

Cronograma – Avançado

Dia	Tema
2ª	Efeitos não lineares
3ª	Eventos Múltiplos
4ª	Eventos competitivos
5ª	Fragilidade

Outline

- 1 Efeitos não-lineares
- 2 Eventos Múltiplos
- 3 Fragilidade

Objetivo

Avaliar a forma funcional de covariáveis.
 Incluir termo de ajuste não-paramétrico *spline* ao modelo de Cox.
 Avaliar o ajuste de modelos com função de suavização *spline*.
 Interpretar modelo com função de suavização *spline*

Bibliografia

- Hastie, T.; Tibshirani, R. *Generalised Additive Models*. Chapman & Hall, 1990.
- Wood, S.N. *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC Texts in Statistical Science Series, 2006.
- Faraway, J.J. *Extending the Linear Model with R*. Chapman & Hall/CRC Texts in Statistical Science Series, 2006.
- Wasserman, L. *All of Nonparametric Statistics*. Springer, 2006

O que é o GAM

- É uma extensão do GLM na qual o preditor linear η não está limitado à essa relação linear – reta – entre variável independente e variável resposta, mas é uma função aditiva dos x_i 's.

$$\eta = \beta_0 + f_1(x_1) + f_2(x_2) + \dots$$

- Quando usar? Quando o efeito da variável muda conforme seu valor.

Exemplos

- Idade
- Renda
- Distância do Serviço de Saúde
- Tempo
- Espaço
- Adesão :

Alternativas quando o efeito muda

- dividir a variável em classes – literatura ou outros pontos de corte (por exemplo, percentis)
- transformações paramétricas (quadrado ou log) – interpretabilidade
- função de suavização

A função de suavização

- Qualquer função de suavização que linearize a relação $x \rightarrow$ tempo de sobrevivida
- Pressupostos:
 - Os mesmos: independência dos erros.
 - A forma como x atua sobre *tempo* deve ser suave, sem grandes saltos.
 - Identificabilidade: duas variáveis x e z incluídas através de f_x e f_z somente são estimáveis dentro de uma constante. Um valor fixo acrescentado em x e subtraído de z dará o mesmo preditor linear: confundimento!

As funções f

- Kernel density: estimativa de densidade de probabilidade – equivale a uma média ponderada.
- Loess: locally weighted scatterplot smoothing – estende a mesma idéia, mas os valores entram em uma regressão ponderada (não apenas média, ou seja, inclui uma inclinação)
- Splines: cúbica, p-splines, thin plate – classe de funções polinomiais estimadas para intervalos de x e conectadas entre si suavemente.
- Wavelets: não me perguntem nada!
- Podem ser uni ou multi-dimensionais

Comparando

http://en.wikipedia.org/wiki/Kernel_Smoothing

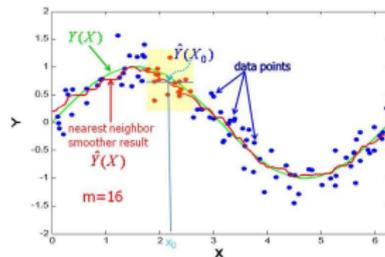


Figura: Vizinho Mais Próximo

Comparando

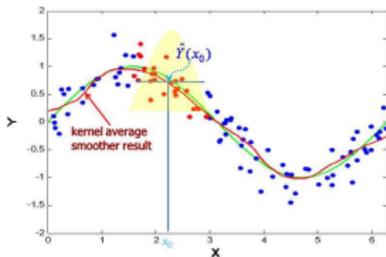


Figura: Média Móvel Ponderada

Comparando

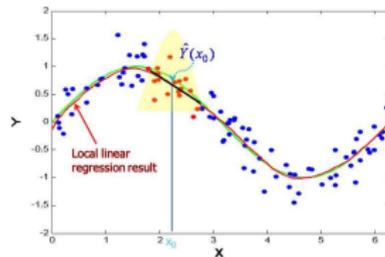


Figura: Regressão Linear Local Ponderada

Lowess – 3 conceitos essenciais

- vizinhança – ajuste **local**;
- **ponderação** – definida em função da distância entre o ponto para o qual se deseja estimar um valor suavizado e as observações vizinhas;
- **regressão** – que expressa a forma da relação entre z e y dentro da vizinhança.
- A função de regressão pode ser um polinômio de grau 1 (linear), de grau 2 (quadrática) ou de grau zero = uma média móvel.

Lowess – os passos

Para estimar o valor da suavização *lowess* no ponto z_0 :

- 1 definir a **largura de banda** h , que será a **vizinhança** do ponto z_0 ;
- 2 identificar os pontos que pertencem a essa vizinhança de z_0 ;
- 3 atribuir pesos para cada ponto da vizinhança de z_0 através de uma função k , chamada **kernel**;
- 4 calcular o valor suavizado de z_0 , $s(z_0)$ através da estimação por mínimos quadrados ponderados para a regressão adotada dentro da vizinhança.

Lowess

- funções kernel são usualmente simétricas
- eventualmente gerem quebras pouco suaves, ver gráfico
- escolha da forma da função de ponderação afeta pouco a suavização final estimada, mais importante é a largura de banda.
- parâmetro que define a largura de banda não é estimado automaticamente
- validação cruzada, mas mesmo assim...

