goodman (1962), Bishop, Fienberg e Holland (1975), Agresti (1990), Fitzmaurice e Laird (1993), Lindsey (1995, 2004).

Tabela 2: Estudo longitudinal com resposta binária, amostra homogênea

Indivíduo	Tempo			
	0	1	...	T
1	0	1	...	1
2	1	1	...	0
3	0	0	...	0
4	1	0	...	1
⋮	⋮	⋮	⋮	⋮
N	0	0	...	1

Tabela 3: Tabela de contingência 2×2 , na t -ésima transição, de um estudo longitudinal com resposta binária e amostra homogênea

$Y_{(t-1)}$	Y_t		Total marginal (t-1)
	0	1	
0	$n_{00}(t)$	$n_{01}(t)$	$n_{0.}(t-1)$
1	$n_{10}(t)$	$n_{11}(t)$	$n_{1.}(t-1)$
Total marginal (t)	$n_{.0}(t)$	$n_{.1}(t)$	N

Estimativas das probabilidades: $\hat{\pi}_{ab}(t) = \frac{n_{ab}(t)}{n_{a.}(t-1)}$

$$\hat{P}(t) = \underbrace{\begin{pmatrix} \hat{\pi}_{00}(t) & \hat{\pi}_{01}(t) \\ \hat{\pi}_{10}(t) & \hat{\pi}_{11}(t) \end{pmatrix}}_{\text{processo não estacionário}}$$

Tabela 4: Tabela de contingência 2×2 , do total de transições de um estudo longitudinal com resposta binária e amostra homogênea: processo estacionário

		Y_t		
$Y_{(t-1)}$	0	1	Total marginal	
0	$\sum_{t=1}^T n_{00}(t)$	$\sum_{t=1}^T n_{01}(t)$	$\sum_{t=1}^T n_{0.}(t-1)$	
1	$\sum_{t=1}^T n_{10}(t)$	$\sum_{t=1}^T n_{11}(t)$	$\sum_{t=1}^T n_{1.}(t-1)$	
Total marginal	$\sum_{t=1}^T n_{.0}(t)$	$\sum_{t=1}^T n_{.1}(t)$	TN	

Estimativas das probabilidades: $\hat{\pi}_{ab} = \frac{\sum_{t=1}^T n_{ab}(t)}{\sum_{t=1}^T n_{a.}(t-1)}$

$$\hat{P} = \underbrace{\begin{pmatrix} \hat{\pi}_{00} & \hat{\pi}_{01} \\ \hat{\pi}_{10} & \hat{\pi}_{11} \end{pmatrix}}_{\text{processo estacionário}}$$

MODELOS LINEARES GENERALIZADOS DE TRANSIÇÃO

Cox (1970); Korn e Whittemore (1979); Azzalini (1983); Zeger, Liang e Self (1985); Ware, Lipsitz e Speizer (1988); Zeger e Qaqish (1988); Zeger e Liang (1992); Diggle et al. (2002), Molenberghs e Verbeke (2005).

EM GERAL, A AMOSTRA DIFICILMENTE É HOMOGÊNEA.

Alternativa:

- especificar modelos para descrever a relação funcional de dependência das probabilidades de transição em relação às covariáveis;
- as probabilidades são estimadas a partir das estimativas dos parâmetros do modelo \Rightarrow que ainda descrevem o peso que cada variável explicativa exerce sob a probabilidade de transição.

DEFINIÇÃO DE UM MLG DE TRANSIÇÃO

Um modelo de transição de Markov especifica um MLG para a distribuição condicional de Y_{it} dado $\mathbf{h}_{it} = (y_{i1}, y_{i2}, \dots, y_{i,(t-1)})$, o vetor de dimensão q das respostas prévias. Então, a distribuição condicional de $Y_{it} | \mathbf{h}_{it}$ é dada por:

$$f(y_{it} | \mathbf{h}_{it}) = \exp \left\{ \frac{w_i}{\phi} \left[y_{it} \theta_{it} - b(\theta_{it}) \right] + c(y_{it}, \phi) \right\},$$

em que

$$\mu_{it}^C = E(Y_{it} | \mathbf{h}_{it}) = b'(\theta_{it}) \quad \text{e} \quad v_{it}^C = \text{Var}(Y_{it} | \mathbf{h}_{it}) = b''(\theta_{it})\phi,$$

denotam a média e variância condicionais a \mathbf{h}_{it} .

A média e a variância condicional satisfazem as equações:

$$g(\mu_{it}^C) = \eta_{it} = \mathbf{x}'_{it} \boldsymbol{\beta} + \sum_{r=1}^s f_r(\mathbf{h}_{it}; \boldsymbol{\alpha}) \quad \text{e} \quad v_{it}^C = v(\mu_{it}^C)\phi,$$

em que $g(\mu_{it}^C)$ e $v(\mu_{it}^C)$ denotam a função de ligação e a função de variância, respectivamente.

