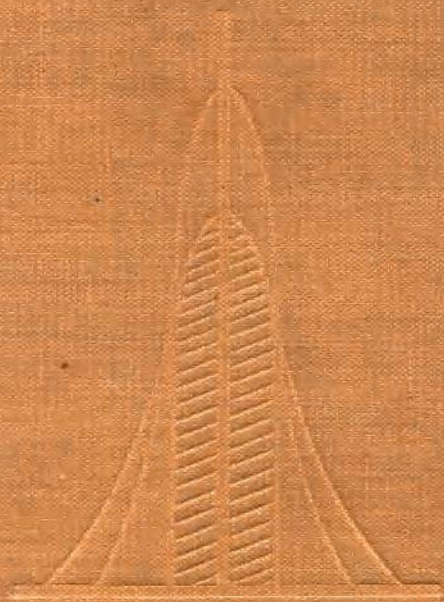
实验物理学丛书



李惕碚 著

实验的数学处理

科学出版社

实验物理学丛书

实验的数学处理

李惕碚 著

科学出版社
1980

内 容 简 介

本书介绍实验数据处理的数学方法。在引入有关概率论和数理统计的基本概念、常用分布和计算公式的基础上，讲述了测量误差理论、参数估计方法(包括贝叶斯方法、置信区间法和置信分布法)、统计假设检验、拟合观测数据的最小二乘法，以及模拟随机现象的蒙特卡罗方法。最后用附录形式简要地介绍了数据处理中常用的数值计算方法，包括插值、数值微分、数值积分和最优化。为了说明数学方法的应用，书中分析了不少实例，它们取自普通物理测量、核物理、高能物理实验及工业生产中的实际问题。

本书对象为从事物理实验的科学工作者和大专院校有关专业师生，也可供理论物理工作者及有关学科的科研人员和工程技术人员参考。

实验物理学丛书

实 验 的 数 学 处 理

李 惕 碚 著

科学出版社出版
北京朝阳门内大街137号

中国科学院印刷厂印刷

新华书店北京发行所发行　各地新华书店经售

1980年2月第一版　　开本：850×1168 1/32
1980年2月第一次印刷　　印张：12　插页：精2
印数：精1—14,850　平1—21,750　　字数：311,000

统一书号：13031·1090
本社书号：1530·13—3

定价：布面精装：2.30元
平　　装：1.50元

序

张文裕

要使自然科学在我国真正地生根和健康地成长，就必须十分重视和大力发展科学实验工作．诚然，理论是很重要的．但是应该指出，理论也必须来源于实践，才能反过来指导实践．科学发展的历史证明：只有有了科学实验和生产斗争，才有科学理论，才有自然科学．

我国是从长期沦为封建、半封建和半殖民地中解放出来的国家．近十余年来，由于受到"四人帮"严重干扰、破坏，造成了很大的思想混乱．有些人往往把理论和实验本末倒置，或错误地认为理论可以脱离实验而独自发展．实际上，实验是理论的泉源、科学的基础．因此，十分重视和大力发展科学实验，实为当务之急．

科学实验的任务是观察自然现象，定量测量有关的物理量；通过误差的数学处理而使测量得到的物理量更接近于真实值；然后通过理论分析，总结出这些物理量之间的相互关系，以求得对自然现象本质的认识．因此，要研究一个自然科学问题，首先必须有测量仪器，有数据处理的数学方法，最后才能作出正确的理论分析．这是科学研究的一般过程．至于如何选题，当然是个极其重要而又困难的问题，这里姑且不谈．

在研究物质微观结构，如研究作为近代物理核心的第三层次，即高能物理——基本粒子过程中，上面三方面都牵涉到许多新问题．在我国，科学实验比较落后，连介绍一般常用的探测仪器和探测方法的书籍也不多见；论述实验数据处理的书籍尤其缺乏；至于高能物理实验研究中所需要的这两方面的书籍更是凤毛麟角．实际上，这两类书籍的有无或多少，在某种意义上反映着一个国家的科学实验因而也是自然科学发达的水平．

李惕碚同志的这本《实验的数学处理》的出版是值得欢迎的.这本书和毛慧顺、王运永两同志正在编写的《多丝正比室和飘移室》一书，以及前几年出版的王祝翔同志写的《气泡室》和《核物理探测器及其应用》两本书，将有助于进一步促进我国正在发展且将大力发展的以高能物理为代表的科学实验工作.

