

Aplicação de modelos multiestados para análise de dados de sobrevivência

Trabalho de Conclusão de Curso em Estatística

Henrique Aparecido Laureano Emerson Rigoni
Profa. Dra. Suely Ruiz Giolo
(Orientadora)

Universidade Federal do Paraná
UFPR

23 de agosto de 2016

Roteiro

- 1 Conceitos básicos e metodologia usual
- 2 Representações de um modelo multiestado
- 3 Modelos multiestados: conceitos básicos
- 4 Inferências disponíveis
- 5 Qualidade do ajuste
- 6 Aplicação I
- 7 Aplicação II
- 8 Considerações finais

Conceitos básicos e metodologia usual

Análise de sobrevivência ou de confiabilidade:

Variável resposta é o tempo até a ocorrência do evento de interesse

Censura:

Não ocorrência do evento de interesse durante o período de acompanhamento ou ocorrência por causa distinta das envolvidas no estudo

Metodologia usual

Estimador de Kaplan-Meier¹ e Modelo de Cox²:

Consideram a existência de apenas um tipo de desfecho

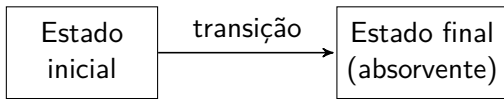
Aplicação limitada quando se quer estudar as causas do evento de interesse ou mais de um evento de interesse

¹Kaplan EL, Meier P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457-481

²Cox DR. (1972). Regression models and life tables. *Journal of the Royal Statistical Association, Series B*, 34(2):187-220

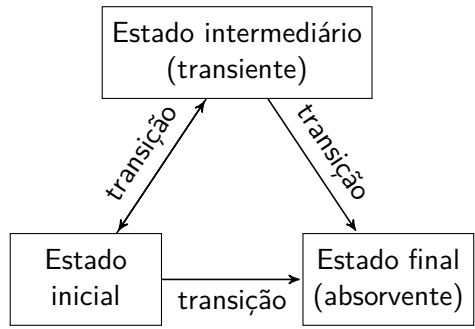
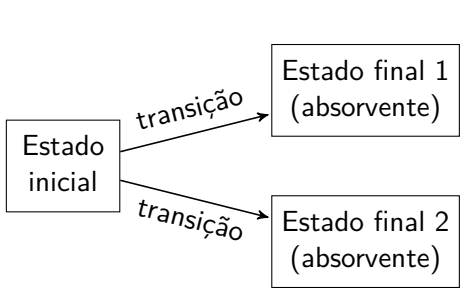
Representações de um modelo multiestado

Representação usual da situação de um dado de sobrevivência:



Multiestado:

Riscos competitivos:



Modelos multiestados: conceitos básicos

Modelo usual com vetor de covariáveis \mathbf{Z} em que é assumido efeito comum a todas as transições:

$$q_{rs}(\mathbf{Z}) = q_{rs}^0 \exp(\boldsymbol{\beta}^\top \mathbf{Z}),$$

em que q_{rs}^0 é a intensidade de transição ou taxa de falha de base para a transição do estado r para o estado s ,

$$q_{rs}(t) = \lim_{\delta t \rightarrow 0} \frac{P(X(t + \delta t) = s | X(t) = r)}{\delta t}.$$

Um modelo multiestado pode ser:

- Não markoviano
- Semimarkoviano
- **Markoviano**

Modelos markovianos são fundamentados no pressuposto markoviano de que uma futura transição depende apenas do estado atual

Um modelo multiestado markoviano pode ser de três tipos:

- Paramétrico
- Não paramétrico
- Semiparamétrico

Modelo paramétrico (*package* msm³ do R⁴):

Foco da implementação: dados de sobrevivência intervalar

Distribuição de probabilidade assumida para o tempo médio de permanência em cada estado transiente r , comumente exponencial

Ele pode ser de dois tipos:

Tempo homogêneo: intensidades de transição constantes ao longo do tempo (independentes de t)

Tempo não homogêneo: intensidades de transição variáveis ao longo do tempo, constantes sob segmentos

³Jackson CH. (2011). Multi-State Models for Panel Data: The msm Package for R. *Journal of Statistical Software*, 38(8):1-29. <http://www.jstatsoft.org/v38/i08/>

⁴R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Modelo (não e) semiparamétrico (*package* `mstate`^{5 6 7} do R):

Foco da implementação: dados de sobrevivência não intervalar

Modelo de Cox estratificado por transição.

