

ESTUDO DAS PROPRIEDADES DOS ESTIMADORES DOS MODELOS DE SOBREVIVÊNCIA ESTENDIDOS E SUAS APLICAÇÕES.

Leonardo Romaioli Barros (PIBIC/CNPq/FA/Uem), Carlos Aparecido dos Santos (Orientador), e-mail: casantos@uem.br, Daniele Cristina Tlta Granzotto (Coorientador), e-mail: dctgranzotto@uem.br

Universidade Estadual de Maringá / Centro de Ciências Exatas e da Terra/Maringá, PR.

Ciências Exatas e da Terra/Probabilidade e Estatística

Palavras-chave: Sobrevivência, risco, modelagem.

Resumo:

Na última década, pesquisas relacionadas com a proposição, derivação ou extensão de novos modelos probabilísticos para o estudo do tempo de sobrevivência cresceram exponencialmente. O grande desafio dos pesquisadores envolvidos nestas pesquisas é o de derivar modelos de probabilidade usuais, amplamente usados na modelagem dos tempos até a ocorrência de um evento de interesse, para que estes representem de forma mais consistente a ampla gama de comportamento das curvas de riscos e sobrevivência dos tempos observados. Assim, o modelo resultante desta extensão ou derivação deve ser simples (com o menor número de parâmetros), interpretáveis e que sejam mais flexíveis e tratáveis que a distribuição de base (aquela usada para a composição da nova distribuição/família, que se propõem). No contexto de modelagem de dados de sobrevivência e proposição de suas extensões, neste projeto de iniciação científica foi estudado a metodologia de análise de sobrevivência e os principais modelos probabilísticos usados para a modelagem dos tempos de sobrevida, além dos novos modelos estendidos Fisher ortogonais. Aplicações com dados reais fizeram parte deste trabalho.

Introdução

A análise estatística do tempo até o evento de interesse em estudo é uma ferramenta chave em pesquisas nas áreas de engenharia e saúde (biomédicas). Preocupados com esta particular área da estatística, uma ampla gama de modelos estatísticos vem sendo desenvolvidos e propostos na literatura. Em particular, muitos são os autores que buscam, a partir de modelos mais simples e amplamente aplicados na área, a derivação ou construção de novas distribuições.

Nos últimos anos, dezenas de modelos transmutados foram apresentados na literatura, usando por base a composição funcional da função de distribuição acumulada e da função de distribuição acumulada inversa (quantil) de um outro modelo, proposta por Shaw e Buckley (2009). Dentre outros artigos publicados nesta área, citamos as transmutações propostas por Louzada e Granzotto (2015).

Apesar de amplamente usada na construção/derivação de novas distribuições, esta metodologia apresenta algumas limitações quanto ao espaço paramétrico do parâmetro a ser incluído ($-1 < \lambda < 1$) e também problemas de identificabilidade e tratabilidade em casos muito particulares. Diante de problemas apresentados na classe transmutada Granzotto *et. Al* (2017) propuseram uma nova família de distribuições, os modelos de sobrevivência e estendidos (ou exponencial estendidos). Apesar de mais flexível que o modelo transmutado, o modelo supracitado não consegue resolver o grande desafio que é o de representar de forma mais consistente a ampla gama de comportamento das curvas de riscos e sobrevivência dos tempos observados pois continua com limitações semelhantes no espaço paramétrico do parâmetro a ser incluído.

Na busca de modelos flexíveis e com boas propriedades desejáveis (simplicidade, tratabilidade e interpretabilidade), neste projeto de iniciação científica foi estudado as extensões Fisher ortogonais, Casella e Berger (2002).

Assim, neste projeto de iniciação científica o objetivo foi estudar e aplicar as metodologias de inferência clássica, análise de sobrevivência e de modelagem de dados, em particular as extensões de modelos. Toda a metodologia estudada foi aplicada em dados reais.

Materiais e métodos

Neste projeto, para desenvolver e atingir os objetivos em sua plenitude, inicialmente foi realizado um levantamento bibliográfico contemplando as áreas de inferência, sobrevivência e modelagem estatística, bem como alguns dos principais modelos de probabilidade usados no estudo do tempo. Paralelamente a este levantamento, a caracterização dos modelos e suas extensões, que contemplam as propriedades desejadas e já descritas, foi realizada por meio das funções de risco e de sobrevivência, momentos e outras características de interesse.

Uma vez caracterizados tais modelos, a estimação dos parâmetros, aplicações e simulações necessárias, foram investigadas por meio do *software* R utilizando-se de bibliotecas.

Resultados e Discussão

Realizamos transformações das funções de probabilidade das famílias Gompertz e Log-logística para as respectivas funções de sobrevivência e risco e uma série de plotagens destas utilizando o *software* R, a fim de

analisar suas respectivas curvas, utilizando valores de θ fixo e variando o parâmetro λ . Análises gráficas das funções de densidade, sobrevivência e risco das composições proposta por Granzotto et al. (2019) foram realizadas. Um exemplo da função de sobrevivência no modelo Gompertz é ilustrado na Figura 1.

