

ANALYSIS OF LONGITUDINAL DATA – P. L. Diggle, K.-L. Liang, & S. L. Zeger (1994). Oxford: Clarendon Press.

Este livro aborda modelos estatísticos e métodos de análise de dados longitudinais usando fundamentalmente ilustrações no campo da biologia e das ciências da saúde. Dados longitudinais são aqui considerados como consistindo num número elevado de séries cur-

tas, uma correspondente a cada um dos sujeitos envolvidos, resultados da mesma quantidade medida numa sequência de tempos de observação.

O livro demarca-se do tratamento de outro tipo de informação muitas vezes associado a problemas de dados repetidos. Assim, estudos transversais, ensaios clínicos, análises de sobrevivência e séries temporais não são aqui abordados.

A questão fundamental nesta obra é a de modelar o conjunto de observações de cada sujeito tendo em atenção que estas medidas tendem a estar intercorrelacionadas. Os estudos exemplificados são formulados como problemas de regressão em que se descreve a dependência da variável de resposta das variáveis explanatórias. O tratamento de cada um dos exemplos tem em conta o tipo de variável de resposta, o número de unidades de experimentação e de observação por unidade e, finalmente, também os objectivos específicos de cada estudo.

Esta obra começa por incluir a abordagem da escolha do design tendo em vista os objectivos a atingir incidindo, fundamentalmente no problema da dimensão da amostra requerida para se conseguir um poder estatístico previamente especificado. Seguem-se métodos de análise de cariz mais exploratório – a representação gráfica das medidas no tempo, o ajustamento de curvas aos dados longitudinais e a exploração da sua estrutura correlacional. A abordagem de modelos lineares (GLM) para este tipo de dados é aqui extensivamente apresentada e é também estabelecido o contraste deste tipo de modelos com a popular análise de variância (ANOVA) para medidas repetidas. O tratamento estatístico de variáveis de resposta de tipo binário e frequencial é exemplificado nas situações correspondentes a modelos marginais, de efeitos aleatórios e de transição (cadeias de Markov). O livro termina com a discussão do problema dos dados omissos («missing values») frequentes e de inquestionável relevância em estudos longitudinais.

Através da exploração de um conjunto de 6 exemplos apresentados logo no início e que cobrem problemas no âmbito da biologia e das ciências da saúde, Digg e colaboradores concretizam a aplicação dos modelos estatísticos teoricamente apresentados ao longo dos vários capítulos.

Sem perder o rigor e a elegância na apresentação dos modelos estatísticos em jogo, os autores conseguem comunicar ao leitor um sentido de interiorização dos materiais expostos, fundamentalmente através da estrutura adoptada: ela constitui-se numa sequência de pequenas introduções que anunciam apresentações teóricas mais detalhadas posteriormente exemplificadas através de situações concretas. A linguagem utilizada está completamente descodificada ao longo da própria obra, não pressupondo do leitor grandes conhecimentos matemáticos. A noção básica de modelos de regressão e o domínio de cálculo matricial elementar parecem ser os requisitos mínimos para uma boa compreensão do texto.

Para nós, que nos interessamos mais por temáticas da Psicologia, fica-nos a pena de não terem sido escolhidas situações exemplificativas correspondentes a este domínio do conhecimento. Mas a transposição não nos parece difícil e a pertinência de uma reflexão mais aprofundada sobre o desenvolvimento deste tipo de modelos, correspondentes a medidas repetidas, no âmbito da Psicologia parece-nos justificar o esforço exigido nesta transposição.

*Glória Ramalho*