Na ausência de covariáveis temos um modelo não paramétrico

Além do modelo usual, permite a especificação de modelos mais elaborados:

- Diferentes efeitos das covariáveis em cada transição
- Intensidades de transição proporcionais
- Covariáveis que aparecem apenas em algumas transições

⁵de Wreede LC, Fiocco M, Putter H. (2011). `mstate`: An R Package for the Analysis of Competing Risks and Multi-State Models. *Journal of Statistical Software*, 38(7):1-30. <http://www.jstatsoft.org/v38/i07/>

⁶de Wreede LC, Fiocco M, Putter H. (2010). The `mstate` Package for Estimation and Prediction in Non- and Semi-Parametric Multi-State and Competing Risks Models. *Computer Methods and Programs in Biomedicine*, (99):261-274

⁷Putter H, Fiocco M, Geskus RB. (2007). Tutorial in Biostatistics: Competing Risks and Multi-State Models. *Statistics in Medicine*, 26:2389-2340

Inferências disponíveis

Diferenciais em relação ao estimador de Kaplan-Meier e ao modelo de Cox:

Probabilidades de transição e de sobrevivência:

Modelo paramétrico: Equações diferenciais de Kolmogorov ⁸
(KFE, *Kolmogorov Forward Equations*)

Modelo (não e) semiparamétrico: estimador similar ao de Aalen-Johansen

Tempos médios esperados:

Modelo paramétrico: de permanência em estados,
através das intensidades de transição

Modelo (não e) semiparamétrico: para transição entre estados,
através das probabilidades de transição

⁸Cox DR, Miller HD. (1965). *The theory of stochastic processes*. Chapman & Hall

Qualidade do ajuste

Modelo paramétrico:

Métodos formais (estatística de teste e um valor p): em geral de difícil obtenção, ou assintóticos, ou instáveis, o que os tornam viáveis em poucas situações

Métodos informais: comparação das curvas de sobrevivência;
contagens de prevalência

Modelo (não e) semiparamétrico:

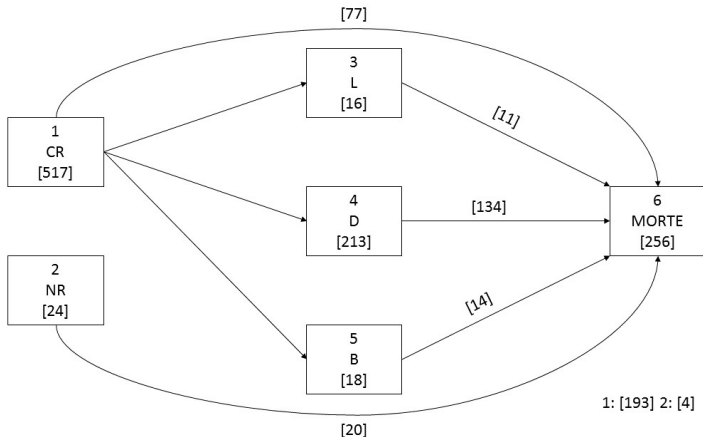
Verificação usual: análise gráfica dos resíduos e verificação da suposição de taxas de falha proporcionais (na presença de covariáveis)

Resíduos:

- Cox-Snell
- *Martingal*
- *Deviance*
- Schoenfeld

Aplicação I

Follic^{9 10}: 541 pacientes com linfoma de células foliculares no *Princess Margaret Hospital* de Toronto



CR: pacientes com remissão

NR: pacientes sem remissão

L: recidiva local D: recidiva distante

B: recidiva local e distante

⁹Pintilie, M. (2006). *Competing risks: a practical perspective*. John Wiley & Sons

