

O Estimador Kaplan-Meier Aplicado à Inferência Causal com o Uso de Escores de Propensão

ARTHUR RIOS DE AZEVEDO * MARCELO TADDEO LEILA AMORIM
Universidade Federal da Bahia

Abstract

A análise de impacto tem despertado interesse crescente entre pesquisadores nos campos sociológico, econométrico e epidemiológico. Estudos que utilizam dados observacionais, especialmente aqueles que consideram o tempo até o evento no modelo de respostas potenciais de Rubin, têm impulsionado o desenvolvimento de métodos para identificar o efeito causal em dados de sobrevivência. Apesar de se basearem frequentemente em técnicas clássicas, como o modelo de Cox, a estimativa direta da intervenção pode apresentar viés devido às características inerentes dos dados e do modelo. Neste trabalho, discutiremos várias técnicas, incluindo o estimador de Kaplan-Meier como referência, o modelo de Cox como G-fórmula, o balanceamento de unidades por meio do escore de propensão, modelos estruturais marginais e estimação com restrições usando verossimilhança empírica. Por meio de estudos de simulação, avaliaremos a estimativa do efeito médio e a curva de sobrevivência em geral, além de explorar as vantagens das pseudo-observações nessas simulações.

References

- [1] IMBENS, G.; RUBIN, D. **Causal inference in statistics, social, and biomedical sciences**. [S.l.]: Cambridge University Press, 2015
- [2] BANG, H.; ROBINS, J. M. (2005). **Doubly robust estimation in missing data and causal inference models**. *Biometrics*, Wiley Online Library, v. 61, n. 4, p. 962–973.
- [3] XIE, J.; LIU, C. (2005) **Adjusted kaplan–meier estimator and log-rank test with inverse probability of treatment weighting for survival data**. *Statistics in medicine*, Wiley Online Library, v. 24, n. 20, p. 3089–3110.
- [4] ZHOU, M. (2015) **Empirical likelihood method in survival analysis**. [S.l.]: CRC Press, v. 79.

Tipo de Apresentação: PÔSTER

*e-mail: arthur.rios@ufba.br

Detecção de Outliers no Modelo de Regressão Circular

CARLOS ANDRÉ PEREIRA DE JESUS SILVA *
Universidade Federal da Bahia

Resumo

Nas últimas décadas, diversos estudos tem sido elaborados na área da Estatística que trata da modelagem usada para explicar a relação entre uma covariável circular e uma variável de resposta circular, chamada de regressão circular. Dentre esses estudos, têm sido propostos na literatura métodos para detecção de outliers em modelos de regressão circular, dado o impacto dessas observações nas estimativas dos parâmetros do modelo. Neste trabalho, consideram-se a estatística COVRATIO e a estatística erro médio circular para detecção de outliers no contexto do modelo de regressão circular desenvolvido por [1]. Por meio de estudos de simulação, os pontos de corte para os procedimentos são obtidos, e seu desempenho é investigado. Um exemplo da aplicação dos procedimentos é apresentado usando um conjunto de dados real da direção do vento.

Referências

- [1] Jammalamadaka, S. R., Sarma, Y. R. (1993). Circular regression. In: Matsusita, K., ed. Statistical Science and Data Analysis. Utrecht: VSP, pp. 109–128.
- [2] IBRAHIM A. RAMBLI, A. G. H. S.; MOHAMED, I. Outlier detection in a circular regression model using covratio statistic. Communications in Statistics - Simulation and Computation, Taylor Francis, v. 42, n. 10, p. 2272–2280, 2013. Disponível em: <https://doi.org/10.1080/00949655.2011.602679>.
- [3] Abuzaid, Ali & Hussin, Abdul & Mohamed, Ibrahim. (2008). Identifying single outlier in linear circular regression model based on circular distance. Journal of Applied Probability and Statistics. 3. 107-117.
- [4] A. H. Abuzaid , A. G. Hussin I. B. Mohamed (2013). Detection of outliers in simple circular regression models using the mean circular error statistic, Journal of Statistical Computation and Simulation, 83:2, 269-277, DOI: 10.1080/00949655.2011.602679

Tipo de Apresentação: PÔSTER.

*e-mail: cpibic@gmail.com

Processo de Wiener para Modelagem da Degradação em Múltiplos Sistemas Reparáveis com Diferentes Efeitos de Manutenção

GEORGE ANDERSON ALVES DOS SANTOS *
UNIVERSIDADE FEDERAL DA BAHIA - UFBA

Resumo

A confiabilidade e a manutenção tornaram-se cruciais em sistemas industriais, levando ao desenvolvimento de teorias e metodologias associadas [3]. Em um mercado globalizado e altamente competitivo, produzir produtos com alta confiabilidade é essencial para maximizar lucros e atender à demanda dos consumidores. A análise de confiabilidade tradicional muitas vezes se baseia em dados de falhas para selecionar modelos de tempo de vida, mas os recentes avanços em técnicas de monitoramento e a alta confiabilidade dos produtos deslocaram o foco para modelos de degradação, que podem fornecer informações valiosas mesmo na ausência de falhas.