PROCESSOS NÃO ESTACIONÁRIOS

Proposta de Ware, Lipsitz e Speizer (1988) \Rightarrow **Ajustar um modelo para cada ocasião**

\Rightarrow **O processo de ajuste é análogo ao de um MLG para dados independentes.**

Para dados categorizados ordinais, as probabilidades de transição são dadas por:

$$\pi_{ab}(t-1, t) = P(I_{ibt} \mid I_{ia(t-1)}, \mathbf{x}_{it}), \quad a, b = 1, 2, \dots, k; \quad t = 0, 1, \dots, T,$$

e podem ser estimadas a partir do modelo de chances proporcionais (McCULLAGH, 1980).

As probabilidades de transição acumuladas são especificadas por:

$$\gamma_{abt}(\mathbf{x}) = \frac{\exp(\lambda_{abt} - \beta'_t \mathbf{x})}{1 + \exp(\lambda_{abt} - \beta'_t \mathbf{x})} \quad b = 1, 2, \dots, k - 1.$$

\Rightarrow **Tem como caso particular a regressão logística para dados binários.**

ESTRUTURA DE DADOS

PROCESSOS NÃO ESTACIONÁRIOS

centro	id	treat	sex	idade	base	resp1	resp2	resp3	resp4
1	1	P	M	46	0	0	0	0	0
1	2	P	M	28	0	0	0	0	0
1	3	A	M	23	1	1	1	1	1
1	4	P	M	44	1	1	1	1	0
.....									
2	108	A	F	39	0	1	1	1	1
2	109	A	M	68	0	1	1	1	1
2	110	A	F	63	1	1	1	1	1
2	111	A	M	31	1	1	1	1	1

ESTIMAÇÃO POR MÁXIMA VEROSSIMILHANÇA

PROCESSOS ESTACIONÁRIOS

Assumindo um modelo de transição de Markov de ordem q , a distribuição condicional de Y_{it} é expressa por:

$$f(y_{it} | \mathbf{h}_{it}) = f(y_{it} | y_{i(t-1)}, y_{i(t-2)}, \dots, y_{i(t-q)}),$$

tal que a contribuição na verossimilhança do i -ésimo elemento é dada por:

$$\underbrace{f(y_{i1}, y_{i2}, \dots, y_{iq})}_{\text{não especificada}} \prod_{t=q+1}^{n_i} \underbrace{f(y_{it} | y_{i(t-1)}, y_{i(t-2)}, \dots, y_{i(t-q)})}_{\text{especificada pelo MLG}}.$$

Como alternativa para estimação dos parâmetros, maximiza-se a verossimilhança condicional:

$$\prod_{i=1}^N f(y_{i(q+1)}, \dots, y_{it_i} | y_{i1}, \dots, y_{iq}) = \prod_{i=1}^N \prod_{t=q+1}^{n_i} f(y_{it} | \mathbf{h}_{it}).$$

O logaritmo da função de verossimilhança condicional é:

$$\ell(\boldsymbol{\theta}_{it}) = \sum_{i=1}^N \sum_{t=q+1}^{n_i} \frac{1}{\phi} \left[y_{it} \boldsymbol{\theta}_{it} - b(\boldsymbol{\theta}_{it}) \right] + c(y_{it}, \phi).$$

Observa-se que

$$\ell = f(\boldsymbol{\theta}_{it}); \quad \boldsymbol{\theta}_{it} = q(\boldsymbol{\mu}_{it}^C); \quad \boldsymbol{\mu}_{it}^C = g^{-1}(\boldsymbol{\eta}_{it}) \quad \text{e} \quad \boldsymbol{\eta}_{it} = \mathbf{x}'_{it} \boldsymbol{\beta} + \sum_{r=1}^q \alpha_r y_{it-r}.$$

Portanto, usando a regra da cadeia obtém-se:

$$U^C(\boldsymbol{\delta}) = \sum_{i=1}^N \sum_{t=q+1}^{n_i} \frac{\partial \boldsymbol{\mu}_{it}^C}{\partial \boldsymbol{\delta}} (v_{it}^C)^{-1} (y_{it} - \boldsymbol{\mu}_{it}^C) = 0.$$

Pelo método score de Fisher, pode-se descrever o processo iterativo matricialmente:

$$\hat{\boldsymbol{\delta}}^{(m+1)} = \left(\sum_{i=1}^N \mathbf{X}_i^{*'} \mathbf{W}_i^{(m)} \mathbf{X}_i^* \right)^{-1} \left(\sum_{i=1}^N \mathbf{X}_i^{*'} \mathbf{W}_i^{(m)} \mathbf{Z}_i^{(m)} \right).$$

(Mínimos quadrados ponderados iterativamente)

ESTRUTURA DE DADOS

PROCESSOS ESTACIONÁRIOS

Obs	centro	id	trat	sexo	idade	visit	resp	prev1	prev2
1	1	1	P	M	46	1	0	.	.
2	1	1	P	M	46	2	0	0	.
3	1	1	P	M	46	3	0	0	0
4	1	1	P	M	46	4	0	0	0
5	1	1	P	M	46	5	0	0	0
.....									
551	2	111	A	M	31	1	1	.	.
552	2	111	A	M	31	2	1	1	.
553	2	111	A	M	31	3	1	1	1
554	2	111	A	M	31	4	1	1	1
555	2	111	A	M	31	5	1	1	1

