第一图书网, tushu007.com

<<多水平模型及其统计诊断>>

图书基本信息

书名：<<多水平模型及其统计诊断>>

13位ISBN编号：9787030212788

10位ISBN编号：7030212789

出版时间：2008-5

出版时间：科学出版社

作者：石磊

页数：177

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：http://www.tushu007.com

第一图书网, tushu007.com

<<多水平模型及其统计诊断>>

内容概要

本书系统介绍了多水平模型中的统计模型诊断问题。
主要内容是：介绍了异常值和影响点的定义和历史发展，数据删除法和局部影响分析的一些主要方法和结果，以及多水平模型的定义、参数估计及统计推断；总结了已知协方差结构的广义线性模型的统计诊断结果，并采用删除多个数据点的方法研究了多水平模型中固定效应及随机效应参数的诊断统计量；研究了多水平模型下基于均值漂移模型的异常点探测问题、高水平单元的局部影响分析和单个观察值的局部影响分析。
本书还介绍了目前关于多水平模型的统计软件，同时给出了利用Matlab语言编写的计算程序。
　　本书可作为数理统计专业本科生、硕士生和博士生的教材或参考书，也可作为数学、生物、医学、工程、经济领域的教师及相关科技工作者的参考书。

第一图书网, tushu007.com

<<多水平模型及其统计诊断>>

书籍目录

前言第1章 引论 1.1 基本概念　1.1.1 异常值　1.1.2 影响点　1.1.3 异常值和影响点的处理　1.2 预备知识　1.2.1 一些矩阵代数　1.2.2 矩阵微商　1.2.3 分布、估计及检验理论　1.3 数据删除法　1.3.1 Cook距离　1.3.2 基于影响函数的研究　1.3.3 残差　1.4 局部影响分析　1.4.1 Cook的局部影响分析方法　1.4.2 广义影响函数及局部影响分析第2章 多水平模型 2.1 引言 2.2 参数估计　2.2.1 IGLS估计理论　2.2.2 RIGLS估计理论 2.3 假设检验及置信区间　2.3.1 固定效应参数　2.3.2 随机效应参数　2.3.3 似然比检验 2.4 残差 2.5 数据分析及建模　2.5.1 “小学项目”(JSP)数据　2.5.2 血清胆红素数据 2.6 其他多水平模型　2.6.1 多元多水平模型　2.6.2 非线性多水平模型　2.6.3 离散数据的多水平模型第3章 GLM模型的影响分析 3.1 均值漂移模型及异常值检验 3.2 数据删除法 3.3 “删除=替代”方法　3.3.1 条件残差　3.3.2 “删除=替代”诊断 3.4 残差及单个数据点的影响度量.　3.4.1 预测残差　3.4.2 影响函数及单个数据点的影响度量 3.5 GLM模型的局部影响分析　3.5.1 协方差矩阵扰动　3.5.2 响应变量扰动　3.5.3 解释性变量的扰动　3.5.4 实例分析：血清胆红素数据 3.6 小结第4章 多水平模型的数据删除　4.1 数据删除度量 4.2 两水平模型下的结论 4.3 线性混合模型中的应用 4.4 实例分析　4.4.1 血清胆红素数据　4.4.2 JSP数据 4.5 小结第5章 多水平模型的异常点检验 5.1 异常点检验　5.1.1 均值漂移模型和检验统计量 5.1.2 两水平模型中异常点的探测 5.2 随机部分异常点单元的探测 5.3 计算问题 5.4 实例分析　5.4.1 JSP数据　5.4.2 血清胆红素数据 5.5 小结第6章 多水平模型高水平单元的局部影响分析 6.1 模型和符号 6.2 扰动理论　6.2.1 V的一般结构　6.2.2 V的块对角 6.3 局部影响分析　6.3.1 扰动的结果和局部影响测度　6.3.2 计算问题 6.4 局部影响的一步近似 6.5 实例分析 6.6 小结第7章 多水平模型观测点的局部影响分析 7.1 扰动理论及结果　7.1.1 协方差矩阵的扰动模型　7.1.2 响应变量的扰动模型　7.1.3 解释性变量的扰动模型 7.2 三种扰动模式下的局部影响分析　7.2.1 局部诊断统计量　7.2.2 计算问题　7.2.3 两种特殊的模型 7.3 实例分析　7.3.1 JSP数据　7.3.2 血清胆红素数据 7.4 小结第8章 多水平模型软件及Matlab计算程序 8.1 多水平模型软件介绍 8.2 基于Matlab多水平模型的计算程序　8.2.1 对角协方差矩阵　8.2.2 一般协方差矩阵参考文献附录 实例中的数据 附表A 血清胆红素数据fSerum Bilirubin Datal 附表B JSP(junior school project)数据