Lowess

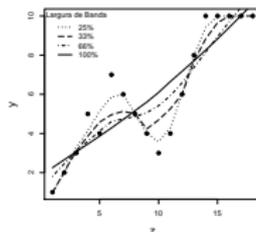


Figura: Suavização lowess com diferentes larguras de banda

Splines

- Splines são um tipo particular de função de suavização, definida como polinômios segmentados (*piecewise*)
- Restrição: cada curva parcial se une à seguinte obrigatoriamente de forma suave, sem ângulos
- Definição dos segmentos no caso do *spline* é dada através de **nós**
- Os polinômios mais utilizados são funções cúbicas \rightarrow flexibilidade

Splines

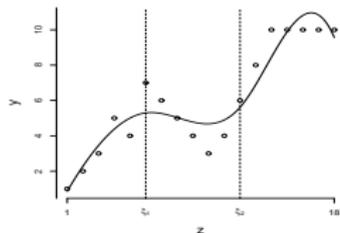


Figura: Polinômios segmentados de grau 3 (*spline* cúbico) ajustado com dois nós

Como chegar ao Spline

- funções de base $h_1(z) = 1$, $h_2(z) = z$, $h_3(z) = z^2$ e $h_4(z) = z^3$
- serão ajustados 3 polinômios:
 - um no intervalo abaixo de ξ_1
 - o segundo no intervalo entre ξ_1 e ξ_2
 - o terceiro polinômio no intervalo acima de ξ_2
- O problema em ajustar estes três polinômios separadamente é que não há como garantir a continuidade nos nós ξ_1 ξ_2

Como chegar ao Spline

- Para garantir continuidade, definimos funções adicionais para compor a base:

$$h_5(z) = (z - \xi_1)_+^3$$

$$h_6(z) = (z - \xi_2)_+^3$$

$$(z - \xi)_+^3 = \begin{cases} (z - \xi)^3, & \text{se } z > \xi; \\ 0, & \text{caso contrário.} \end{cases}$$

- O *spline* cúbico é dado pela combinação linear dessas funções da base
- $s(z) = \sum_{j=1}^6 \beta_j h_j(z)$

Como chegar ao Spline: polinômios?

$$s(z) = \begin{cases} \beta_1 + \beta_2 z + \beta_3 z^2 + \beta_4 z^3, & \text{se } z \leq \xi_1 \\ \beta_1 + \beta_2 z + \beta_3 z^2 + \beta_4 z^3 + \beta_5 (z - \xi_1)^3, & \text{se } \xi_1 < z \leq \xi_2 \\ \beta_1 + \beta_2 z + \beta_3 z^2 + \beta_4 z^3 + \beta_5 (z - \xi_1)^3 + \beta_6 (z - \xi_2)^3, & \text{se } z > \xi_2 \end{cases}$$

Como chegar ao Spline: estimar β s

- Começamos com o tantos nós quantas observações, conectadas por um polinômio cúbico único \rightarrow interpolação suave e perfeita sobre os dados observados \rightarrow "overfitting"
- Precisamos de curva que permita entender a relação entre z e y
- minimização de:

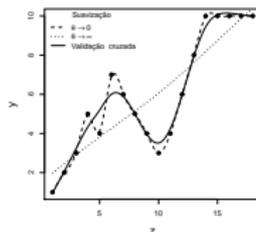
$$\sum_{i=1}^n (y_i - s(z_i))^2 + \theta \int (s''(z))^2 dz,$$

- θ é o parâmetro que controla o grau de suavização.
- A primeira parcela mede a proximidade da curva aos dados
- o termo $\int (s''(z))^2 dz$ penaliza a curvatura ou a rugosidade da função.

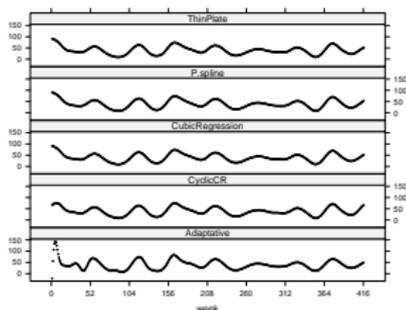
Como chegar ao Spline: cenários

- se $\theta = 0$, temos que minimizar $\sum_{i=1}^n (y_i - s(z_i))^2$, o que fornece a solução $s(z_i) = y_i$ e portanto, a curva passa por todos os pontos, ou seja, interpola os dados;
- se $\theta \rightarrow \infty$, a expressão é minimizada se $s''(z)^2 = 0$ para todo z e isto é obtido se $s(z)$ for uma reta.

Como chegar ao Spline: cenários

Figura: Curvas ajustadas por spline com diferentes valores de θ

Tipos de splines



Escolha do parâmetro de suavização

- *lowess* e *spline* dependem dos respectivos parâmetros de suavização h e θ
- escolha arbitrária do grau de suavização:
 - pode ser vantagem: permite explorar diferentes comportamentos, em diferentes escalas.
 - da tendência mais global (θ é pequeno ou h grande) até comportamentos locais
 - Por outro lado, muita subjetividade nessa escolha pode ser indesejável, valorizando características locais pouco importantes, particularmente quando os dados são esparsos em determinados intervalos da variável suavizada.

Validação cruzada

- escolher h ou θ que minimize as diferenças entre um conjunto de ajustes nos quais foi retirada aleatoriamente uma observação

$$\frac{1}{n} \sum_{i=1}^n (y_i - s_{-i}(z_i))^2$$

- $s_{-i}(z_i)$ é o estimador obtido quando se omite o par (z_i, y_i)

Por que NÃO usar sempre o GAM?

- Modelos estatísticos são uma forma de **explicar** o observado, não simplesmente reproduzir – cuidado com o “overfitting”.
- Modelar, além de estimar o efeito, inclui estimar a confiança (erro padrão, intervalos de confiança, comparação de modelos), e os modelos paramétricos são em geral melhores nisso.
- Os modelos paramétricos são mais eficientes se a especificação do modelo é correta.
- Exigem menor número de observações.