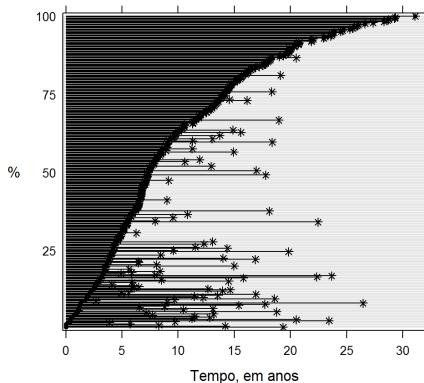
¹⁰<https://www.jstatsoft.org/article/view/v038i02>

Follic

Entre os pacientes que não entraram em óbito, cerca de 75% tiveram um tempo de primeira falha (recidiva ou censura) menor ou igual a quinze anos

Pacientes que não entraram em óbito (285 ou 53%)

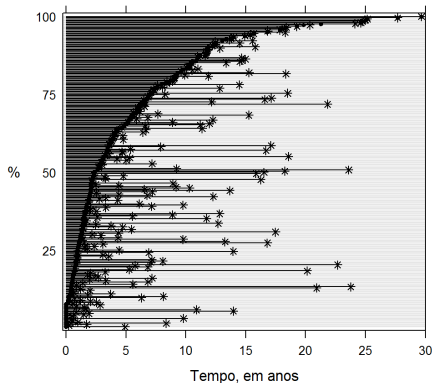
- Tempo até a primeira falha
- * Tempo de sobrevivência



Entre os pacientes que entraram em óbito, cerca de 90% tiveram um tempo de primeira falha (recidiva ou óbito) menor ou igual a quinze anos

Pacientes que entraram em óbito (256 ou 47%)

- Tempo até a primeira falha
- * Tempo de sobrevivência



*Aqui, o termo **falha** foi utilizado num sentido mais amplo, englobando também as censuras*

Modelagem

Covariáveis: Estágio clínico (I [362] ou II [179]),
Tratamento (radioterapia [423] ou radio e quimioterapia [118]),
Idade (< 58 anos (mediana) [263] ou ≥ 58 anos [278]),
Hemoglobina (< 140 g/l (mediana) [247] ou ≥ 140 g/l [294])

Covariáveis significativas:

541 pacientes (dois estados iniciais)

Modelo paramétrico: todas as quatro

Modelo (não e) semiparamétrico: Estágio clínico, Tratamento e Idade

517 (um estado inicial,
sem os pacientes que não entraram em remissão)

Modelo paramétrico: todas as quatro

Modelo (não e) semiparamétrico: todas as quatro

Follic: modelo multiestado markoviano paramétrico

Probabilidades de transição entre estados para o período de 1 ano

56 pacientes

Estágio clínico I, hemoglobina < 140 g/l, idade < 58 anos e tratados com radioterapia

		DESTINO					
		1	2	3	4	5	6
ORIGEM	1	0.949	0	0.005	0.041	0.002	0.001
	2	0	0	0	0	0	1
	3	0	0	0.952	0	0	0.048
	4	0	0	0	0.945	0	0.055
	5	0	0	0	0	0.869	0.13
	6	0	0	0	0	0	1

13 pacientes

Estágio clínico II, hemoglobina ≥ 140 g/l, idade ≥ 58 anos e tratados com radioterapia e quimioterapia

		DESTINO					
		1	2	3	4	5	6
ORIGEM	1	0.855	0	0,003	0.091	0.015	0,036
	2	0	0.347	0	0	0	0,653
	3	0	0	0.03	0	0	0.97
	4	0	0	0	0.648	0	0.352
	5	0	0	0	0	0.552	0.448
	6	0	0	0	0	0	1

1: remissão; 2: sem remissão; 3: recidiva local; 4: recidiva distante; 5: recidiva local e distante; 6: morte

Tempos médios de permanência, em anos, para cada estado transiente

56 pacientes

Estágio clínico I, hemoglobina < 140 g/l, idade < 58 anos e tratados com radioterapia

		Estimativa pontual	Erro padrão	Mínimo (IC de 95%)	Máximo (IC de 95%)
		ESTADO	1	19.31	2.57
2	0.03		35392	0	∞
3	20.23		11.07	6.92	59.11
4	17.69		3.51	11.99	26.11
5	7.15		5.55	1.56	32.77