第一图书网, tushu007.com

章节摘录

第1章 引论统计模型诊断是20世纪70年代中期发展起来的统计学领域一个新的研究方向，其主要目的是评价统计模型的适当性以及识别数据中可能存在的异常值和强影响点。
在模型适当性的评价方面，线性模型中目前主要采用残差分析来判断模型拟合的好坏。
异常值的识别主要在一定的异常模型假设下进行统计检验。
而影响点的识别主要是研究数据点f或数据集）对我们关注的某个内容的影响程度并识别数据中的特殊结构，这项工作也称之为影响分析。
统计模型诊断可以为统计模型的改进提供重要的参考信息。
在某些领域中，异常值及影响点还可以为我们提供某些特殊信息：如在地质找矿中，异常值及影响点可能对应着矿产资源富集信息；而在经济领域，异常的出现还可能是某种预警信息的表现。
本书主要讨论多水平模型中异常值和影响点的识别，但主要集中于多水平线性模型。
本章主要给出统计模型诊断中涉及的基本概念和方法。
l.1节主要回顾了异常值及影响点识别的发展及相关概念；l.2节给出了本书中经常使用的一些矩阵知识；1.3节以线性模型为例，介绍了标准线性模型中数据删除法及相关结果；l.4节是关于局部影响分析的介绍。
1.1.1 异常值1.1 基本概念异常值对我们现代人来说并不陌生，我们甚至不自觉地会采用一些手段来处理现实生活中出现的异常现象。
最典型的就是在体育比赛中对裁判打分的平均算法：去掉一个最高分和最低分，再作平均。
人们对异常值的认识可以追朔到l6世纪，Bernoulli提到：“ 在200多年前，人们丢掉异常值的处理方法已经是常见的现象。
”处理和识别异常值的统计方法可追溯到l850年。
虽然异常值的识别和处理方法在现代已经发展很快，但是对异常值的定义依然有不同的理解和争论。
例如Edgeworth（1887）认为：不一致观测值（discordant observation）可以定义为那些与所在样本中其他数据点遵从的频率规则flaw of frequency）不一致的观测值。
82年后，Grubbs（1969）又这样表述：一个异常的观测值，即异常值，是严重偏离所在样本其他数据点的观测值。
这些表述实质上认为异常值是有目的的、后验的．这种有目的性的识别异常值的方法，一般只能在数据中的异常值可以预先通过视觉观察时才能使用（在一元小样本中较多）。
事实上，对样本量较大或较为复杂的数据集，比如回归、多元数据、试验设计等，预先观察到异常值是很困难的．因此，就有在观察到异常值之前制定一种客观的准则，这种准则大多依赖于异常值模型（outlier model）。
由于近几年来强调统计建模的重要性，许多研究者认为异常值是那些来自于非目标总体（某种统计模型）的观察值．Hawkins（1980）给出了一种比较明确的定义：异常值是指污染的观测值或不一致观测值的总称。
不一致观测值是指那些让调查者感到吃惊或有较大偏差的数据点。
而污染的观测值是指来自非目标总体的观测值。

第一图书网, tushu007.com

编辑推荐

《多水平模型及其统计诊断》由科学出版社出版。

第一图书网, tushu007.com

<<多水平模型及其统计诊断>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:http://www.tushu007.com