《实验的数学处理》一书的主要特色是，数学为物理服务，与物理密切结合的思想贯穿全书. 发展数学方法是为了更好地解决物理问题，即更好地处理数据，使测量值经处理后更接近于真实值.因此，本书并不显得太抽象化和太数学化. 它深入地、系统地介绍了实验数据处理和分析中所用的数学方法，从基本概念、基础知识直到这个领域的最新发展以及各种应用例题，都有很好的阐述和系统的介绍. 作者还着重分析和比较了各种方法之间的联系和各自的优缺点，并加以评价；在此基础上引入了置信密度的概念和参数估计的置信分布法，从而对实验中的稀有事例给出了较好的处理方法. 书中还引用了许多高能物理例题，既有助于读者理解有关的概念和方法，同时也有现实性.

本书无论从科学水平、内容安排和应用方面来说，都正好适合实验物理工作者的需要. 我们相信，这本书的出版，将对我国实验物理工作水平的提高起很好的作用，同时对于理论物理工作者和其他学科的研究者，也是一本很好的参考书.

前　言

实验的数学处理方法，包括实验设计、实时控制和数据处理的数学方法三个部分．本书除少量属于实验设计的内容外，主要讲述如何应用数理统计方法对实验进行事后数据处理．

所谓实验的数据处理，就是从带有偶然性的观测值中用数学方法导出规律性结论的过程．在不少实验中，尤其是在现代物理实验中，现象的随机性质是十分突出的，使物理过程的规律性往往被现象表面的偶然性所掩盖，因而必须运用适当的数学工具才能恰当地设计实验，才能由实验观测数据得出正确的结论．所以，实验数学处理的内容和重要性大大超出了古典物理实验的所谓误差处理．另一方面，对于不直接从事实验工作的人，要正确地理解实验结果报道的意义，也必须对数据分析方法有所了解．因此，本书虽然是为实验工作者写的，但对于理论工作者也是有用的．

本书第一、第二、第三章介绍了有关概率论和数理统计学的基本知识，其余部分论述了测量误差理论、参数估计方法、统计假设检验、曲线拟合技术和随机现象的数学模拟——蒙特卡罗法等．但数据处理中的大量计算工作必须借助于电子计算机才能完成，而且函数关系的复杂性也常常使解析方法求解成为不可能．所以，数值方法是完成数据处理的一个必要的数学工具．为此，我们在附录中简要地介绍了数据处理中一些常用的数值计算方法．

本书假定读者已熟悉微积分以及向量和矩阵的基本运算．书中例子大多数取自高能物理实验．不熟悉高能物理的读者，可以绕过物理理论和实验技术的细节，从如何应用数学处理方法的角度来阅读这些例子．认真地从事过数据分析工作的科技人员，大概都会感到现成的数学工具往往不敷应用，有必要进一步发展实验数据处理的数学理论．一般说来，实验数学处理方法常常是在

解决具体的科学和技术问题中发展起来的，所以，实验工作者应当善于从应用于其它学科的数学方法中找到对自己有用的东西.

本书部分内容曾在云南高山宇宙线观测站对在该站实习的北大物理专业72届基本粒子班学员讲授过，高能物理研究所在云南站工作的同志也参加了听课. 在教学过程中和他们进行的多次热烈讨论，使作者获得不少启发，在此向听课同志表示谢意.

张文裕先生始终热情关心和鼓励本书的写作，从全书内容到表述方式，都提出了不少宝贵意见，书中论述蒙特卡罗法的一章，就是按照他的建议补写的，他还于百忙中为本书写序，作者谨致谢意. 作者在从事物理实验和学习数理统计的过程中，曾得到肖健先生的指导，他同作者深入地讨论了粒子质量估计问题，从而使作者提出了置信分布的概念和方法. 在最后定稿过程中，何泽慧先生、高崇寿同志以及高能物理研究所高能天体实验组的全体同志，都给予了十分宝贵的支持，作者谨向他们表示衷心感谢.