MÉTODOS

DEFINIÇÃO DO MODELO PARA DADOS BINÁRIOS

A distribuição condicional de Y_{it} , dado \mathbf{h}_i segue um MLG:

$$f(y_{it} | \mathbf{h}_i) = \exp \left\{ y_{it} \ln \left(\frac{\mu_{it}^C}{1 - \mu_{it}^C} \right) + \ln(1 - \mu_{it}^C) \right\}, \quad (i = 1, 2, \dots, N)$$

para o qual especificam-se as funções de ligação e variância:

$$g(\mu_{it}^c) = \eta_{it} \quad \text{e} \quad v(\mu_{it}^c) = \mu_{it}^c(1 - \mu_{it}^c).$$

Para o caso em que a ordem da cadeia é 1, assumindo função de ligação canônica, tem-se como preditor linear:

$$\eta_{it} = \text{logito} [P(Y_{it} = b | Y_{i(t-1)} = a)] = \ln \left(\frac{\pi_{ab}(t)}{1 - \pi_{ab}(t)} \right) = \mathbf{x}'_{it} \boldsymbol{\beta} + \alpha y_{i(t-1)}.$$

MATRIZES DE PROBABILIDADES DE TRANSIÇÃO

A caracterização das matrizes de probabilidades de transição, 2×2 , com elementos $\pi_{ab}(t) = P(Y_{it} = b \mid Y_{i(t-1)} = a)$, em que $a, b \in S = \{0, 1\}$ é feita a partir de uma função de ligação conveniente.

Para a função de ligação logito, tem-se:

$$\pi_{ab}(t) = \frac{\exp(\eta_{it})}{1 + \exp(\eta_{it})}.$$

Logo:

$$P(t) = \begin{pmatrix} \pi_{00}(t) = \frac{1}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta})} & \pi_{01}(t) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta})} \\ \pi_{10}(t) = \frac{1}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \alpha)} & \pi_{11}(t) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \alpha)}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \alpha)} \end{pmatrix}.$$

Uma forma mais concisa para estimação é adotar a estrutura linear:

$$\eta_{it} = \text{logito} [P(Y_{it} \mid Y_{i(t-1)})] = \ln \left(\frac{\pi_{ab}(t)}{1 - \pi_{ab}(t)} \right) = \mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)}\mathbf{x}'_{it}\alpha.$$

Para a função probito, tem-se:

$$\eta_{it} = \text{probito}(\pi_{ab}(t)) = \Phi^{-1}(\pi_{ab}(t)) = \mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)}\mathbf{x}'_{it}\boldsymbol{\alpha},$$

com probabilidades especificadas pela matriz:

$$\mathbf{P}(t) = \begin{pmatrix} \pi_{00}(t) = 1 - \Phi(\mathbf{x}'_{it}\boldsymbol{\beta}) & \pi_{01}(t) = \Phi(\mathbf{x}'_{it}\boldsymbol{\beta}) \\ \pi_{10}(t) = 1 - \Phi(\mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)}\mathbf{x}'_{it}\boldsymbol{\alpha}) & \pi_{11}(t) = \Phi(\mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)}\mathbf{x}'_{it}\boldsymbol{\alpha}) \end{pmatrix}.$$

E, para um modelo com função de ligação complemento *log-log*, especifica-se:

$$\eta_{it} = \ln(-\ln(1 - \pi_{ab}(t))) = \mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)}\mathbf{x}'_{it}\boldsymbol{\alpha},$$

levando a:

$$\mathbf{P}(t) = \begin{pmatrix} \pi_{00}(t) = 1 - \pi_{01}(t) & \pi_{01}(t) = 1 - \exp(-\exp(\mathbf{x}'_{it}\boldsymbol{\beta})) \\ \pi_{10}(t) = 1 - \pi_{11}(t) & \pi_{11}(t) = 1 - \exp(-\exp(\mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)}\mathbf{x}'_{it}\boldsymbol{\alpha})) \end{pmatrix}.$$

Uma extensão para um modelo de ordem q tem como preditor linear:

$$\eta_{it} = g[P(Y_{it} = b \mid \mathbf{h}_{it})] = \mathbf{x}'_{it}\boldsymbol{\beta} + \sum_{r=1}^q \alpha_r y_{i(t-r)} \mathbf{x}_{it},$$

Em particular, uma cadeia de Markov de alcance 2, com função de ligação canônica tem-se:

$$\begin{aligned} \eta_{it} &= \text{logito} [P(Y_{it} = c \mid Y_{i(t-1)} = b, Y_{i(t-2)} = a)] \\ &= \mathbf{x}'_{it}\boldsymbol{\beta} + y_{i(t-1)} \mathbf{x}_{it}\boldsymbol{\alpha}_1 + y_{i(t-2)} \mathbf{x}_{it}\boldsymbol{\alpha}_2 + y_{i(t-1)} y_{i(t-2)} \mathbf{x}_{it}\boldsymbol{\alpha}_3. \end{aligned}$$

