



EMENTA E PROGRAMA DE DISCIPLINA

Nome: ANÁLISE DE DADOS CATEGÓRICOS

Código: GET00073

Carga horária semestral

Total: 80 horas

Teórica: 60 horas

Prática: 20 horas

Validade: A partir de 2010

Ementa:

Introdução aos dados categóricos. Modelos para variáveis com respostas binárias. Modelos log-lineares. Modelos para variáveis com respostas multinomiais. Análise de dados com respostas categóricas repetidas. Seleção de modelos e medidas de qualidade de ajuste. Teoria assintótica para modelos paramétricos

Programa

1. Conceitos Básicos
 - a. Características de dados categóricos
 - i. Distinção entre variável resposta e variável explicativa
 - ii. Escalas nominal e ordinal
 - b. Tabelas de contingência: forma bidimensional de uma tabela de contingência genérica
2. Modelos probabilísticos para dados categóricos
 - a. Binomial
 - b. Multinomial
3. Inferência estatística para uma proporção
 - a. Função de verossimilhança e estimação por máxima verossimilhança
 - b. Teste de significância para uma proporção binomial
 - c. Intervalos de confiança para uma proporção binomial
4. Mais sobre inferência estatística para dados discretos
 - a. Estatística Wald
 - b. Teste da razão de verossimilhanças
 - c. Teste score
5. Tabelas de contingência
 - a. Probabilidades conjunta, marginal e condicional
 - b. Sensibilidade e especificidade em testes diagnósticos
 - c. Independência
 - d. Amostragem binomial e multinomial
 - e. Comparando proporções em tabelas 2x2
 - i. Diferença de proporções
 - ii. Risco relativo
 - f. Razão de chances
 - i. Propriedades da razão de chances
 - ii. Inferência para razão de chances e \log da razão de chances
 - iii. Relação entre razão de chances e risco relativo
 - iv. A razão de chances aplicada a estudos caso-controle
 - v. Tipos de estudos observacionais
 - g. Testes qui-quadrado de independência
 - i. Estatística de Pearson e a distribuição qui-quadrado
 - ii. Estatística da razão de verossimilhanças
 - iii. Testes de independência
 - iv. Resíduos numa tabela de contingência

- h. Testes de independência para dados ordinais
 - i. Tendência linear alternativa para independência
 - ii. Força extra com testes ordinais
 - iii. Testes de tendência para tabelas $I \times 2$ e $2 \times J$
 - iv. Tabelas nominal-ordinal
 - i. Inferência exata para pequenas amostras
 - i. Teste exato de Fisher para tabelas 2×2 (*Fisher's Tea Taster*)
 - ii. P -valores e erro tipo I
 - iii. Intervalo de confiança para razão de chances
 - j. Associação em tabelas de tripla entrada
 - i. Tabelas parciais
 - ii. Associações condicionais versus marginais
 - iii. Paradoxo de Simpson
 - iv. Razões de chance condicional e marginal
 - v. Independência condicional versus independência marginal
 - vi. Associação homogênea
6. Regressão logística
- a. Interpretação do modelo de regressão logística
 - b. Interpretação da razão de chances
 - c. Inferência para regressão logística
 - d. Regressão logística com preditores categóricos (teste de Cochran-Mantel-Haenszel)
 - e. Regressão logística múltipla
 - f. Estratégias para seleção do modelo (*AIC- Akaike Information Criterion*)
 - g. Tabelas de classificação
 - h. Curvas ROC – *Receiver Operating Characteristic*
7. Modelos log-lineares para tabelas de contingência
- a. Modelo log-linear de independência para tabela 2×2
 - b. Modelos log-lineares para tabelas de tripla entrada
 - c. Inferência para modelos log-lineares

Bibliografia

- Agresti, A. *An Introduction to Categorical Data Analysis*. 2nd edition. New York: Wiley, 2007.
- Agresti, A. *Categorical Data Analysis*. 2nd edition. New York: Wiley, 2002.
- Le, C. T. *Applied Categorical Data Analysis and Translational Research*. 2nd edition. New York: Wiley, 2010.
- Paulino, C. D.; Singer, J. M. *Análise de Dados Categorizados*. Editora Edgard Blucher, 2006



ANAMARIA LIMA DE FARIAS
Chefe-Deptº Estatística
SIAPE 0311506