13 pacientes

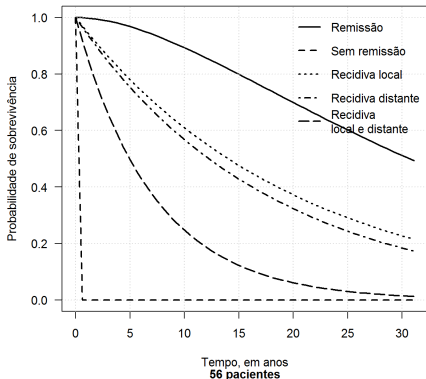
Estágio clínico II, hemoglobina ≥ 140 g/l, idade ≥ 58 anos e tratados com radioterapia e quimioterapia

		Estimativa pontual	Erro padrão	Mínimo (IC de 95%)	Máximo (IC de 95%)
		ESTADO	1	6.4	1.28
2	0.94		1.09	0.1	9.12
3	0.28		0.23	0.06	1.36
4	2.3		0.79	1.18	4.51
5	1.68		1.44	0.31	9.03

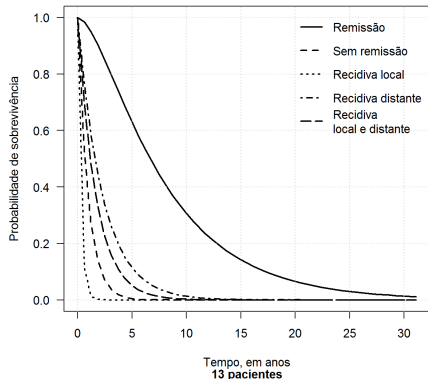
Follic: modelo multiestado markoviano paramétrico

Curvas de sobrevivência para os estados transitórios

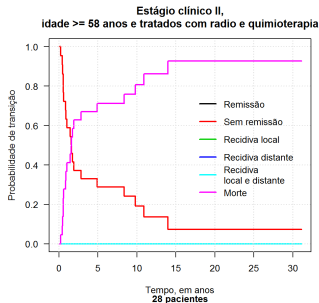
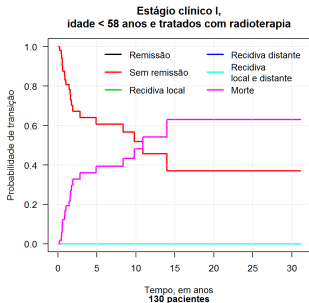
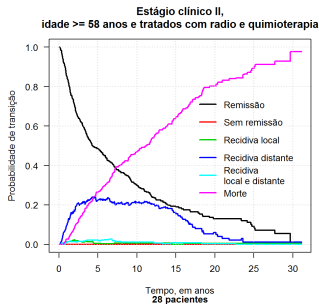
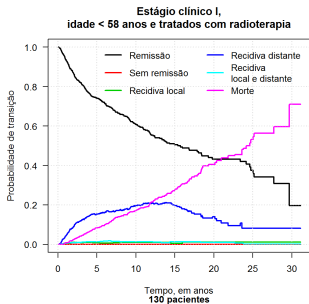
**Estágio clínico 1, hemoglobina < 140 g/l,
idade < 58 anos e tratados com radioterapia**



**Estágio clínico 2, hemoglobina >= 140 g/l,
idade >= 58 anos e tratados com radio e quimioterapia**



Follic: modelo multiestado markoviano (não e) semiparamétrico



Follic: modelo multiestado markoviano (não e) semiparamétrico

Tempos esperados de permanência, em anos, para cada transição entre estados

130 pacientes

Estágio clínico I, idade < 58 anos e tratados com radioterapia

		DESTINO					
		1	2	3	4	5	6
ORIGEM	1	16.67	0	0.3	4.16	0.28	9.68
	2	0	14.8	0	0	0	16.3
	3	0	0	10.73	0	0	20.37
	4	0	0	0	13.25	0	17.85
	5	0	0	0	0	6.88	24.22
	6	0	0	0	0	0	31.1

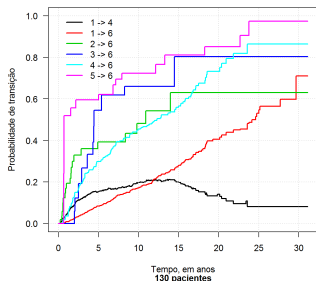
1: remissão; 2: sem remissão; 3: recidiva local; 4: recidiva distante; 5: recidiva local e distante; 6: morte