A análise de confiabilidade baseada em degradação postula que o monitoramento da degradação de características de qualidade (PCs) ao longo do tempo pode revelar informações importantes sobre a confiabilidade do equipamento, mesmo sem falhas observadas. Vários processos estocásticos, como o Processo Gama e o Processo de Wiener, têm sido utilizados para modelar a degradação do sistema. Esses modelos foram generalizados para incorporar covariáveis, efeitos aleatórios e impactos de manutenção. Estudos recentes, como os de Leroy et al.[2], investigam os efeitos de diferentes esquemas de observação sobre a qualidade da inferência estatística em modelos de degradação, especialmente no contexto de sistemas reparáveis com manutenção imperfeita.

O presente trabalho propõe um modelo de degradação para descrever múltiplos sistemas submetidos a ações de manutenção com diferentes efeitos, assumindo que a degradação segue um processo de Wiener. O modelo incorpora a influência das manutenções por meio do conceito de Redução Aritmética da Degradação com Memória 1 (ARD1), onde a manutenção impacta diretamente o nível de degradação do sistema. Esse processo permite que a degradação seja mensurada em três momentos-chave: imediatamente antes, imediatamente após e entre manutenções, o que possibilita um acompanhamento preciso da condição do sistema ao longo do tempo. A proposta é uma alternativa às tradicionais análises de confiabilidade que se baseiam exclusivamente em dados de falha, sendo útil especialmente em cenários onde ocorrem censuras, ou seja,

*e-mail: george_13031995@hotmail.com

onde não há falhas observadas no período de teste. A abordagem de degradação permite o uso de dados sobre a evolução de características do sistema que indicam sua deterioração ao longo do tempo, como trincas ou desgaste [1]. A simulação realizada no estudo mostra que os estimadores obtidos por máxima verossimilhança têm boas propriedades assintóticas, validando a eficácia do modelo. Essa abordagem proporciona uma alternativa para a análise de confiabilidade em sistemas de maior complexidade, especialmente em ambientes onde as falhas são raras ou onde há censuras significativas nos testes de vida.

Referências

- [1] L. H. M. Morita, “Degradation modeling for reliability analysis with time-dependent structure based on the inverse gaussian distribution”.
- [2] M. Leroy, C. Bérenguer, L. Doyen, and O. Gaudoin, ‘Statistical inference for a Wiener-based degradation model with imperfect maintenance actions under different observation schemes’, *Appl. Stoch. Models Bus. Ind.*, vol. 39, no. 3, pp. 352–371, May 2023.
- [3] X. Wang, O. Gaudoin, L. Doyen, C. Bérenguer, and M. Xie, ‘Modeling multivariate degradation processes with time-variant covariates and imperfect maintenance effects’, *Appl. Stoch. Models Bus. Ind.*, vol. 37, no. 3, pp. 592–611, May 2021.

Tipo de Apresentação: PÔSTER

Aplicando a Teoria da Resposta ao Item para Respostas Distais em dados do Enade de 2019

INGO DUBE SOUZA *
UNIVERSIDADE FEDERAL DA BAHIA

Resumo

A Teoria da Resposta ao Item (TRI) é uma metodologia amplamente utilizada para trabalhar com variáveis latentes, tendo como seu maior uso, a área educacional. Segundo Andrade, et al. (2000), o Modelo Logístico de 3 parâmetros utiliza indicadores binários, representando sucesso ou fracasso em cada item, para estimar o construto da habilidade que não é mensurado diretamente. As variáveis latentes estimadas pela TRI podem ser incorporadas como covariáveis em modelos de regressão, resultando em respostas distais. No contexto do presente trabalho, foi realizado um estudo com dados simulados em cenários distintos para verificar a subestimação do efeito da resposta distal, já discutida na literatura por Schofield (2015). Na aplicação com dados reais, a TRI foi utilizada para estimar o nível socioeconômico de estudantes do curso de Engenharia Mecânica os quais fizeram a prova do Exame Nacional de Avaliação de Desempenho de Estudantes (Enade) de 2019. Para tanto, 5 indicadores acerca da situação socioeconômica de cada estudante foram selecionados para mensurar essa variável latente. Após isso, os níveis socioeconômicos entraram como variáveis exógenas no modelo de regressão, tendo como desfecho a nota final na prova de conhecimentos específicos. Os objetivos do trabalho são os de discutir metodologicamente a subestimação do efeito da resposta distal, utilizar o método de estimação em duas etapas da TRI e da resposta distal utilizando as regressões linear e logística, oferecendo duas interpretações distintas do efeito do nível socioeconômico na nota final de cada aluno.

Referências

- [1] de Andrade, D. F., Tavares, H. R., & da Cunha Valle, R. (2000). Teoria da Resposta ao Item: conceitos e aplicações. ABE, São Paulo.
- [2] Schofield, L. S. (2015). Correcting for measurement error in latent variables used as predictors. *The annals of applied statistics*, 9(4), 2133.

