

Proposta 1: Função SOBREVIVE

Introdução

Análise de sobrevivência é um ramo da estatística que analisa o tempo até a ocorrência de determinado fenômeno, como morte ou diagnóstico de doença. Análises de sobrevivência obedecem modelos de regressão logística e têm um ponto de partida - geralmente definido quando as observações se iniciam - e um término - quando o evento de interesse acontece. O período entre o início das observações e a ocorrência do evento de interesse compreende no tempo de observação, que é a variável independente (preditora) do modelo. Contudo, o tempo para a ocorrência de um dado evento pode ser maior que o tempo disponível para observação e coleta de dados, gerando observações incompletas. Processos produzindo esse tipo de observação incompleta é chamado de “*censoring*”, ao passo que a observação incompleta em si é chamada “*censored data*”.

“*Censored data*” são comuns de acontecer em diversos tipos de trabalho devido à limitação de recursos para a realização do estudo, como tempo, equipe e financiamento. Desse modo, análises de sobrevivência devem utilizar modelos que estimam a probabilidade de sobrevivência de um determinado grupo considerando “*censored data*”. A estimativa padrão para funções de sobrevivência é a de Kaplan-Meier (Kaplan & Meier, 1958), a qual utiliza os dados observados para estimar a probabilidade de sobrevivência em cada grupo analisado e para extrair os parâmetros de interesse do pesquisador.

Sobre a função

Descrição

Baseada na introdução acima, proponho uma função que implemente a estimativa de Kaplan-Meier para curvas de sobrevivência provenientes de diferentes conjuntos de dados (com diferentes níveis de tratamentos, diferentes populações ou espécies), considerando “*censored data*”.

Inputs

- *Dataframe* com a descrição do grupo estudo (espécie, população, tratamento, réplica, etc), os dias de observação e a quantidade de indivíduos vivos para cada um dos dias.

Argumentos

- Omissão de dados faltantes (“TRUE”/“FALSE”).
- Número de grupos de estudo (mínimo de 1).
- Número de dias de observação.

Outputs

- Gráfico com a curva de sobrevivência para cada grupo de estudo.

- Parâmetro meia-vida da curva de sobrevivência.

Etapas

1. Cálculo da estimativa Kaplan-Meier para cada grupo de estudo do conjunto de dados disponível*.
2. Construção do gráfico com a curva de sobrevivência para cada grupo de estudo.
3. Extração do parâmetro meia-vida da curva de sobrevivência (tempo t em que $S(t) = 50$).

*O cálculo da estimativa será feito seguindo a fórmula:



Exemplo de uso

A função poderá ser usada por estudiosos de diversas áreas do conhecimento, da área médica até estudos ecológicos ou de fisiologia. Particularmente, a função SOBREVIVE será de grande utilidade para o meu projeto de mestrado, pois terei que construir e analisar curvas de sobrevivência de 48 espécies de *Drosophila* sob dois tipos de tratamento (infectadas e não infectadas com vírus DAV).

Referências

Hosmer, DW; Lemeshow, S; May, S. Applied survival analysis : regression modeling of time-to-event data - 2nd ed. 411p.

Ótima proposta que captura bem o espírito desse trabalho.

Eu gostaria de ver uma descrição mais precisa do dados de entrada e do que exatamente a função vai calcular. A ideia é estimar um numero de eventos por tempo? É isso? É bom começar a função sabendo exatamente oq vc gostaria que saísse.

Seria legal ter também comparações estatísticas entre os grupos, as vezes que eu usei curvas de sobrevivência foi pra diferenciar as curvas entre grupos. Alguma ideia de como fazer isso?

resolvidos esses detalhes pode mandar bala!

Valeu,

—Ogro

From:
<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2016:alunos:trabalho_final:camila.beraldo:proposta_a 

Last update: **2020/08/12 06:04**