$$P(t) = \begin{pmatrix} \pi_{000}(t) = 1 - \pi_{001}(t) & \pi_{001}(t) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta})}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta})} \\ \pi_{010}(t) = 1 - \pi_{011}(t) & \pi_{011}(t) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}_{it}\boldsymbol{\alpha}_1)}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}_{it}\boldsymbol{\alpha}_1)} \\ \pi_{100}(t) = 1 - \pi_{101}(t) & \pi_{101}(t) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}_{it}\boldsymbol{\alpha}_2)}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}_{it}\boldsymbol{\alpha}_2)} \\ \pi_{110}(t) = 1 - \pi_{111}(t) & \pi_{111}(t) = \frac{\exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}_{it}\boldsymbol{\alpha}_1 + \mathbf{x}_{it}\boldsymbol{\alpha}_2 + \mathbf{x}_{it}\boldsymbol{\alpha}_3)}{1 + \exp(\mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}_{it}\boldsymbol{\alpha}_1 + \mathbf{x}_{it}\boldsymbol{\alpha}_2 + \mathbf{x}_{it}\boldsymbol{\alpha}_3)} \end{pmatrix}.$$

TESTE PARA ALCANCE DE UMA CADEIA

A hipótese nula de interesse é:

H_0 : A cadeia é de $(q - 1)$ -ésima ordem;

contra a alternativa de que a cadeia é de q -ésima ordem.

Testes usuais (Anderson e Goodman, 1957)

$$\xi = \frac{\sum_{a,b,\dots,\ell,m} n_{ab,\dots,\ell}^* (\hat{\pi}_{ab,\dots,\ell,m} - \hat{\pi}_{b,\dots,\ell,m})^2}{\hat{\pi}_{b,\dots,\ell,m}} \sim \chi_{(k-1)k(q-1)}^2,$$

ou

$$\lambda = -2 \log \Lambda = 2 \sum_{ab,\dots,\ell,m} n_{ab,\dots,\ell,m} [\log \hat{\pi}_{ab,\dots,\ell,m} - \log \hat{\pi}_{b,\dots,\ell,m}] \sim \chi_{(k-1)k(q-1)}^2.$$

Com estrutura de um MLG (Proposta)

Denotem $\ell(\hat{\beta}, \hat{\alpha}, \mathbf{y})_{(q-1)}$ e $\ell(\hat{\beta}, \hat{\alpha}, \mathbf{y})_{(q)}$ os logaritmos das funções de verossimilhança maximizadas sob as condições de ordem $(q - 1)$ e q , respectivamente. Tem-se, assintoticamente que:

$$\lambda = 2(\ell(\hat{\beta}, \hat{\alpha}, \mathbf{y})_{(q)} - \ell(\hat{\beta}, \hat{\alpha}, \mathbf{y})_{(q-1)}) \sim \chi_v^2.$$

TESTES PARA COMPARAÇÃO DE TRATAMENTOS

1-TESTE DA RAZÃO DE CHANCES

Tabela 5: Distribuição das probabilidades de transição, considerando os grupos placebo e ativo (tratamento), com estado prévio fixado $Y_{(t-1)} = a$, num modelo de Markov

Tratamentos	Y_{it}	
	0	1
Placebo (P)	$\pi_{a0}(P)$	$\pi_{a1}(P)$
Ativo (A)	$\pi_{a0}(A)$	$\pi_{a1}(A)$

Razão de chances:

$$\psi_{Y_{(t-1)}=a} = \frac{\pi_{a0}(P) \times \pi_{a1}(A)}{\pi_{a0}(A) \times \pi_{a1}(P)},$$

Definição das hipóteses de interesse:

$$\text{Hipóteses : em termos de razão de chances} \left\{ \begin{array}{l} H_0 : \psi = 1; \\ H_a : \psi > 1. \end{array} \right.$$

HOMOGENEIDADE DA RAZÃO DE CHANCES EM RELAÇÃO À h_{it}

Para cadeia de ordem 1:

$$H_0 : \psi_{Y_{(t-1)=0}} = \psi_{Y_{(t-1)=1}}.$$

Essa hipótese pode ser verificada através do modelo estocástico:

$$\eta = \ln \left(\frac{\pi_{ab}}{1 - \pi_{ab}} \right) = \beta_0 + \beta_1 \text{tratamento} + \alpha y_{(t-1)} + \gamma(y_{(t-1)}) * \text{tratamento}$$

$$H_0 : \gamma = 0 \begin{cases} M(1) : \eta_{(1)} = \beta_0 + \beta_1 \text{tratamento} + \alpha y_{(t-1)} \\ M(2) : \eta_{(2)} = \beta_0 + \beta_1 \text{tratamento} + \alpha y_{(t-1)} + \gamma(y_{(t-1)}) * \text{tratamento} \\ \text{Desvio } (M(1)) - \text{Desvio } (M(2)) \sim \chi_{(1)}^2. \end{cases}$$

HOMOGENEIDADE DA RAZÃO DE CHANCES EM RELAÇÃO À h_{it}

Para cadeia de ordem 2:

$$H_0 : \psi_{\mathbf{h}_t=(0,0)} = \psi_{\mathbf{h}_t=(0,1)} = \psi_{\mathbf{h}_t=(1,0)} = \psi_{\mathbf{h}_t=(1,1)}.$$