作　者

符 号 表

$P_r(A)$	事件 A 的概率	(3)
$P_r(A\|B)$	A 对于 B 的条件概率	(5)
$P(x)$	随机变量 x 的分布函数	(12)
$p(x)$	随机变量 x 的概率(密度)函数	(12)
$x\sim f(x)$	随机变量 x 服从分布 $f(x)$	(16)
$\langle x\rangle$	随机变量 x 的期待值	(20)
$\langle x^n\rangle$	随机变量 x 的 n 阶矩	(20)
$\mathrm{Var}(x)$, $\sigma^2(x)$	随机变量 x 的方差	(20)
$\mathrm{Cov}(x, y)$	随机变量 x 和 y 的协方差	(22)
$\rho(x, y)$	随机变量 x 和 y 的相关系数	(22)
$\phi_x(t)$	随机变量 x 的特征函数	(26)
$N(x;\mu, \sigma^2)$	正态分布函数	(37)
$n(x;\mu, \sigma^2)$	正态密度函数	(37)
u	标准正态变量	(38)
V	协方差阵	(51)
W	权重矩阵	(52)
Q	正态变量的二次型	(52)
r	均匀分布随机变量	(59)
u_ξ	标准正态变量的误差限	(72)
fwhm	半峰宽度	(74)
S_x^2	样本标准偏差	(79)
$S_{\bar{x}}^2$	平均值标准偏差	(81)
$\chi^2(\nu)$	自由度 ν 的 χ^2 分布	(81)
$\chi_\xi^2(\nu)$	自由度 ν 的 χ^2 分布的 ξ 点	(82)
t_ξ	t 分布的 ξ 点	(92)
S_{xy}	样本共差	(97)
$\hat{\theta}$	θ 的估计量	(119)
L	似然函数	(125)
$\bar{x}$	样本加权平均值	(134)
$P_c(\theta\|\hat{\theta})$	θ 的置信分布函数	(204)
$p_c(\theta\|\hat{\theta})$	θ 的置信概率(密度)函数	(204)

H	统计假设,原假设	(215)
H'	备择假设	(216)
χ^2	χ^2 量	(222)
$S_N(x)$	样本分布函数	(225)
$\chi^2_{\min}$	拟合的最小 χ^2 值	(266)

目　录

第一章 随机变量

§1.1 测量结果的随机性

使用相同的测量仪器，在一定的实验观测条件下重复观测某一个物理量的数值，每次测量得到的结果也会各不相同．因此，物理实验的测量结果具有偶然性，或者叫做随机性．这种随机性的原因在于：

（一）测量的偶然误差　所谓“一定的实验观测条件”，是相对于观测者所能控制的程度而言．实际上，由于实验技术水平的限制，总是存在着观测者尚不能完全控制的某些偶然因素，使得各次观测的实验条件发生变化，例如观测时的环境条件，测量仪器的状况，等等，从而造成测量的偶然误差．另一方面，观测者本身感官分辨本领的限制，也是偶然误差的一个来源．偶然误差的存在使实测数值带有随机性．

例如，用一个测微仪重复测量某一金属丝的直径，各次测得的值并不完全相同，而是在某个平均值附近摆动；又例如，用云雾室测量荷电粒子在磁场中运动径迹的曲率，推算粒子的动量，由于云雾室内气流运动的无规则变化，往往导致径迹曲率的无规则畸变，使得即使对动量相同的粒子，每次测得的曲率也各不相同．虽然提高实验技术水平（例如使用精度更高的测微仪仔细地读数；更严格地控制云雾室工作条件，以减小气流的无规则运动等），可以使各次测量结果彼此靠近，但偶然误差总是存在的，并不能完全消除．