Modelos de regressão penalizados spline

- **p-splines** – estão incorporados ao modelo de Cox, na biblioteca *survival*.
- penalização direta do parâmetro β
- intervalos entre os nós devem ser espaçados regularmente
- menor número possível de nós E a menor soma dos resíduos quadráticos para a covariável
- função de penalização tem um parâmetro $\theta \rightarrow$ indica quão afastado de uma reta os dados estão
 - $\theta \rightarrow 0$, a solução converge para uma reta;
 - quando $\theta \rightarrow 1$, a curva passa por todos os pontos.

Exemplo – CTI-INCA

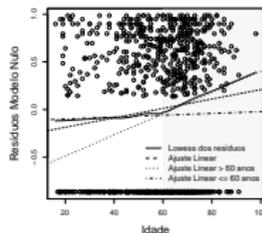


Figura: Resíduos martingale do modelo nulo contra a variável idade

```
Call:
coxph(formula = Surv(tempo, status) ~ ppspline(idade),
      data = cti, x = T)
n= 862
```

	coef	se(coef)	se2	Chisq	DF	p
ppspline(idade),linear	0.0089	0.0027	0.0027	10.39	1.00	0.0013
ppspline(idade),nonlin				3.27	3.06	0.3600

	exp(coef)	exp(-coef)	lower	.95 upper	.95
ps(idade)2	0.962	1.039	0.450	2.05	
ps(idade)3	0.918	1.089	0.291	2.89	
ps(idade)4	0.866	1.154	0.251	2.99	
ps(idade)5	0.916	1.092	0.275	3.05	
ps(idade)6	0.983	1.018	0.307	3.14	
ps(idade)7	0.960	1.042	0.304	3.03	
ps(idade)8	1.116	0.896	0.353	3.53	
ps(idade)9	1.287	0.777	0.406	4.08	
ps(idade)10	1.495	0.669	0.465	4.81	
ps(idade)11	1.736	0.576	0.507	5.95	
ps(idade)12	2.066	0.484	0.461	9.26	
ps(idade)13	2.482	0.403	0.317	19.42	

```
Iterations: 6 outer, 15 Newton-Raphson
Theta= 0.921
Degrees of freedom for terms= 4
Rsquare= 0.016 (max possible= 0.999 )
```

```
> idade.linear <- coxph(Surv(tempo,status) ~
idade, data=cti)
> anova(idade.linear, idade.spl)
```

```
Analysis of Deviance Table
Cox model: response is Surv(tempo, status)
Model 1: ~ idade
Model 2: ~ ppspline(idade)
```

```
loglik Chisq Df P(>|Chi1)
1 -3183.3
2 -3181.4 3.7879 11 0.9757
```

Alternativas

- Particionar variável
 - o β estimado corresponde ao valor médio da variável \rightarrow interessa?
 - critérios \rightarrow OLHEM!
- Regressão segmentada \rightarrow criar duas (ou +) variáveis, modeladas em conjunto (como uma dummy) assim:

$$\begin{aligned} var1 &\leftarrow \begin{cases} \text{se variável} \leq \text{corte} & = \text{variável} \\ \text{caso contrário} & = 0 \end{cases} \\ var2 &\leftarrow \begin{cases} \text{se variável} > \text{corte} & = \text{variável} - \text{valor corte} \\ \text{caso contrário} & = 0 \end{cases} \end{aligned}$$

- E se os efeitos das variáveis não for semelhante para os dois grupos? Modelar em separado cada um.

Outline

- 1 Efeitos não-lineares
- 2 **Eventos Múltiplos**
- 3 Fragilidade

Objetivos

- Identificar os diversos modelos existentes para análise de sobrevivida com eventos múltiplos.
- Definir conceitualmente os diversos modelos, descrever seus pressupostos e propriedades.
- Indicar o tipo de modelo adequado a cada situação observada.
- Interpretar as estimativas ajustadas em cada modelo.

Exemplos

- Eventos que podem acontecer mais de uma vez para um mesmo indivíduo:
 - gestações,
 - internações,
 - cáries,
 - infartos do miocárdio
 - fraturas.
- Diferentes tipos de eventos decorrentes de um mesmo fator de risco em estudo:
 - efeitos adversos de medicamentos,
 - doenças oportunistas da aids,
 - óbito de um paciente, quando são possíveis várias causas.

Perguntas

- Quais são os fatores de risco associados aos tempos até as ocorrências de eventos em um mesmo indivíduo?
- Como analisar diferentes eventos como desfecho de uma mesma situação de risco?

O que difere

- o registro de mais de um tempo por indivíduo → tempo até cada cárie, tempo até o óbito por diferentes causas
- a utilização direta do modelo de Cox **não** serve, mesmo com processo de contagem, pois os intervalos de tempo de cada indivíduo **não** são independentes entre si
- não havendo independência entre observações → erro na estimação dos parâmetros e dos intervalos de confiança
- composição do grupo de risco → os indivíduos podem (ou não) se manter no grupo de risco mesmo após a ocorrência de algum evento

Estratégias possíveis

- ajustar Poisson → variável dependente é uma contagem dos eventos
 - não diferencia eventos próximos no tempo de distantes
- modelar apenas **um** evento
 - até o primeiro, entre o segundo e o terceiro, ...
 - apenas uma das causas de interesse censurando a outra e vice-versa
 - simplifica, mas pode não ser adequada
- modelos multinível (efeitos aleatórios, fragilidade, mistos) → última aula

Modelos Marginais

Classificação

O que deve ser observado:

- a população em risco → montagem do banco de dados
- o risco basal (λ_0) → no modelo semi-paramétrico não é estimado, mas pode ser o mesmo para todos os eventos ou ser diferente conforme o estrato
- a estrutura temporal (ordenação) → definida simultaneamente pela forma do banco de dados e pela indicação de presença de estratos ordenados
- a estrutura de dependência entre os eventos → métodos robustos para corrigir a variância dos estimadores

Classificação dos eventos

- **competitivos** → só é possível observar o tempo até a ocorrência do primeiro evento que impede a realização de qualquer outro (óbito por diferentes causas e um mesmo fator de risco) → próxima aula
- **paralelos** → a ocorrência de um evento não exclui a ocorrência de outro evento e não há qualquer ordem preferencial (doenças oportunistas, efeitos colaterais, perda de dente)
- **ordenados** → a sucessão de tempos segue obrigatoriamente uma ordem, que é dada seja pela estrutura de datas de início e fim de cada evento, seja por assumir-se uma ordem nos estratos de risco relacionados a cada desfecho ou ambos

Eventos ordenados

Combinando definição de grupo sob risco e estrutura do risco basal (λ_0) são três tipos de problemas:

Eventos ordenados independentes - o risco basal é o mesmo em todos os intervalos de tempo analisados. O indivíduo retorna ao grupo de risco após cada evento.

Ex: IRA. Ordenado, pois não podem ocorrer simultaneamente, e o indivíduo sai do risco na vigência de um episódio.

Modelo de incrementos independentes ou de Andersen-Gill (**AG**)

Eventos ordenados

Eventos ordenados estruturados - o indivíduo só entra em risco de sofrer o n -ésimo evento depois que o evento de ordem $n - 1$ tiver ocorrido. O risco basal para um segundo evento é zero até que o primeiro ocorra enquanto que o risco do terceiro evento é zero até que o segundo ocorra.

Ex. IAM. O risco de base se altera à medida em que o indivíduo sofre novos eventos, ou seja, os eventos, além de ordenados no tempo, são ordenados segundo o risco basal.

Modelo **condicional**, ou modelo **PWP** (Prentice, Williams & Peterson) (todos são marginais!)

Eventos ordenados

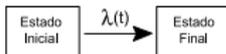
Eventos ordenados com risco concomitante - o indivíduo está sempre em risco de sofrer o primeiro, segundo, terceiro, n -ésimo evento. Cada evento tem um risco basal diferente, dependendo de sua posição na sequência de eventos, mas não do tipo de evento.

Ex. reações adversas (cefaléia, náuseas, tremores e sonolência). O risco basal para cada efeito adverso depende apenas da ordem em que ele aparece, e não do tipo de efeito. O tempo até a ocorrência é contado sempre a partir de zero para todos os indivíduos (não se utiliza a estrutura de contagem).

Modelo **marginal**, também chamado modelo **WLW** (Wei, Lin & Weissfeld) Na prática não utilizamos, retirado da segunda edição

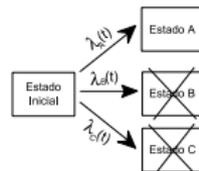
Modelos multiestados

A - Modelo de Cox



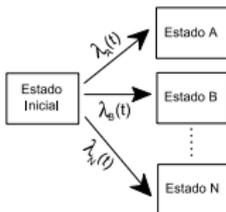
Modelos multiestados

B - Eventos Competitivos



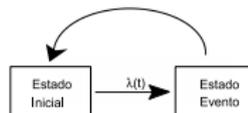
Modelos multiestados

C - Eventos Paralelos



Modelos multiestados

D - Incrementos Independentes ou AG



Modelos multiestados

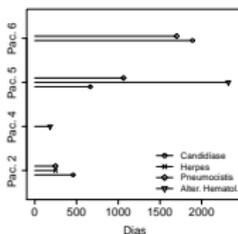


Modelos marginais

- Modelar-se a resposta média como função das covariáveis
- Interpretação igual
- Estimativa robusta da variância por *Jackknife*
- Etapas:
 - identificar conceitualmente o modelo
 - definir está sob risco em cada momento → construir banco de dados
 - ajustar modelo de Cox simples
 - modelar usando mais de uma estrutura

Eventos paralelos

Tempo até a ocorrência de algumas doenças oportunistas em quatro pacientes portadores do vírus HIV.



Eventos paralelos – banco de dados

Sem ordenação ou sem interesse em considerar a ordenação.

reg	sex	esc	idade	udi	sexual	ini	oport	fim	status	tempo
2	M	2	38	1	0	2884	Candida	3343	1	459
2	M	2	38	1	0	2884	Herpes	3133	1	249
2	M	2	38	1	0	2884	Pneumo	3133	1	249
4	M	3	43	0	0	3780	Hemato	3964	1	184
5	F	3	32	0	1	2923	Candida	3588	1	665
5	F	3	32	0	1	2923	Hemato	5243	1	2320
5	F	3	32	0	1	2923	Pneumo	3987	1	1064
6	M	2	33	0	0	2938	Candida	4828	1	1890
6	M	2	33	0	0	2938	Pneumo	4639	1	1701

Eventos paralelos – aids

- o termo `cluster` é necessário para corrigir a variância da estimativa
- para preservar a não ordenação no tempo o modelo deve ser `Surv(tempo,status)`, nunca usando `Surv(inicio,fim,status)`
- quem estaria censurado?
- pode-se ou não estratificar, dependendo do risco de base para cada evento ser igual ou não:

$$\lambda_A(t) = \lambda_B(t)$$

$$\lambda_A(t) \neq \lambda_B(t)$$

Modelando separadamente

Pode-se ajustar um modelo para cada doença oportunista, estimando os diferentes riscos para cada uma delas:

```
> hem.fit <- coxph(Surv(tempo,status)~udi+sexual+sex+
  factor(esc),data=op,subset=(oport=="Hemato"))
> her.fit <- coxph(Surv(tempo,status)~udi+sexual+sex+
  factor(esc),data=oport,subset=(oport=="Herpes"))
```

```
:
```