Essa hipótese pode ser verificada através do modelo estocástico:

$$\eta = \beta_0 + \beta_1 x + \alpha_1 y_{(t-1)} + \alpha_2 y_{(t-2)} + \gamma_1 y_{(t-1)} x + \gamma_2 y_{(t-2)} x + \gamma_3 y_{(t-1)} y_{(t-2)} x,$$

em que $x = \text{tratamento}$.

$$H_0 : \gamma = \mathbf{0} \left\{ \begin{array}{l} \psi_{\mathbf{h}_t=(0,0)} = \exp(\beta_1) \\ \psi_{\mathbf{h}_t=(0,1)} = \exp(\beta_1 + \gamma_1) \\ \psi_{\mathbf{h}_t=(1,0)} = \exp(\beta_1 + \gamma_2) \\ \psi_{\mathbf{h}_t=(1,1)} = \exp(\beta_1 + \gamma_1 + \gamma_2 + \gamma_3) \end{array} \right.$$

EXTENSÃO DO TESTE DE MANTEL-HAENSZEL

Tabela 6: Estrutura usual de uma tabela de contingência 2×2 para o teste de Mantel-Haenszel

Estratos	Resposta		Total marginal
	1	2	
A	y_1	$n_1 - y_1$	n_1
B	y_2	$n_2 - y_2$	n_2
Total marginal	m	$n - m$	n

Sob essa estrutura, a distribuição condicional $(y_1 \mid m, \psi)$ produz um modelo pertencente à família de distribuições hipergeométricas não centrais:

$$f(y_1 \mid m, \psi) = \frac{\binom{n_1}{y_1} \binom{n_2}{m - y_1} \psi^{y_1}}{\sum_s \binom{n_1}{s} \binom{n_2}{m - s} \psi^s},$$

em que s varia entre o máximo $(0, m - n_2)$ e o mínimo (n_1, m) .

Em particular, para $\psi = 1$:

$$f(y_1 | m, \psi = 1) = \frac{\binom{n_1}{y_1} \binom{n_2}{y_2}}{\binom{n_1 + n_2}{m}},$$

Assintoticamente, essa distribuição tende a uma distribuição normal com média e variância dadas por:

$$E(Y_1 | m, \psi = 1) = \frac{mn_1}{n} \quad \text{e} \quad \text{Var}(Y_1 | m, \psi = 1) = \frac{n_1 n_2 m (n - m)}{n^3}.$$

Assim,

$$Z = \frac{y_1 - \frac{mn_1}{n}}{\sqrt{\frac{n_1 n_2 m (n - m)}{n^3}}} \sim N(0, 1),$$

é uma estatística apropriada para avaliar a hipótese de que $\psi = 1$.

Tabela 7: Tabela de contingência 2×2 , considerando os estratos de tratamento e as frequências de transições para os estados 0 e 1, dado $Y_{(t-1)} = a, a \in \{0, 1\}$

Tratamento	Y_t		Total marginal
	0	1	
Placebo	$n_{a0(j)}^*(P)$	$n_{a1(j)}^*(P)$	$n_{a.(j)}^*(P)$
Ativo	$n_{a0(j)}^*(A)$	$n_{a1(j)}^*(A)$	$n_{a.(j)}^*(A)$
Total marginal	$n_{.0(j)}^*$	$n_{.1(j)}^*$	$n^{*..(j)}$

Estatística do teste:

$$Z_{MH} = \frac{\sum_{j=1}^2 n_{a0(j)}^*(P) - \sum_{j=1}^2 (E)_j}{\sqrt{\sum_{j=1}^2 (\text{Var})_j}} \sim N(0, 1),$$

$$\text{em que } (E)_j = \frac{n_{.0(j)}^* n_{a.(j)}^*(P)}{n^{*..(j)}} \text{ e } (\text{Var})_j = \frac{n_{a.(j)}^*(P) n_{a.(j)}^*(A) n_{.0(j)}^* (n^{*..(j)} - n_{.0(j)}^*)}{(n^{*..(j)})^3}.$$

Assim para $H_a : \psi > 1$, o nível descritivo do teste é dado por $p^* = P(Z > Z_{MH})$.

2-TESTE DA DIFERENÇA DAS PROBABILIDADES DE TRANSIÇÃO

Definição das hipóteses:

$$\text{Hipóteses : mudanças para o estado 1} \begin{cases} H_0 : \pi_{01}(A) - \pi_{01}(P) = 0; \\ H_a : \pi_{01}(A) - \pi_{01}(P) > 0. \end{cases}$$

Pela equação de Chapman-Kolmogorov:

$$\pi_{01}^{(r)} = \sum_{k \in S} \pi_{0k}^{(s)} \pi_{k1}^{(v)},$$

podem-se formular as hipóteses:

$$\text{Hipóteses : mudanças para o estado 1 em r passos} \begin{cases} H_0 : \pi_{01}^{(r)}(A) - \pi_{01}^{(r)}(P) = 0; \\ H_a : \pi_{01}^{(r)}(A) - \pi_{01}^{(r)}(P) > 0. \end{cases}$$