Tipo de Apresentação: PÔSTER

*e-mail: ingo.dube@ufba.br

MODELANDO EPISÓDIOS DE ABSENTEÍSMO POR
DME: UMA APLICAÇÃO EM ANÁLISE DE
SOBREVIVÊNCIA

LARISSA REIS SOUSA SANTOS *
VERÔNICA MARIA CADENA LIMA †
ROSEMEIRE LEGIOVILDO FIACCONE ‡
UNIVERSIDADE FEDERAL DA BAHIA - UFBA

Resumo

O presente trabalho explora a abordagem de dados que lidam com eventos recorrentes sob a perspectiva da área de análise de sobrevivência. Os dados utilizados são provenientes de um estudo longitudinal de coorte fixa, com seguimento de 10 anos, realizado com 2661 servidores públicos com informações sobre múltiplos episódios de afastamento de licença médica relacionadas ao trabalho.^[2] Os afastamentos de interesse foram aqueles relacionados aos distúrbios musculoesqueléticos (DME), que compõem um grupo de doenças do sistema osteomuscular que são caracterizados pela inflamação e a degeneração em músculos, nervos, tendões, articulações e cartilagens, acarretando dor e limitação funcional, que podem ter longa duração e ocorrer de forma recorrente. O objetivo é identificar os fatores de riscos associados à incidência dos múltiplos episódios de afastamentos através do modelo Andersen-Gill (AG) ^[1]. Por meio do modelo é possível notar que a taxa de incidência do sexo feminino tem um aumento de aproximadamente 44% em relação ao sexo masculino, ajustado pelas demais covariáveis. Ter nível de escolaridade superior, reduz a taxa em cerca de 28% em relação ao nível inferior, controlando pelas outras variáveis. Considerando as faixas de idades, é possível observar aumento da taxa a medida que a faixa de idade aumenta, utilizando como base trabalhadores de até 30 anos: taxa de 1,51 para a faixa de idade 30 a 40 anos; 3,02 para 40 a 50 anos e 4,08 para maiores de 50 anos. Esses resultados indicam a relação que existe entre a taxa de incidência de episódios de afastamento por DME com as características dos trabalhadores e a influência de cada uma nos episódios.

*e-mail: larissarss@ufba.br

†e-mail: cadena@ufba.br

‡e-mail: fiacccone@ufba.br

Referências

- [1] ANDERSEN, P. K.; GILL, R. D. **Cox's Regression Model for Counting Processes: A Large Sample Study**. Institute of Mathematical Statistics, v. 10, n. 4, p. 1100–1120, 1982.
- [2] FRANÇA, R.; FERNANDES, R.; LIMA, V. **Fatores de risco para absenteísmo devido a doenças musculoesqueléticas em trabalhadores do setor judiciário**. Rev Bras Med Trab, v. 19, n. 4, p. 454-464, 2021.

Tipo de Apresentação: Poster

Modelagem Logística com Respostas Distais Utilizando o Modelo de Resposta Gradual: Uma Abordagem Bayesiana

PAULO GOMES DE JESUS LINO¹, LILIA CAROLINA CARNEIRO COSTA²
SILVIA REGINA RIBEIRO LEMOS MORAIS³ *
UNIVERSIDADE FEDERAL DA BAHIA - UFBA

Resumo

Métodos com abordagem frequentista têm sido propostos para estimação de efeitos latentes em desfechos distais. O objetivo do presente trabalho foi estimar a variável latente e avaliar seu efeito na variável distal, por meio de métodos bayesianos. O procedimento foi realizado em duas etapas, em que na primeira etapa é estimada a variável latente, por meio da Teoria de Resposta ao Item (TRI) para itens ordinais, e na segunda etapa avaliar o seu efeito em um modelo de regressão logístico. Estudos de simulação foram conduzidos para avaliar a performance do método ingênuo. Considerou-se diferentes cenários com variações no número de itens do questionário e tamanhos amostrais. A partir dos dados simulados, os resultados mostraram que as amostras com menor número de itens apresentaram um viés maior e mais variável, sugerindo estimativas de habilidade menos precisas. À medida que o tamanho da amostra aumenta, o viés tende a se aproximar de zero, como esperado, indicando que as estimativas se tornam mais consistentes e confiáveis. Os resultados obtidos fornecem informações importantes para o modelo de resposta gradual em estudos psicométricos, contribuindo para o avanço do conhecimento na área.

Referências

- [1] ANDRADE, D. F. de; TAVARES, H. R.; VALLE, R. da C. **Teoria da resposta ao item: conceitos e aplicações**. ABE, Sao Paulo, 2000.
- [2] CORDEIRO, G. M.; DEMÉTRIO, C. G. **Modelos lineares generalizados e extensões**. Piracicaba: USP, p. 31, 2008.
- [3] TONG, Y.; KOLEN, M. J. Irt proficiency estimators and their impact. In: **annual conference of the National Council of Measurement in Education, Denver, CO**. [S.l.: s.n.], 2010.

Tipo de Apresentação: PÔSTER

*e-mail: pgjlino@gmail.com¹, liliacosta@ufba.br², slemosufba@gmail.com³