Desvantagem: não é possível estimar o efeito de algum fator de risco comum a todas

Efeitos de covariáveis – eventos paralelos

Efeitos de covariáveis sobre o risco de desenvolver algumas doenças oportunistas

Covariáveis	hemo.fit	herp.fit	cand.fit	pneu.fit	tubc.fit
<i>udi</i>	0.8188	0.4534	0.9206	0.5870	2.5973*
<i>sexual</i>	0.6990*	0.7792	0.7776	0.6453	0.9357
<i>sexM</i>	1.2981*	0.8633	1.0657	1.1142	0.9339
<i>esc1</i>	0.8294	0.6974	0.9392	2.3267*	4.3446
<i>esc2</i>	0.8045	0.7854	0.9796	2.6867**	4.4126
<i>esc3</i>	0.7497	0.7441	0.8231	3.2377**	5.3651
<i>esc4</i>	0.5458	0.5512	0.8055	NA	1.2952

Efeito significativo ao nível de 10% (*), 5% (*), 1% (**)

Modelando conjuntamente

Verificar se a linha de base é diferente entre as doenças

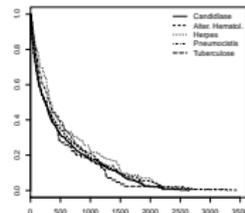


Figura: Curva de sobrevivência segundo doenças oportunistas

Modelo para eventos paralelos

```
> summary(coxph(Surv(tempo,status)~udi+sexual+
sex+factor(esc)+cluster(reg),data = op))

[...]
```

	exp(coef)	exp(-coef)	lower .95	upper .95
udi	0.7909	1.2643	0.5529	1.1314
sexual	0.7385	1.3540	0.5500	0.9917
sexM	1.1451	0.8733	0.9797	1.3384
factor(esc)1	1.0026	0.9974	0.4989	2.0148
factor(esc)2	1.0288	0.9720	0.5114	2.0699
factor(esc)3	0.9649	1.0364	0.4771	1.9516
factor(esc)4	0.6546	1.5276	0.3164	1.3545

```
Rsquare= 0.028 (max possible= 1 )
[...]
```

Eventos paralelos – comparando

- efeito médio de todas as covs não significativo, por que?
- diferentes linhas de base?
- modelo para cada doença separadamente não permite comparar riscos
- ajustar cada doença como um estrato

Eventos paralelos – diferentes estratos

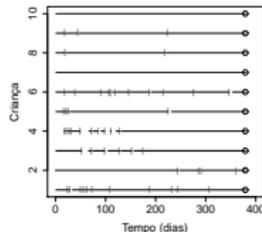
```
[...]
```

	exp(coef)	exp(-coef)	lower .95	upper .95
udi	0.7999	1.2501	0.5548	1.1533
sexual	0.7379	1.3553	0.5548	0.9813
sexM	1.1383	0.8785	0.9721	1.3329
factor(esc)1	1.1114	0.8997	0.5112	2.4166
factor(esc)2	1.1263	0.8878	0.5159	2.4592
factor(esc)3	1.0698	0.9348	0.4886	2.3421
factor(esc)4	0.7117	1.4050	0.3185	1.5906

```
Rsquare= 0.028 (max possible= 1 )
[...]
```

Eventos ordenados

Tempos observados entre episódios recorrentes de diarreia



Eventos ordenados – BD

```

numcri grupo sexo idade ini fim enum status
2 pla fem 48 1 244 1 1
2 pla fem 48 247 286 2 1
2 pla fem 48 289 291 3 1
2 pla fem 48 293 361 4 1
2 pla fem 48 363 379 5 0

5 pla fem 42 1 16 1 1
5 pla fem 42 18 21 2 1
5 pla fem 42 23 26 3 1
5 pla fem 42 29 225 4 1
5 pla fem 42 232 379 5 0

```

Duração do evento

- pontual: sem duração mensurável → o tempo do indivíduo sob risco é contínuo, sem interrupção
- IAM: o indivíduo pode ainda não ter se recuperado de um infarto, e já estar em risco de outro
- com duração: o evento tem início e fim → o tempo do indivíduo sob risco exclui o tempo de duração (diarréia)
- evitar a fragmentação do evento

Incrementos independentes – AG

- o indivíduo retorna ao estado inicial sempre com o mesmo risco $\lambda(t)$ de sofrer qualquer evento
- pressuposto de risco igual para qualquer evento assume que o histórico do indivíduo não afeta seu risco presente
- Ex: risco de engravidar, diarréias, IRAs.
- a formulação é a mesma dos eventos paralelos, sendo necessário incluir o termo para a correção da variância (`cluster()`), pois os tempos até a ocorrência de cada evento se superpõem
- não se usa `strata` – a ordenação é dada pelo `Surv(ini.fim,status)`

Incrementos independentes – diarréia

Call:

```

coxph(formula = Surv(ini, fim, status) ~ grupo+sexo+idade+
      cluster(numcri), data = diarrreia)
n= 5592

```

	coef	exp(coef)	se(coef)	robust se	z	p
grupopla	0.1621	1.176	0.02911	0.06394	2.536	0.011
sexomasc	0.0597	1.061	0.02914	0.06398	0.933	0.350
idade	-0.0339	0.967	0.00122	0.00256	-13.265	0.000

	exp(coef)	exp(-coef)	lower .95	upper .95
grupopla	1.176	0.850	1.037	1.333
sexomasc	1.061	0.942	0.936	1.203
idade	0.967	1.035	0.962	0.971

Rsquare= 0.135 (max possible= 1)

Para avaliar se este modelo é adequado, comparar `se(coef)` e `robust se`. Se forem muito diferentes é porque deve haver uma estrutura de correlação entre os tempos, logo não seriam incrementos independentes.