28 pacientes

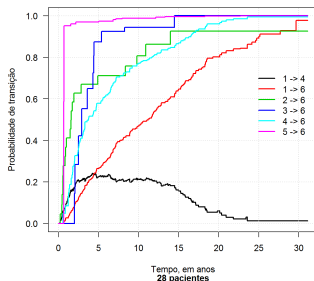
Estágio clínico II, idade ≥ 58 anos e tratados com radioterapia e quimioterapia

		DESTINO					
		1	2	3	4	5	6
ORIGEM	1	8.28	0	0.19	3.55	0.24	18.84
	2	0	5.58	0	0	0	25.51
	3	0	0	3.85	0	0	27.24
	4	0	0	0	6.36	0	24.74
	5	0	0	0	0	0.97	30.13
	6	0	0	0	0	0	31.1

Estágio clínico I, idade < 58 anos e tratados com radioterapia

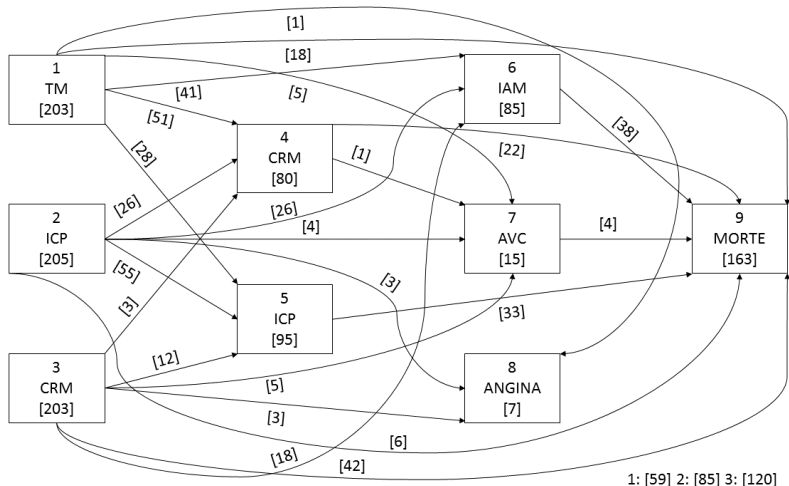


Estágio clínico II, idade ≥ 58 anos e tratados com radio e quimioterapia



Aplicação II

DACM: 611 pacientes com doença arterial coronariana multiarterial



(1) TM: Tratamento Médico

(2) ICP: Intervenção Coronária Percutânea (angioplastia)

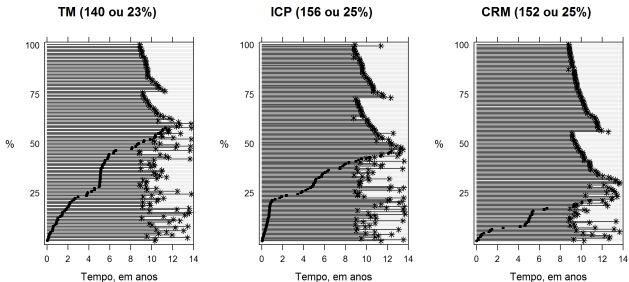
(3) CRM: Cirurgia de Revascularização Miocárdica

Evento isquêmico: (6) IAM: Infarto Agudo do Miocárdio

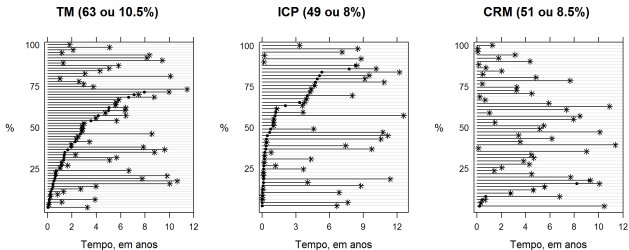
(7) AVC: Acidente Vascular Cerebral

(8) Angina

Pacientes que não entraram em óbito (448)



Pacientes que entraram em óbito (163)



