

MAE514 - Introdução a Análise de Sobrevivência

1º Semestre/2020

5ª Lista de Exercícios

Informações Importantes

- Data de entrega: 05/08 (Quarta-feira)
- Forma de entrega: pelo sistema e-disciplinas exclusivamente.
- Acrescente o código do programa em todas as questões quando um *software* for utilizado.

1. Considere o conjunto de dados de pacientes com leucemia aguda que receberam transplante de medula óssea, apresentado na seção 1.3 de Klein e Moeschberger (2003). Os pacientes que recebem transplante de doador compatível (alogênico) podem desenvolver uma doença conhecida como DECH (*doença do enxerto contra o hospedeiro*) ou GVHD (*graft-versus-host disease*), que pode ser muito grave. No entanto, suspeita-se que DECH tenha um efeito anti-leucêmico nos pacientes. Para verificar essa hipótese, deseja-se ajustar um modelo de Cox aos dados, considerando-se como variável resposta o tempo até a recorrência da doença (*relapse*). Pacientes que apresentaram óbito antes da recorrência são considerados como observações censuradas. Os dados que devem ser utilizados estão disponíveis no arquivo `BMT-Data.csv` e a descrição das variáveis está no arquivo `BMT-Data.des`.
 - (a) Defina duas variáveis dependentes do tempo indicadoras da ocorrência de DECH aguda e crônica. Escreva o modelo semiparamétrico de Cox incluindo as variáveis idade do paciente, sexo do paciente, tempo até o transplante, classificação morfológica Franco-Americana-Britânica (FAB) e tratamento profilático para DECH (MTX), além das variáveis dependentes do tempo definidas.
 - (b) Ajuste um modelo de Cox com as variáveis do item (a). Selecione as variáveis significativas e faça a verificação da proporcionalidade dos riscos. O que pode ser concluído?
 - (c) Repita os itens (a) e (b) utilizando o tempo até o óbito como variável reposta.

2. Um estudo foi conduzido para estudar se uma determinada droga era cancerígena ou não. Para isso, foram selecionadas 50 ninhadas de ratos e, de cada ninhada, foram selecionados 3 ratos. Dos ratos de cada ninhada, selecionou-se aleatoriamente um deles para receber a droga e os demais receberam placebo. Observou-se o tempo, em semanas, até o desenvolvimento de um tumor. Os dados estão apresentados no arquivo `litter-data.csv`.
 - (a) Obtenha as curvas de Kaplan-Meier para os dois grupos e comente.
 - (b) Ajuste um modelo de riscos proporcionais de Cox, ignorando o fato de existirem ratos pertencentes a uma mesma ninhada. Obtenha as estimativas dos parâmetros e interprete os resultados. Obtenha os resíduos de Schoenfeld para testar a proporcionalidade dos riscos. Faça um teste estatístico para a proporcionalidade dos riscos, utilizando a transformação KM, ou seja, a curva de sobrevivência de Kaplan-Meier (versão contínua à esquerda).
 - (c) Ajuste um modelo estratificado por ninhada, obtenha as estimativas e interprete os parâmetros. Obtenha os resíduos de Schoenfeld e faça o mesmo teste do item anterior para testar a proporcionalidade dos riscos. Compare os resultados.
3. Um estudo foi feito sobre o efeito da radiação na sobrevivência de ratos. Um grupo de ratos recebeu uma dose de 300 rad de radiação quando tinham entre 5 e 6 semanas de vida e foram acompanhados até o óbito. Quando morriam, os ratos eram necropsiados e a causa da morte determinada. Em particular, os pesquisadores tinham interesse em estudar as mortes por um tipo específico de linfoma um tipo de sarcoma.

Os tempos de vida, em dias, dos ratos e as causas das mortes estão apresentadas na tabela abaixo.

Causa da morte	Idade ao morrer (dias)
Linfoma tímico	158, 192, 193, 194, 195, 202, 212, 215, 229, 230, 237, 240, 244, 247, 259, 300, 301, 337, 415, 444, 485, 496, 529, 537, 624, 707, 800
Sarcoma de células reticulares	430, 590, 606, 638, 655, 679, 691, 693, 696, 747, 752, 760, 778, 821, 986
Outras causas	136, 246, 255, 376, 421, 565, 616, 617, 652, 655, 658, 660, 662, 675, 681, 734, 736, 737, 757, 769, 777, 801, 807, 825, 855, 857, 864, 868, 870, 873, 882, 895, 910, 934, 942, 1015, 1019

- (a) Obtenha as curvas de incidência acumulada dos três riscos competitivos. Comente.
- (b) Obtenha os valores estimados das funções de incidência acumulada dos três riscos competitivos nos instantes $t = 200, 300, 400, 600, 800$ e 1000 .
- (c) Obtenha a curva de Kaplan Meier da sobrevivência global dos dados. Calcule o valor da função de sobrevivência nos instantes $t = 200, 300, 400, 600, 800$ e 1000 . Observe que a soma dos valores das funções de incidência acumulada em cada instante é igual a 1 menos o valor da curva de Kaplan-Meier naquele ponto.
- (d) Obtenha as curvas de sobrevivência de Kaplan Meier marginal dos dados, para cada tipo de evento, considerando a ocorrência dos outros eventos como censuras à direita. Coloque em um mesmo gráfico as curvas de incidência acumulada e 1 menos a curva de Kaplan-Meier. Compare e comente as diferenças, enfatizando o que cada curva de fato estima.