

# 《实验数据多元统计分析》

## 图书基本信息

书名：《实验数据多元统计分析》

13位ISBN编号：9787030236760

10位ISBN编号：7030236769

出版时间：2009-2

出版社：朱永生 科学出版社 (2009-02出版)

作者：朱永生

页数：188

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：[www.tushu000.com](http://www.tushu000.com)

## 前言

复杂大系统的科学研究往往都需要收集和处理大量反映系统特征和运行状态的数据信息，这类原始数据集合由于样本数量巨大，刻画系统特征的指标变量众多，并且带有随机性质，以致于形成了规模宏大、复杂难辨的数据海洋。利用统计学和数学方法对多维复杂数据集合进行科学的分析，挖掘出隐藏在复杂海量数据中的规律和信息，就是多元统计分析研究的基本内容。大型高能物理实验就是典型的复杂大系统的科学研究工作。20世纪80年代末北京正负电子对撞机（BEPC）和北京谱仪（BES）的建成，是中国高能加速器实验物理的真正开端。在北京谱仪上进行实验工作的研究组是以谱仪的名称（Beijing Spectrometer）命名的，简称BES合作组，它是由多国物理学家组成的国际合作研究组，我国物理学家在其中占有主导性的地位。北京谱仪成功地运行到2004年，获取了百—粲能区海量的高能物理实验数据。在此基础上，应用多元统计分析方法对实验数据进行分析，获得了大量居于当时世界领先水平的物理成果。其中，百轻子质量的精确测量、2—5 GeV能区R值的精确测量、共振态X(1835)的实验观察、 $\psi'$ 粒子的实验确定，更是引起当时国际高能物理界广泛瞩目的重大成就。为了保持和发展我国在高能物理百粲能区实验研究的领先地位，我国政府又拨巨资对北京正负电子对撞机和北京谱仪进行升级改造，称为BEPCII和BESIII。BEPCII的设计指标是产生粒子反应的强度约为原对撞机的100倍，BESIII的性能则比原北京谱仪有大幅度的提高。目前，BEPCII和BESIII已经完成安装，并在2008年开始实验取数。有理由期望，利用升级改造后的：BESIII，可以获得比原北京谱仪更多、更精细、更重要的物理成果。为了达到这一目标，应用比原北京谱仪数据分析更为精细、更为有效的多元统计分析方法成为一个十分重要和急迫的任务。事实上，多元统计分析方法应用于高能物理实验数据分析近年来已经成为国际高能物理界的一种普遍趋势。

# 《实验数据多元统计分析》

## 内容概要

《实验数据多元统计分析》介绍实验或测量数据的多元统计分析方法，内容包括：贝叶斯决策、线性判别方法、决策树判别、人工神经网络、近邻法、概率密度估计量法、"矩阵判别、函数判别分析、支持向量机法等，以及不同判别方法的比较。此外，还简要介绍了将多种多元统计分析方法的计算机程序汇集在一起的程序包TMVA（ toolkit for multivariate data analysis ），并分析了粒子物理实验数据分析中应用多元统计分析方法的一些实例。