Modelo condicional ou PWP

- condicional apenas no nome, porque considera-se que o evento de ordem j somente pode acontecer se o de ordem $j - 1$ tiver acontecido
- não é condicional na forma de estimar – também é marginal

Modelo PWP – diarreia

Call:

```
coxph(formula = Surv(ini, fim, status) ~ grupo+sexo+idade+
      strata(enum)+cluster(numcri), data = diarreia)
```

```
n= 5592
      coef exp(coef) se(coef) robust se      z      p
grupola 0.0470   1.048  0.02991  0.03574  1.31 0.19
sexomasc 0.0529   1.054  0.02974  0.03571  1.48 0.14
idade   -0.0160   0.984  0.00132  0.00153 -10.48 0.00
```

```
exp(coef) exp(-coef) lower .95 upper .95
grupola   1.048   0.954  0.977   1.124
sexomasc  1.054   0.948  0.983   1.131
idade     0.984   1.016  0.981   0.987
```

```
Rsquare= 0.027 (max possible= 0.999 )
```

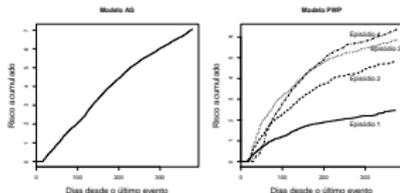
```
Likelihood ratio test= 155 on 3 df, p=0
```

```
Wald test = 112 on 3 df, p=0
```

```
Score (logrank) test = 154 on 3 df, p=0, Robust = 105 p=0
```

Modelo PWP – diarreia

Risco acumulado basal para ocorrência de novo episódio diarreico



Não existe modelo perfeito!

- Pensar bem sobre o modelo conceitual para o problema (estratificação, covariáveis tempo-dependentes, etc).
- Montar o banco de dados é trabalho e sujeito a erros, defina cuidadosamente qual o modelo que será ajustado.
- Sempre verificar no banco:
 - a estrutura temporal,
 - a variável indicativa de censura,
 - o tempo que indica término do estudo,
 - as mudanças de covariáveis no tempo,
 - a variável indicativa de estrato.
- Ajustar o modelo de Cox simples.
- Corrigir a estimativa de variância, calculando uma variância robusta, caso necessário.
- Introduzir a estratificação.

Ajuste e resíduos

Tudo igual

Resumo

Tabela: Características dos modelos de eventos múltiplos

Modelo	Desfecho	Risco de base
Eventos paralelos	Desfechos independentes, em número variável	Pode variar entre os desfechos
Incrementos independentes AG	Observações independentes para um indivíduo	Não varia de um evento para o outro
Estruturado PWP	Assume estrutura de dependência entre os eventos	Variável, cada evento pertence um estrato

Banco de dados para análise de eventos múltiplos

Modelo	Individuo	Início	Fim	Tempo	Status	Evento
Evento	1	1	10	9	1	A
Paralelo	1	1	12	11	1	B
com ou sem	1	1	16	15	1	C
estratificação	2	1	15	14	1	A
	2	1	15	14	1	C
	3	8	20	12	1	A
AG – tempos independentes	1	0	12	-	1	X
sem	1	12	33	-	1	X
estratificação	1	33	54	-	1	X
	2	3	20	-	1	X
	2	20	43	-	0	X
PWP – estruturado	1	0	12	-	1	1
estrato	1	12	33	-	1	2
ordenado	1	33	54	-	1	3
	2	3	20	-	1	1
	2	20	43	-	0	2

Síntaxe no R

- AG:


```
coxph(Surv(inicio,fim,status) ~ covariáveis+
cluster(id),data=banco
```
- PWP:


```
coxph(Surv(inicio,fim,status) ~ covariáveis+
cluster(id)+strata(enum),data=banco
```

Outline

- 1 Efeitos não-lineares
- 2 Eventos Múltiplos
- 3 **Fragilidade**

Objetivos

- Conceituar fragilidade ou efeito aleatório.
- Indicar adequadamente a inclusão de efeitos aleatórios na modelagem da sobrevivência.
- Identificar as distribuições mais utilizadas e as formas de estimação do efeito aleatório.
- Ajustar os modelos utilizando o R.
- Interpretar as estimativas obtidas no ajuste de modelo com efeitos aleatórios.

Algo não medido....

- Modelos de sobrevivência explicam quanto da variabilidade?
Superdispersão
- E se o serviço de saúde for o mesmo?
Multinível

Efeitos Aleatórios

- Efeitos aleatórios atuando multiplicativamente sobre o risco de base, da mesma forma que as covariáveis fixas.
- Se o indivíduo é um *cluster* de observações, como nos eventos múltiplos, não pode ser tratado como efeito aleatório.
- É possível incluir dois níveis de efeito aleatórios, por exemplo, o do indivíduo e o da unidade, mas a estimação fica complicada.