Resultado da Teoria Assintótica

$$\sqrt{n\varsigma_a}(\hat{\pi}_{ab} - \pi_{ab}) \rightarrow N(0, \pi_{ab}(1 - \pi_{ab})),$$

$$\text{em que } \varsigma_a = \sum_{j=1}^k \sum_{t=1}^T \frac{n_j(0)}{n} \pi_{ja}^{[t-1]}.$$

Para os grupos ativo e placebo tem-se assintoticamente:

$$\hat{\pi}_{01}^{(r)}(A) \sim N \left[\pi_{01}^{(r)}(A), \frac{\pi_{01}^{(r)}(A)[1 - \pi_{01}^{(r)}(A)]}{n(A)\zeta_a(A)} \right]$$

e

$$\hat{\pi}_{01}^{(r)}(P) \sim N \left[\pi_{01}^{(r)}(P), \frac{\pi_{01}^{(r)}(P)[1 - \pi_{01}^{(r)}(P)]}{n(P)\zeta_a(P)} \right].$$

Como há independência entre esses grupos, a estatística $\hat{\pi}_{01}^{(r)}(A) - \hat{\pi}_{01}^{(r)}(P)$, também tem como limite uma distribuição normal:

$$\hat{\pi}_{01}^{(r)}(A) - \hat{\pi}_{01}^{(r)}(P) \sim N(\pi_D, \sigma_D^2),$$

$$\text{em que } \pi_D = \pi_{01}^{(r)}(A) - \pi_{01}^{(r)}(P) \text{ e } \sigma_D^2 = \frac{\pi_{01}^{(r)}(A)[1 - \pi_{01}^{(r)}(A)]}{n(A)\zeta_a(A)} + \frac{\pi_{01}^{(r)}(P)[1 - \pi_{01}^{(r)}(P)]}{n(P)\zeta_a(P)}.$$

Sob H_0 e considerando teste unilateral à direita, o nível descritivo é $p^* = P \left(Z > \frac{\hat{\pi}_D}{\hat{\sigma}_D} \right)$.

RESULTADOS

AJUSTES DE MODELOS EXPLORATÓRIOS

$$\eta_{it} = \beta_0 + \beta_1 \text{tratamento} + \alpha y_{i(t-1)} + \gamma (y_{i(t-1)} * \text{tratamento}).$$

Resultados dos testes de estacionariedade (ou homogeneidade): $\xi = 9,62$ (ativo) e $\xi = 5,93$ (placebo). Resultados n.s., uma vez que $\chi^2_{6; 0,01} = 16,81$.

Tabela 8: Estimativas dos parâmetros fixos e respectivos erros-padrão, para o modelo de transição estacionário de primeira ordem, no estudo sobre a condição respiratória.

Parâmetros	Estimativa	Erro-padrão
intercepto	-0,3248	0,4104
centro 1	-0,7622	0,2477
sexo (feminino)	0,0506	0,3111
tratamento (ativo)	0,8382	0,2413
idade	-0,0177	0,0093
$Y_{i(t-1)}$	2,2277	0,2337

Os efeitos de centro clínico, tratamento, e resposta prévia são significativos ($p < 0,01$). Os efeitos das interações da resposta prévia com as covariáveis foram não significativos. Desse modo, tem-se as matrizes:

$$\hat{\mathbf{P}}_{(ativo)} = \begin{pmatrix} 0,5618 & 0,4382 \\ 0,1214 & 0,8786 \end{pmatrix} \quad \text{e} \quad \hat{\mathbf{P}}_{(placebo)} = \begin{pmatrix} 0,7478 & 0,2522 \\ 0,2422 & 0,7578 \end{pmatrix}.$$

Por Chapman-Kolmogorov, tem-se:

$$\hat{\mathbf{P}}_{(ativo)}^{(4)} = \begin{pmatrix} 0,2464 & 0,7536 \\ 0,2088 & 0,7911 \end{pmatrix} \quad \text{e} \quad \hat{\mathbf{P}}_{(placebo)}^{(4)} = \begin{pmatrix} 0,5232 & 0,4768 \\ 0,4578 & 0,5422 \end{pmatrix}.$$

sendo $\pi_{01}^{(4)}(A) - \pi_{01}^{(4)}(P) = 0,2768$, com erro-padrão estimado em 0,1527, levando à estatística $Z = 1,81$, (nível descritivo = 0,036).