《实验数据多元统计分析》可供实验物理工作者和大专院校相关专业师生、理论物理研究人员、工程技术人员及从事自然科学和社会科学的数据测量和分析研究人员参考。

## 书籍目录

前言第一章 绪论1.1 模式和模式识别1.2 模式识别系统1.2.1 原始数据获取1.2.2 原始数据的预处理1.2.3 特征提取和选择1.2.4 分类决策1.3 数据矩阵与样本空间1.3.1 数据矩阵与样本空间1.3.2 模式的相似性度量1.3.3 样本点的权重和特征向量数据的预处理1.4 主成分分析1.4.1 主成分分析的基本思想1.4.2 主成分分析算法1.4.3 降维处理及信息损失第二章 贝叶斯决策2.1 基于最小错误率的贝叶斯决策2.1.1 决策规则2.1.2 错误率2.1.3 分类器设计2.2 Neyman-Pearson决策2.3 正态分布时的贝叶斯决策2.4 分类器的效率和错误率2.4.1 分类器的效率、错误率和判选率矩阵2.4.2 错误率的上界2.4.3 利用检验样本集估计判选率矩阵和错误率2.4.4 训练样本集和检验样本集的划分2.4.5 利用判选率矩阵估计各类“真实”样本数2.4.6 分类器判定的“信号”样本中错判事例的扣除2.5 讨论第三章 线性判别方法3.1 线性判别函数3.1.1 线性判别函数的基本概念3.1.2 广义线性判别函数3.1.3 线性分类器的设计3.2 Fisher线性判别3.3 感知准则函数3.3.1 几个基本概念3.3.2 感知准则函数3.4 最小错分样本数准则函数3.5 最小平方误差准则函数3.5.1 平方误差准则函数及其MSE解3.5.2 MSE准则函数的梯度下降算法3.5.3 随机MSE准则函数及其随机逼近算法3.6 多类问题第四章 决策树判别4.1 超长方体分割法4.1.1 超长方体分割法的基本思想4.1.2 超长方体分割法中阈值的确定4.1.3 超长方体分割法的优缺点及其改进4.1.4 超长方体分割法用于高能物理实验分析4.2 决策树法4.2.1 决策树法的基本思想4.2.2 信号/本底二元决策树的构建4.2.3 决策树的修剪4.3 决策树林法4.3.1 决策树林的构建4.3.2 决策树林对输入事例的分类4.3.3 重抽样法构建决策树林第五章 人工神经网络5.1 概述5.1.1 生物神经元和人工神经元5.1.2 人工神经网络的构成和学习规则5.2 感知器5.2.1 单输出单元感知器5.2.2 多输出单元感知器5.3 多层前向神经网络和误差逆传播算法5.3.1 BP网络学习算法5.3.2 BP网络学习算法的改进5.4 Hopfield神经网络5.4.1 离散Hopfield网络5.4.2 连续Hopfield网络5.4.3 Hopfield网络在优化计算中的应用5.5 随机神经网络5.5.1 随机神经网络的基本思想5.5.2 模拟退火算法5.5.3 Boltzmann机及其工作规则5.5.4 Boltzmann机学习规则5.5.5 随机神经网络小结5.6 神经网络用于粒子鉴别5.6.1 用于带电粒子鉴别的特征变量5.6.2 带电粒子鉴别的神经网络的架构5.6.3 网络的训练和粒子鉴别效果第六章 近邻法6.1 最近邻法6.2 尼近邻法6.3 剪辑近邻法6.3.1 两分剪辑近邻法6.3.2 重复剪辑近邻法6.4 可作拒绝决策的近邻法6.4.1 具有拒绝决策的k近邻法6.4.2 具有拒绝决策的剪辑近邻法第七章 其他非线性判别方法7.1 概率密度估计量方法7.1.1 基本思想7.1.2 总体概率密度的非参数估计7.1.3 投影似然比估计7.1.4 多维概率密度估计7.1.5 近邻体积中样本数的确定7.1.6 概率密度估计法与神经网络的性能对比7.2 日矩阵判别7.3 函数判别分析7.4 支持向量机7.4.1 最优分类面7.4.2 广义最优分类面7.4.3 支持向量机第八章 不同判别方法的比较8.1 不同判别方法的特点8.2 多元统计分析程序包TMVA简介参考文献

## 章节摘录

第四章 决策树判别前面我们讨论了用线性判别函数设计分类器的方法。但是大量实际的模式识别问题并不是线性可分的，比如当两类样本的分布具有多峰性质并相互交错时，简单的线性判别函数往往会导致较大的分类错误。这种情况下就需要采用非线性分类器。从本章开始我们来讨论几种常用的、特别是在高能物理实验数据分析中常用的非线性分类方法。4.1 超长方体分割法我们首先讨论一种对于二类问题的最简单的非线性判别方法——超长方体分割法，它可以认为是决策树判别方法的一种最简单的特例，但是由于它简单、易实行的特点，在实验数据的多元分析中，特别是高能物理实验数据分析中，仍然有比较广泛的应用。4.1.1 超长方体分割法的基本思想在本节的讨论中，为了不失一般性，我们把样本分为信号和本底两个类别，信号指实验中所要研究的过程的事例样本，所有信号以外的样本都属于本底样本。超长方体分割法不是企图用一个决策规则把两类样本一次分开，而是采用分级的方法来解决分类问题。它的基本思想如图4.1所示。

# 《实验数据多元统计分析》

## 编辑推荐

《实验数据多元统计分析》是现代物理基础丛书之一。

# 《实验数据多元统计分析》

## 精彩短评

- 1、挺不错的，可以作为入门的资料

# 《实验数据多元统计分析》

## 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:[www.tushu000.com](http://www.tushu000.com)