

Análise de Dados Amostrais Complexos

Djalma Pessoa e Pedro Nascimento Silva

2016-12-30

Contents

Prefácio

Chapter 1

Introdução

Chapter 2

Referencial para Inferência

Chapter 3

Estimação Baseada no Plano Amostral

Chapter 4

Efeitos do Plano Amostral

Chapter 5

Ajuste de Modelos Paramétricos

Chapter 6

Modelos de Regressão

6.1 Modelo de Regressão Linear Normal

O problema considerado nesta seção é o de estimar os parâmetros num modelo de regressão linear normal especificado para um subconjunto das variáveis da pesquisa. O procedimento de máxima pseudo-verossimilhança, descrito na Seção ??, é aplicado. Os resultados são derivados considerando pesos ordinários dados pelo inverso das probabilidades de inclusão das unidades na amostra. Resultados mais gerais considerando outros tipos de pesos (tais como os derivados de estimadores de razão ou regressão, por exemplo) estão discutidos em Nascimento Silva(1996, cap. 6).

6.1.1 Especificação do Modelo

Vamos supor que os dados da i -ésima unidade da população pesquisada incluam um vetor $\mathbf{z}_i = (z_{i1}, \dots, z_{iP})'$ de dimensão $P \times 1$ com os valores de variáveis \mathbf{z} , que são **preditoras** ou explanatórias num modelo de regressão M . Este modelo tem o objetivo de predizer ou explicar os valores de uma variável da pesquisa y , que é considerada como variável **resposta**. Denotemos por Y_i e \mathbf{Z}_i a variável e o vetor aleatórios que geram y_i e \mathbf{z}_i , para $i \in U$. Sem perda de generalidade, suponhamos também que a primeira componente do vetor \mathbf{z}_i de variáveis preditoras é sempre igual a 1, de modo a incluir sempre um termo de intercepto nos modelos de regressão linear considerados (tal hipótese não é essencial, mas será adotada no restante deste capítulo). Suponhamos agora que $(Y_i, \mathbf{Z}_i)'$, $i \in U$, são vetores aleatórios independentes e identicamente distribuídos tais que

$$f(y_i | \mathbf{z}_i; \beta, \sigma_e) = (2\pi\sigma_e)^{-1/2} \exp \left[- \left(y_i - \mathbf{z}_i' \beta \right)^2 / 2\sigma_e \right] \quad (6.1)$$

onde $\beta = (\beta_1, \dots, \beta_P)'$ e $\sigma_e > 0$ são parâmetros desconhecidos do modelo.

Observe que (??) constitui-se numa especificação (parcial) de um modelo marginal para um conjunto de variáveis da pesquisa, e não faz nenhuma referência direta à forma como elas se relacionam com variáveis auxiliares \mathbf{x} que eventualmente possam estar disponíveis. A atenção é focalizada na estimação de β e σ_e e sua interpretação com respeito ao modelo agregado (??).

Modelos como (??) já foram considerados por vários autores, por exemplo Holt, Smith e Winter(1980), Nathan e Holt (1980), Skinner(1989b, p.81), Chambers(1986, 1995). Eles são simples, mesmo assim frequentemente usados pelos analistas de dados, pelo menos como uma primeira aproximação. Além disto, eles satisfazem todas as condições padrões de regularidade. Assim eles são adequados a uma aplicação de procedimentos de máxima pseudo-verossimilhança descritos na Seção @ref{modpar3}.