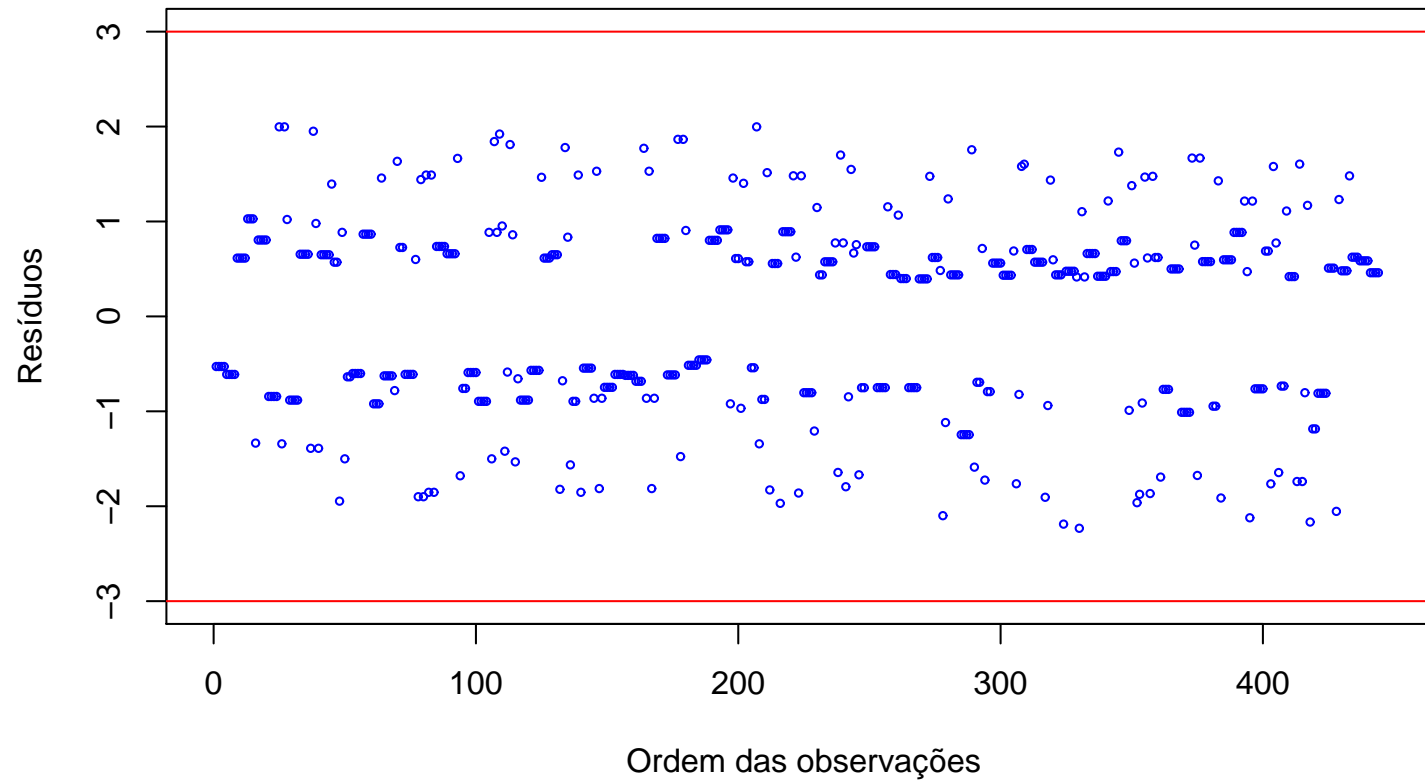


Figura 1: Gráfico dos resíduos de Pearson versus a ordem das observações para o modelo de transição estacionário de 1^a ordem, no estudo da condição respiratória

Modelo de transição de segunda ordem

Tabela 9: Estimativas dos parâmetros e respectivos erros-padrão para o modelo de transição estacionário de segunda ordem, no estudo sobre a condição respiratória

Parâmetros	Estimativa	Erro-padrão
intercepto	-0,5616	0,4940
centro 1	-0,6540	0,2967
sexo (feminino)	0,1945	0,3709
tratamento (ativo)	0,7339	0,2889
idade	-0,0252	0,0111
$Y_{i(t-1)}$	1,9213	0,2979
$Y_{i(t-2)}$	0,9477	0,2974

Teste para alcance da cadeia com os demais efeitos \Rightarrow **razão de verossimilhanças**

$$\lambda = 2(\ell(\hat{\beta}, \hat{\alpha}, \mathbf{y})_{(2)} - \ell(\hat{\beta}, \hat{\alpha}, \mathbf{y})_{(1)}) = 2(-161,85 - (-166,79)) = 9,88$$

(nível descritivo = 0,0016)

Modelo de transição de segunda ordem

Considerando, então, os efeitos significativos desse modelo, as matrizes de probabilidades de transição, para os respectivos grupos ativo e placebo, são:

$$\hat{P}_{(ativo)} = \begin{pmatrix} 0,4570 & 0,5430 \\ 0,2460 & 0,7540 \\ 0,1097 & 0,8903 \\ 0,0455 & 0,9545 \end{pmatrix} \quad \text{e} \quad \hat{P}_{(placebo)} = \begin{pmatrix} 0,6368 & 0,3632 \\ 0,4046 & 0,5954 \\ 0,2042 & 0,7958 \\ 0,0905 & 0,9095 \end{pmatrix}.$$

As razões de chances estimadas para cada possibilidade de história da cadeia, mantendo-se o efeito da estratificação por tratamento são:

$$\psi_{\mathbf{h}_t=(0,0)} = 2,083; \quad \psi_{\mathbf{h}_t=(0,1)} = 2,082; \quad \psi_{\mathbf{h}_t=(1,0)} = 2,082; \quad \psi_{\mathbf{h}_t=(1,1)} = 2,087.$$

Considerando,

$$\text{Hipóteses : } \begin{cases} H_0 : \psi = 1; \\ H_a : \psi > 1. \end{cases}$$

A estatística do teste é $Z_{MH} = 2,43$, com nível descritivo igual a 0,0074, indicando forte evidência a favor do tratamento ativo.