Tempo até a primeira falha • Tempo de sobrevivência *

Modelagem

Covariáveis: Grupo de risco (médio [157] ou alto [454]),
Histórico de IAM (não [341] ou sim [270])

O grupo de risco foi determinado com base nas covariáveis:

sexo, idade, colesterol da lipoproteína de alta densidade, colesterol total, histórico de hipertensão, fumo e histórico de diabetes

Covariáveis significativas:

Modelo completo (9 estados)

Modelo paramétrico: sem covariáveis

Modelo (não e) semiparamétrico: sem covariáveis

Modelo reduzido (7 estados)

Modelo paramétrico: grupo de risco

Modelo (não e) semiparamétrico: sem covariáveis

Modelos marginais completos e reduzidos (7 e 5 estados)

6 modelos paramétricos e 6 modelos (não e) semiparamétricos

DACM: modelo multiestado markoviano paramétrico

7 pacientes tiveram angina (8), mas todos permaneceram nesse estado, assim, a angina foi considerado como um estado absorvente

PROBABILIDADES DE TRANSIÇÃO ENTRE ESTADOS PARA O PERÍODO DE 1 ANO

		DESTINO								
		1 (TM)	2 (ICP)	3 (CRM)	4 (CRM)	5 (ICP)	6 (IAM)	7 (AVC)	8 (ANGINA)	9 (MORTE)
ORIGEM	1 (TM)	0.882 [0.86 - 0.897]	0	0	0.041 [0.03 - 0.053]	0.022 [0.015 - 0.031]	0.032 [0.024 - 0.043]	0.004 [0.001 - 0.009]	0.0008 [0 - 0.005]	0.018 [0.012 - 0.026]
	2 (ICP)	0	0.907 [0.886 - 0.92]	0	0.019 [0.014 - 0.028]	0.041 [0.031 - 0.053]	0.019 [0.013 - 0.028]	0.003 [0.001 - 0.008]	0.002 [0 - 0.007]	0.007 [0.005 - 0.013]
	3 (CRM)	0	0	0.949 [0.935 - 0.958]	0.002 [0 - 0.005]	0.007 [0.004 - 0.012]	0.01 [0.006 - 0.016]	0.003 [0.001 - 0.007]	0.002 [0 - 0.006]	0.026 [0.02 - 0.035]
	4 (CRM)	0	0	0	0.954 [0.929 - 0.968]	0	0	0.002 [0 - 0.013]	0	0.044 [0.03 - 0.065]
	5 (ICP)	0	0	0	0	0.941 [0.918 - 0.958]	0	0	0	0.059 [0.042 - 0.081]
	6 (IAM)	0	0	0	0	0	0.916 [0.885 - 0.939]	0	0	0.083 [0.06 - 0.115]
	7 (AVC)	0	0	0	0	0	0	0.905 [0.769 - 0.964]	0	0.094 [0.035 - 0.23]
	8 (ANGINA)	0	0	0	0	0	0	0	1	0
	9 (MORTE)	0	0	0	0	0	0	0	0	1

(1) TM: Tratamento Médico

(2) ICP: Intervenção Coronária Percutânea (angioplastia)