（二）物理现象本身的随机性质　处于平衡状态的一个宏观物体，其热力学量（温度、密度、压强等）的数值是一个统计平均值，

这些物理量的实际数值时时刻刻都围绕着平均值发生微小起伏.在有些物理现象中,热力学量的起伏可以表现得很明显.在微观现象领域,特别是在高能物理实验中,物理现象本身的统计性更为突出.按照量子力学的原理,对处于同一个态的微观粒子,测量同一个可观测的物理量时,即使不存在任何测量误差,各次测量结果也会不同,除非粒子处于这个可观测量的本征态.例如,同一种基本粒子的寿命,其实测值分布在从相当短到相当长的时间范围内;高能基本粒子在相互作用过程中产生各种不稳定粒子,由各个作用事例测出的同一种不稳定粒子的质量,是在某个质量范围内的一个分布,等等.这些测量结果的离散程度都大大超过了测量的偶然误差可能造成的离散.因此,测量结果的随机性主要反映了物理现象本身固有的随机性质,它更是无法用提高测量精确度的办法予以消除的.

必然性和偶然性是对立的统一,“被断定为必然的东西,是由纯粹的偶然性构成的,而所谓偶然的东西,是一种有必然性隐藏在里面的形式”[1].实验数据是某些物理量的一些观测值,它们是带有偶然性的一些随机量.但是,“在表面上是偶然性在起作用的地方,这种偶然性始终是受内部的隐蔽着的规律支配的,而问题只是在于发现这些规律.”[2] 实验数据处理的任务正是要找出在随机数据中隐蔽着的这些规律性的东西.

既然测量结果是一些随机量,对它们进行处理和解释,就必须应用随机量的数学——定量研究随机现象规律性的科学.否则,就得不出合理的结论,也无法知道结论的可靠程度.在现象本身的随机性很突出的场合,例如在高能物理实验中,如果不利用适当的数学工具去设计实验和对测量结果进行分析,物理现象的规律性甚至会被表面上的偶然性完全掩盖掉,使观测者不知道应该如何去理解他的观测结果.所以,实验的数学处理并不只限于通常所

1) 恩格斯,《路德维希费尔巴哈和德国古典哲学的终结》,人民出版社,1972,第35页.

2) 同上注,第38页.

说的“误差处理”，它是由观测结果导出规律性结论的过程，而测量的偶然误差的规律性只是其中的一个部分.

随机量数学是一门正在迅速发展的学科，它已形成了自己的算术、代数和分析——概率论、数理统计和随机过程理论，它在工农业生产和科学实验中正在得到越来越广泛的应用，生产斗争和科学实验的发展，也不断地对它提出很多需要解决的新课题. 本书介绍概率论和数理统计学的一些基本内容和如何应用它们来处理实验数据.

§1.2 随机事件的概率

1.2.1 随机事件及其概率

在一定的试验条件下，现象 A 可能发生，也可能不发生，并且只有发生或不发生这样两种可能性，这是偶然性现象中一种比较简单的形态，我们把发生了现象 A 的事件叫做随机事件 A，简称事件 A.

如果在既定的条件下进行一组试验，总共试验 N 次，其中现象 A 发生了 N_A 次，则该组试验中事件 A 的频率是比值 N_A/N. 重复进行很多组这样的试验，我们就会发现随机事件的频率具有某种规律性：频率总是在某个值的上下摆动，并且随着每组试验次数 N 的增多，频率上下摆动的平均幅度趋于减小，出现显著大或显著小的频率的可能性减少. 因此，在统计的意义上，随机事件的频率存在着一个极限值，叫做事件 A 的概率，记作 $P_r(A)$：

$$\frac{N_A}{N}\xrightarrow[N\to\infty]{}P_r(A). \tag{1.1}$$

随机事件的频率在其概率值的上下摆动，并且随着试验次数的增多趋近其概率值，是随机现象存在着内在的规律性的表现.

若事件 A 与事件 B 是两个不同的随机事件，A 和 B 的和事件 $A+B$ 指 A 与 B 至少有一个发生的事件. 如果用两个圆分别表示事件 A 和事件 B 的集合，如图 1.1 所示，则两个圆的总和(图