Modelo de fragilidade

- Z – variável aleatória desconhecida que reflete a fragilidade (peculiaridade) do indivíduo ou do grupo
- média igual a 1 e variância ξ
- ξ grande \rightarrow variabilidade não atribuível a covariável fixa alta
- ξ pequeno \rightarrow pouca heterogeneidade, não é necessário incluir fragilidade

Modelo de fragilidade

$$\lambda(t) = z\lambda_0(t)\exp(\mathbf{x}\beta)$$

sendo

$\mathbf{x} \rightarrow$ covariáveis fixas

$Z = z \rightarrow$ efeito aleatório

- Fragilidade $z_i > 1 \rightarrow$ indivíduo (ou grupo) i tendem a experimentar o evento com uma taxa mais rápida que sob o modelo de Cox básico.
- Variância $\xi \Rightarrow 0$ modelo de Cox.

Quem é Z – Variável aleatória com distribuição "sensata"

Distribuição da fragilidade precisa ter plausibilidade biológica:

- **gama** – a distribuição da fragilidade entre os que sofreram o evento em um dado tempo e entre os sobreviventes também é gama (resultado teórico demonstrável)
- **lognormal** – escolha quando fragilidades relacionadas a covariáveis perdidas x_m : dado $Z = \exp(x_m\beta_m)$, ao aplicarmos o teorema do limite central a $x_m\beta_m$, teremos uma distribuição lognormal para Z .
- Ambas distribuições muito flexíveis, caracterizadas por dois parâmetros, variam desde forma exponencial até forma de sino similar à curva normal

Quem é Z – gama

- $\gamma \rightarrow$ parâmetro de forma
- $\alpha \rightarrow$ parâmetro de escala

$$E(Z) = \gamma\alpha$$

$$Var(Z) = \gamma\alpha^2$$

Sendo

$$\gamma = 1/\xi$$

$$\alpha = \xi$$

Então

$$E(Z) = 1$$

$$Var(Z) = \xi$$

Quem é Z - *gama*

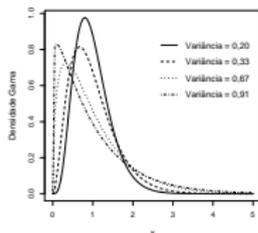


Figura: Distribuição gama com valor esperado 1 e diferentes variâncias ξ

Quem é Z - *lognormal*

- μ -> parâmetro de locação
- σ^2 -> parâmetro de escala

$$E(Z) = \exp(\mu + \sigma^2/2)$$

$$Var(Z) = \exp(\sigma^2) - 1$$

Se

$$\mu = -\sigma^2/2e$$

$$\sigma^2 = \log(1 + \xi)$$

Então

$$E(Z) = 1$$

$$Var(Z) = \xi$$

Quem é Z - *lognormal*

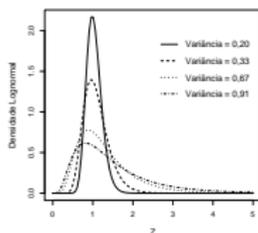


Figura: Distribuição lognormal com valor esperado 1 e diferentes variâncias ξ

Quem é Z

- *lognormal* tem a cauda direita mais pesada do que a *gama* => proporção maior de indivíduos com riscos muito altos
- cauda esquerda é mais leve que *gama* => menor a proporção de indivíduos com riscos muito baixos

Algoritmo EM

- Utiliza-se algoritmo Esperança-Maximização (EM) ou inferência bayesiana.
- Fragilidade tratada como dados não observáveis => estimadas no passo **E** do algoritmo
- $E(Z|t, \delta)$ como estimador.
- No passo **M** são obtidos os valores dos coeficientes de regressão que maximizam a verossimilhança parcial:

$$L_i(\beta) = \left(\frac{z_i \exp(\mathbf{x}_i \beta)}{\sum_{j \in R(t_i)} z_j \exp(\mathbf{x}_j \beta)} \right)^{\delta_i}$$

Algoritmo EM

- Bom para estimar fragilidade individual
- *ipdmeta* (*Individual Patient Data Meta-Analysis of Time-to-Events*)
- *phmm* (*Proportional Hazards Mixed-effects Model*) – algoritmo EM utilizando MCMC

```
modelo.phmm <- phmm(Surv(tempo,status) covariáveis + (1|id),
data=dados)
```

Verossimilhança Parcial Penalizada

- Implementado no R
- Somente para fragilidade compartilhada (ou multinível)

$$\lambda(t) = \lambda_0(t) \exp(\mathbf{x}\beta + \mathbf{U}\omega) \quad (1)$$

- \mathbf{x} e \mathbf{U} são matrizes de covariáveis
- β correspondem aos p efeitos fixos
- ω é um vetor que contém q efeitos aleatórios
- o elemento u_{ij} da matriz \mathbf{U} é igual a 1 se o indivíduo i é um membro do grupo j , e é igual a zero caso contrário

Estimação

- Maximização da função de logverossimilhança parcial penalizada em função de β e ω .

$$LPP = l(\beta, \omega) - g(\omega, \theta) \quad (2)$$

- g é a função que impõe penalidades a valores de ω menos desejáveis.
- θ é um parâmetro de ajustamento de g
- Quando:
 - $g(\omega, \theta) = (1/\theta) \sum_i [\omega_i - \exp(\omega_i)]$ => modelo gama de efeitos aleatórios
 - $g(\omega, \theta) = (1/2\theta) \sum_i (\omega_i^2)$ => modelo lognormal, sendo θ é a variância dos efeitos ω_i .