(3) CRM: Cirurgia de Revascularização Miocárdica

Evento isquêmico: (6) IAM: Infarto Agudo do Miocárdio

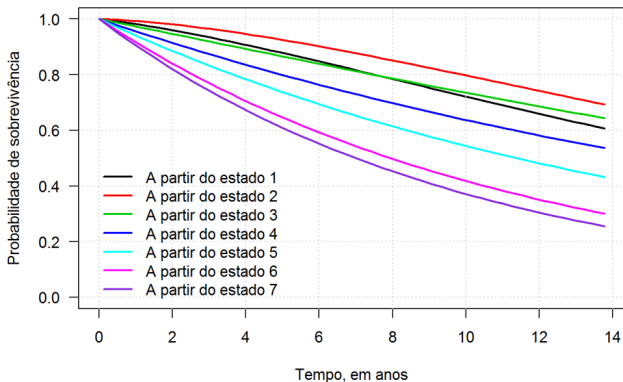
(7) AVC: Acidente Vascular Cerebral

(8) Angina

DACM: modelo multiestado markoviano paramétrico

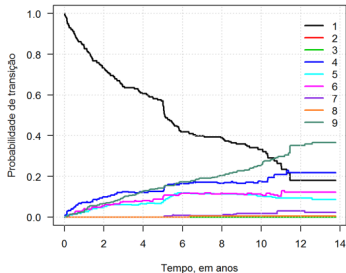
Tempos médios de permanência, em anos, para cada estado transiente					
	Estimativa pontual	Erro padrão	Mínimo - IC de 95%	Máximo - IC de 95%	
ESTADO	1: TM	7.98	0.66	6.71	9.27
	2: ICP	10.26	0.91	8.38	11.99
	3: CRM	19.25	2.13	14.99	23.59
	4: CRM	21.43	4.59	13.51	31.13
	5: ICP	16.41	3.09	11.7	23.98
	6: IAM	11.46	1.9	8.27	16.15
	7: AVC	10.09	5.91	3.9	26.23

Curvas de sobrevivência para os estados transientes

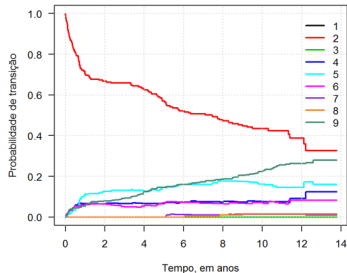


DACM: modelo multiestado markoviano (não e) semiparamétrico

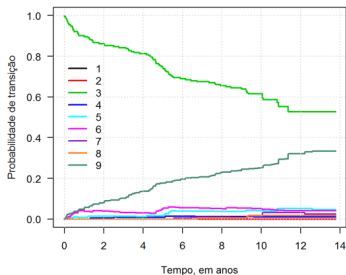
Estado inicial 1 (TM)



Estado inicial 2 (ICP)



Estado inicial 3 (CRM)

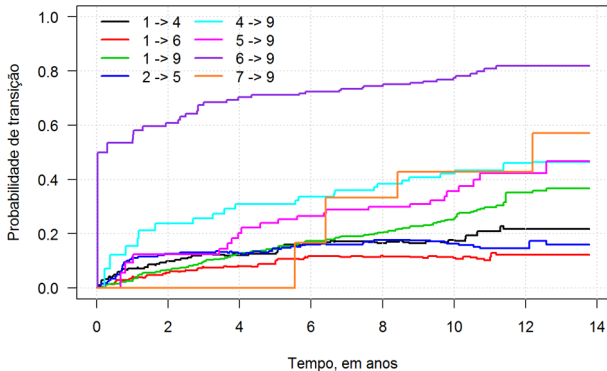


- 1: TM
- 2: ICP
- 3: CRM
- 4: CRM
- 5: ICP
- 6: IAM
- 7: AVC
- 8: ANGINA
- 9: MORTE

DACM: modelo multiestado markoviano (não e) semiparamétrico

Tempos esperados de permanência, em anos, para cada transição entre estados

		DESTINO								
		1: TM	2: ICP	3: CRM	4: CRM	5: ICP	6: IAM	7: AVC	8: ANGINA	9: MORTE
ORIGEM	1: TM	6.38	0	0	2.11	1.14	1.31	0.14	0.04	2.64
	2: ICP	0	7.28	0	1.07	1.97	0.92	0.1	0.09	2.33
	3: CRM	0	0	9.65	0.13	0.43	0.59	0.14	0.09	2.73
	4: CRM	0	0	0	9.04	0	0	0.05	0	4.67
	5: ICP	0	0	0	0	10.05	0	0	0	3.72
	6: IAM	0	0	0	0	0	3.89	0	0	9.88
	7: AVC	0	0	0	0	0	0	10.44	0	3.33
	8: ANGINA	0	0	0	0	0	0	0	13.77	0
	9: MORTE	0	0	0	0	0	0	0	0	13.77



Considerações finais

Considerações finais

- Ambos os modelos geraram inferências muito similares, independente do delineamento (dados de sobrevivência intervalar ou não intervalar)
- Modelo (não e) semiparamétrico se mostrou mais robusto
- Ambos os modelos se mostraram altamente dependentes do tamanho amostral
 - Grande amostra
 - ⊗ Grande amostra em cada transição

Obrigado pela atenção!