CONSIDERAÇÕES PARA TRABALHOS FUTUROS

- explorar a questão da inclusão de efeitos aleatórios convenientes;
- extensão para mais de duas categorias de resposta;
- extensão dos testes para mais de dois grupos;
- estudar poder dos testes;
- estudar um resíduo mais apropriado para o modelo;
- estudar métodos para a possibilidade de observações omissas;
- considerar cadeias de alcance variável.

MODELOS DE TRANSIÇÃO COM EFEITOS ALEATÓRIOS

Sob um modelo de transição de Markov misto, as variáveis aleatórias $Y_{it} \mid \mathbf{h}_{it}, \mathbf{d}_i$ seguem um modelo linear generalizado.

$$Y_{it} \mid (\mathbf{h}_i, \mathbf{d}_i) = \exp \left\{ \frac{w_i}{\phi} \left[y_{it} \theta_{it} - b(\theta_{it}) \right] + c(y_{it}, \phi) \right\},$$

Para uma cadeia de ordem q , tem-se como preditor linear:

$$\eta_{it} = \underbrace{\mathbf{x}'_{it} \boldsymbol{\beta} + \sum_{r=1}^q \alpha_r y_{i(t-r)}}_{\mathbf{x}^{*'}_{it} \boldsymbol{\delta}} + \mathbf{z}'_{it} \mathbf{d}_i,$$

Simplificadamente,

$$\boldsymbol{\eta}_{it} = \mathbf{x}^{*'}_{it} \boldsymbol{\delta} + \mathbf{z}'_{it} \mathbf{d}_i \quad \text{e} \quad g(\mathbf{E}(Y_{it} \mid (\mathbf{h}_i, \mathbf{d}_i))) = g(\mu_{it}^C) = \boldsymbol{\eta}_{it}$$

$\mathbf{d}_i \Rightarrow$ vetor de efeitos aleatórios do i -ésimo indivíduo, tal que $\mathbf{d}_i \sim N_q(\mathbf{0}, \mathbf{G})$.

$\boldsymbol{\delta} \Rightarrow$ vetor de efeitos fixos.

MODELOS DE TRANSIÇÃO COM EFEITOS ALEATÓRIOS

Para dados binários, assumindo cadeia de alcance 1, a função de verossimilhança para δ e \mathbf{G} no caso estacionário é proporcional a:

$$L(\delta, \mathbf{G}; \mathbf{y}) \propto \prod_{i=1}^N \int \prod_{t=2}^{n_i} [\mu_{it}(\delta, \mathbf{G})]^{y_{it}} [1 - \mu_{it}(\delta, \mathbf{G})]^{1-y_{it}} f(\mathbf{d}_i; \mathbf{G}) d(\mathbf{d}_i),$$

Métodos usuais que permitem solução numérica aproximada das integrais e conseqüentemente otimização da função de verossimilhança:

- aproximação dos integrandos (aproximação de Laplace) \Rightarrow R;
- aproximação dos dados (QVP e QVM) \Rightarrow SAS (GLIMMIX);
- aproximação da integral (quadratura gaussiana) \Rightarrow SAS (NLMIXED) e R(glmML).

Para o caso do estudo epidemiológico sobre doença respiratória:

$$\eta_{it} = (\beta_0 + d_{i0}) + \mathbf{x}'_{it}\boldsymbol{\beta} + \alpha y_{i(t-1)}.$$

REFERENCIAL TEÓRICO BÁSICO

ANDERSON, T.W.; GOODMAN, L.A. Statistical Inference about Markov Chains. **Annals of Mathematical Statistics**. Ann Arbor, v. 28, p. 89 – 110, 1957.

AZZALINI, A. Maximum likelihood estimation of order n for stationary stochastic processes. **Biometrika**, Cambridge, v. 70, p. 381 – 388, 1983.

DIGGLE, P.J.; HEAGERTY, P.J.; LIANG, K.Y.; ZEGER, S.L. **Analysis of longitudinal data**. New York: Oxford University Press, 2002, 379 p.

LARA, I.A.R. **Modelos de transição para dados binários**. 2007, 128 p. Tese (Doutorado em Agronomia) - Universidade de São Paulo, Piracicaba, 2007.

MOLENBERGHS, G. VERBEKE, G. **Models for discrete longitudinal data**. New York: Springer-Verlag, 2005, 683 p.

NOLETO, A.Q. **Análise de Dados Longitudinais com Respostas Categorizadas através de Processos de Markov**. 1991, 93 p. Dissertação (Mestrado em Estatística)- Universidade Estadual de Campinas, Campinas, 1986.

WARE, J. H.; LIPSITZ, S.; SPEIZER, F. E. Issues in the Analysis of Repeated Categorical Outcomes. **Statistics in Medicine**, Chichester, v. 7, p. 95 – 107, 1988.

ZEGER, S.L.; LIANG, K.Y. An overview of methods for the analysis of longitudinal data. **Statistics in Medicine**, Chichester, v. 11, p. 1825 – 1839, 1992.