Ajuste no R

- θ pode ser fixado diretamente
- pode-se especificar grau de liberdade para o termo dos efeitos aleatórios de 0 (máxima penalidade) a q (nenhuma penalidade, sendo q o número máximo de grupos)
- critério de informação de Akaike (AIC)
- a variância fragilidade gama pode ser estimada usando-se verossimilhança perfilhada, com resultado idêntico ao do algoritmo EM
- fragilidades lognormais pode ser estimada com base numa equação de verossimilhança restrita aproximada

Testes de Hipóteses

- os mesmos: Wald e teste da Razão de Verossimilhança
- $H_0: v = C(\beta', \omega')' = 0$
- $(\beta', \omega')'$ é o vetor combinado de $p + q$ parâmetros, e C é uma matriz de contrastes de dimensão $k \times (p + q)$
- teste de Wald:

$$W = v'(CH^{-1}C')^{-1}v \quad (3)$$

em que $H = I + G$ é a matriz de derivadas segundas da função de logverossimilhança penalizada com I sendo a matriz de informação usual do modelo de Cox e

$$G = \begin{pmatrix} 0 & 0 \\ 0 & -g'' \end{pmatrix} \quad (4)$$

- sob H_0 , W tem distribuição assintótica qui-quadrado com gl graus de liberdade.

Fragilidade no R

- Tudo igual, incluindo o termo `frailty`

```
modelo <- coxph(Surv(inicio,fim,status) ~ covariáveis + frailty(grupo),
data=dados)
```

- `grupo` => variável que indica o segundo nível, para o qual a fragilidade será estimada

Clínicas de hemodiálise

Objetivo Avaliar qualidade dos centros de hemodiálise

Dados APAC, 1998 a 2001 (mensal), 44 meses

Estado do Rio

Variáveis bancos PAC, AC, LAB, CENTRO...

Indivíduo Idade e sexo

Causa da Doença Renal: Hipertensão (base), diabetes, congênitas, renal, outras

Centro No de máquinas: 11 a 19 Vs 20 e mais

Número de pacientes: até 50 e maior

Proporção de pacientes idosos, oferta de diálise cíclica, ...

Problemas Dado prevalente

Estrutura de dependência por centro

Mudança de centro

Aviso

- Warning message: Loglik converged before variable 23 ; beta may be infinite. in: `fitter(X, Y, strats, offset, init, control, weights = weights,`
- Warning message: Inner loop failed to converge for iterations 6 in: `coxpenal.fit(X, Y, strats, offset, init = init, control, weights = weights,`

Comparando os modelos

- H_0 : efeitos aleatórios = 0;
- Estatística de Wald segue distribuição chi-quadrado (frailty(unidade) = 123.10);
- Razão de verossimilhança segue distribuição chi-quadrado;
- O efeito aleatório tem 1 grau de liberdade

$$2 \times \loglik(modelo.frail) - \loglik(modelo.nofrail)$$

$$2 \times (-1059,8 - (-1099,963)) = 80,33$$

- Medida global de ajuste:
Likelihood ratio test= 210 on 20.4 df, p=0
Não disponível para modelos Gauss

Estimando os modelos

- distribuição gama => verossimilhança perfilhada para θ (`method=EM`)
- distribuição normal => máxima verossimilhança restrita aproximada (`method=REML`)
- minimização AIC (Critério de Akaike)
- Não há método para identificar se uma distribuição é melhor que a outra, ou um método de estimação é melhor que o outro
- Avaliar consistência:
 - mesmas unidades com risco diferenciado
 - estimativas efeitos fixos e intervalos de confiança semelhantes

Comparando os modelos

	Efeito Gama (EM)			Efeito Gama (AIC)		
	$exp(\beta)$.95Inf	.95Sup	$exp(\beta)$.95Inf	.95Sup
idade	1.033	1.022	1.04	1.033	1.022	1.04
causacon	0.393	0.144	1.07	0.388	0.144	1.04
causadia	1.819	1.270	2.60	1.827	1.277	2.61
causaout	1.182	0.659	2.12	1.174	0.656	2.10
causaren	1.279	0.846	1.93	1.293	0.856	1.95

	Efeito Gauss (REML)			Efeito Gauss (AIC)		
	$exp(\beta)$.95Inf	.95Sup	$exp(\beta)$.95Inf	.95Sup
idade	1.033	1.022	1.04	1.033	1.022	1.04
causacon	0.419	0.157	1.12	0.415	0.157	1.09
causadia	1.814	1.267	2.60	1.820	1.273	2.60
causaout	1.168	0.660	2.07	1.161	0.658	2.05
causaren	1.247	0.824	1.89	1.254	0.829	1.90

Comparando os ajustes

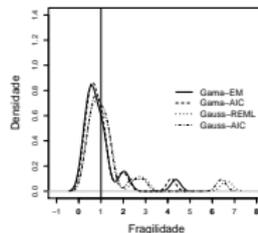
Medidas de ajuste	Cox Clássico	Gama EM	Gama AIC	Gauss EM	Gauss AIC
Variância do efeito aleatório	—	0,734	0,55	0,697	0,55
Wald Test ^{a,c}	—	123,10 (15,6)	118,81 (14,8)	128,26 (14,7)	123,82 (14,1)
Rsquare	0,085	0,216	0,215	0,216	0,215
Log-verossimil.	-1100	-1033	-1034	-1033	-1034
I-likelihood	—	-1059,8	-1059,9	—	—
Qualidade do Ajuste ^{b,c}	76,5 (5)	210 (20,4)	208 (19,5)	210 (19,5)	209 (18,8)

^a Teste para a fragilidade (graus de liberdade entre parênteses)

^b Teste da razão de verossimilhança (graus de liberdade entre parênteses)

^c p-valor <0,001

Comparando os efeitos



Observar o escopo da fragilidade – gama chega perto de zero, gauss com valores muito altos e flutuação

Comparando os efeitos