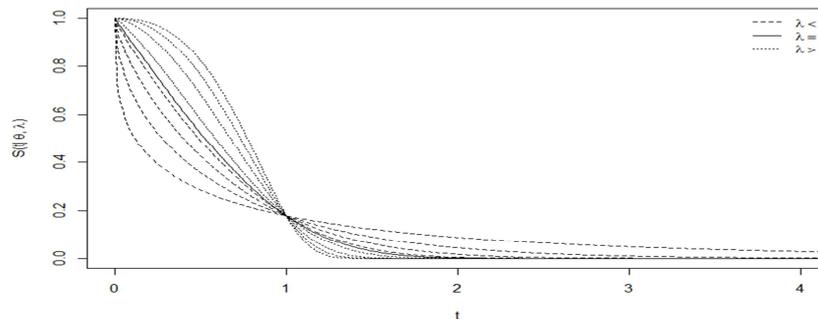


Figura 1. O efeito do parâmetro λ , na função de sobrevivência, com θ fixo em 1.

Na Figura 1 podemos visualizar as diversas formas que a função de sobrevivência do modelo Gompertz para diferentes valores do parâmetro λ . Auxiliando um projeto de mestrado na mesma área desta iniciação científica, realizamos a análise dos dados coletados de ensaio com mancais usados em esteira de transporte de materiais na indústria de grãos. Para Tal 8 experimentos com diferentes cargas e velocidades de rotação foram realizados como segue:

- E1 e E2 receberam cargas de 50 N só que em rotações distintas, 0,3 m/s e 0,6 m/s respectivamente;
- E3 e E4, foram utilizadas as mesmas condições tanto de carga quanto de rotação, 50 N a 0,6 m/s;
- E5 e E6 receberam a mesma rotação de 0,3 m/s mas com cargas distintas, 75 N e 100 N respectivamente;
- E7 e E8 recebeu a mesma carga de 100 N e rotação de 0,6 m/s e 0,9 m/s respectivamente.

Neste experimento buscamos determinar o tempo ótimo para a manutenção das esteiras, a fim de evitar a quebra e consequente parada no processo. Para tal, foram medidos os coeficientes de atrito ao longo de 120 minutos de experimentação. Os dados supracitados foram normalizados e técnicas de sobrevivência empregadas tanto na modelagem quanto na descrição dos resultados. Na Figura 2 apresentamos os coeficientes ao longo do tempo para cada um dos experimentos.

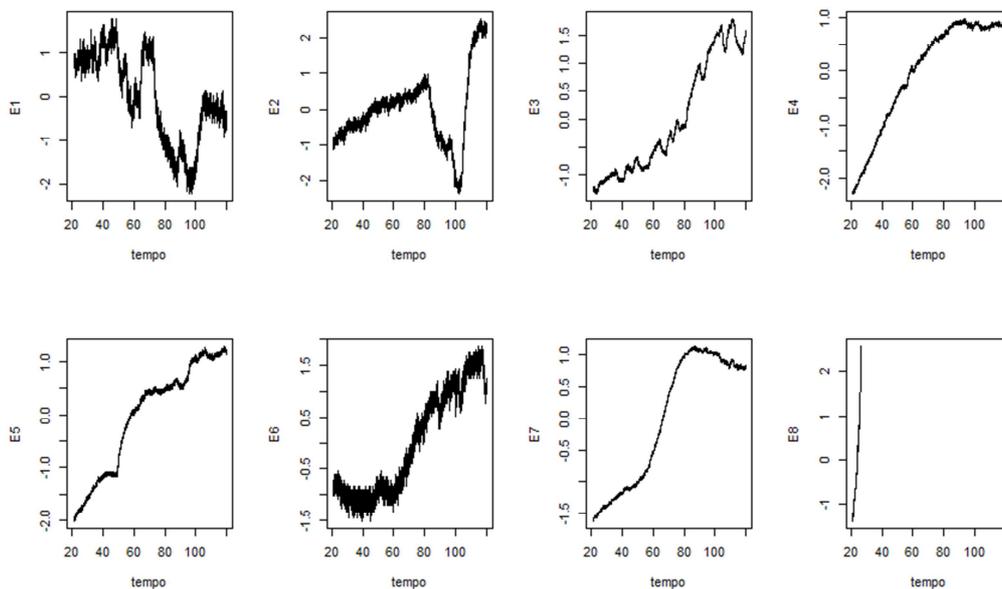


Figura 2. Coeficiente de atrito ao longo do tempo para os experimentos de E1 a E8.

Conclusões

Por se tratar de um estudo preliminar, uma vez que outros experimentos ainda precisavam ser executados, os dados iniciais foram analisados afim de direcionar os demais experimentos na busca do tempo ótimo de manutenção. Para tal análise, foram revistas as principais propriedades dos estimadores e métodos de estimação (neste caso focamos no método de máxima verossimilhança). Estudamos as técnicas de análise de dados de sobrevivência para aplicá-las aos dados fornecidos e esperamos dar prosseguimento a estas análises assim que os novos experimentos forem delineados e realizados.

Agradecimentos

A CAPES/CNPq, Fundação Araucária e UEM.

Referências

Casella, G.; Berger, R. L. Statistical inference. **Pacific Grove**, v. 2, 2002.

Louzada, F.; Granzotto, D. C. T. The transmuted log-logistic regression model: a new model for time up to first calving of cows. **Statistical Papers**, v. 57.3, p. 623-640, 2015.