〔德〕 B.R.弗里德 著

王家道 尹世松 译 王士依 朱 锜 校

# 概率与统计在光学研究中的应用

# 概率与统计在光学研究中的应用

〔德〕 B.R.弗里德 著
王家道 尹世松 译
王士依 朱 镝 校

宇航出版社

## 内 容 提 要

本书用完美而简单的考尔莫高洛夫（Kolmogoroff）公理阐述了概率论的基本概念，通过湍流理论、激光散斑和成象理论等例子说明了概率统计是怎样在光学研究中得到应用的。

本书重点放在：建立"统计模型"，讨论蒙特卡洛(Monte Carlo)计算和各种近似计算。同时也介绍了抽样定理、快速傅里叶变换以及其他简捷算法。本书可供光学及有关科研工作者和大专院校师生参考。

The Computer in Optical Research
Methods and Applications
3. Computational Methods of Probability and Statistics
B. R. Frieden
Springer-Verlag Berlin Heidelberg 1980.

**概率与统计在光学研究中的应用**

〔德〕 B.R.弗里德 著
王家道 尹世松 译
王士依 朱锜 校
责任编辑：张芝

☆

宇航出版社出版
新华书店北京发行所发行
各地新华书店经销
顺义县印刷厂印刷

☆

开本：787×1092 1/32 印张：6.125 字数：142 千字
1989年2月第1版第1次印刷 印数：1—3000册
ISBN 7-80034-171-2/O·001 定价：2.50元

# 译 者 的 话

本书根据《The Computer in Optical Research Methods and Applications》一书的第3章译出。

原书是一本关于计算机在光学研究中的应用的综合性著作，由几位专家联合编写成的。全本共分6章，各章基本独立自成体系，共约30万字。书中内容较为广泛，主要介绍：光学衍射理论中的积分计算；概率与统计的计算方法；光学中的最优化方法；计算机和天文光学；计算机形成全息图等等。

原书第3章“概率与统计的计算方法”系由该书主编B.R.弗里德编写。作者鉴于近年来概率与统计方法在许多研究领域中都得到了成功的应用，然而它在光学研究中的应用很有限，故特编写此章。

作者就概率与统计的基本原理和各种计算方法作了全面而简要的叙述。深入浅出地介绍了用某些概率、统计法分析一些光学问题的实例。说明在电子计算机的应用日益广泛的今天，许多概率律在计算机上可以直接存取，用概率、统计方法处理、分析一些光学问题是比较方便而简单的。

我们认为该章通用性较强，实用价值较高，特单独将此章译出，更名为《概率与统计在光学研究中的应用》，作单行本出版。

由于译校者水平有限，书中难免存在缺点和错误，敬请读者批评指正。

1987年5月

# 目　录

# 绪　论

光学是研究光，特别是研究传递信息的光的科学。光学设计、光谱学、激光器设计等等都有一个共同的目标，要提高光束提取信息或传递信息的能力。由于携带信息的小信使——光子——具有统计性质，所以达到以上所述的提高的目标是可能的。由于光子的量子性质，光子并不是循预定的轨道通过介质，而是产生扩展，甚至不能保持固定的波长；根据海森伯格（Heisenberg）的测不准原理，这个量也有一定的扩展。

这些不确定性导致了光学中的“衍射扩展”与泊松（Poisson）象噪声。实际上，衍射扩展就是用概率密度来描述光子的随机轨迹，而泊松象噪声则反映了有限数量的光子的轨迹。以上两个问题基本上是具有统计性质的。虽然习惯上把衍射扩展看成是确定性的。然而，在低亮度情况下，这种确定性的观点也必须予以放弃（参见4.4节）。

因此，光的随机性和测不准性存在于最基础级的、直至光子级的光学中。虽然光子是一个非随机性的或确定性的实体，但其它不确定性的源也会不知不觉地混进来（记住：宇宙间的熵和混乱程度总是不断增加的！）。

各种测试仪器的性能是不完善的。除了光子的统计性质以外，还有一个在光学课题中经常遇到的基本统计量，那就是起干扰作用的“噪声”。任何一个实际的物理信号的测量结果都包含随机误差。我们把这种误差称为噪声，就象我们

把有害的植物称为“杂草”一样。噪声分析是采用统计学研究的一个重要课题。虽然任何一种测量工作都会遇到无法预测到的噪声量，即个别噪声值是无法预测到的，然而它们的长期的平均值是可以预测到的。通过多次观测，能作出噪声值的分布图，可用它来进一步改善仪器的性能。

光学中不确定性的第三个来源是数据不足。对图象抽样过于粗糙就是一个例子。如果每隔尼奎斯特（Nyquist）间距进行一次抽样，则根据数据可以唯一地把由诺抽样点的连线组成的整个图象再现出来。然而，进行粗糙的抽样，即使数据没有噪声，与这些数据相符的可能出现的图象有无限多个。数据不充分驱使人们寻找概率统计的答案，例如寻找“可能性最大的”象。通常，在统计学的领域里，论述这个问题的理论称为“估计理论”。

我们将要讨论的概率统计方法，除了研究不确定性的问题以外，还可用以解确定性问题！这个方法十分有效，可以用来分析非统计性的情况。例如，在几何光学中，确定光斑图或光强的点扩展函数。对于一个给定的透镜结构，光强分布并不存在不确定性，光强分布图形的界限是分明的。因此这个问题似乎不是统计学的问题。然而，认识光强分布是唯一的是一回事，而在实践中去确定这个光强分布是另一回事。一种常用的估计光强分布的方法是对通过透镜系统的光线进行追迹，建立光线投射到象平面上的分布图。这个分布图接近未知的分布图。然而，如果投射到象面上的光线不多，由此得出的分布图具有很大的误差。还有一个比较好的估计方法，它是根据概率理论推导出来的，此种方法的优点是每一对（指定的）光线都能在理论光点强度分布曲线上以任意的精确度建立一个点。本章将对此方法作进一步的阐述。

到目前为止，我们一直把“概率”和“统计”这两个术语未加区别地使用。事实上，根据它们所处理的问题来看，两者是有区别的。概率论是论述分析性的问题，即一组已知的数据输入到一个已知的物理系统中后，数据将怎样分布。统计学则是处理相反的问题，即已知一个实验的输出数据，要求根据此数据找出输入数据的某些性质。统计学的宗旨是不对有关的物理系统的性质作任何假设的情况下得出答案。我们把这种研究问题的方法称为统计估计法。例如在没有其他信息（如未给出用来测定这些数据的实验方法）的情况下，给出光速$c$的一组测量值，要求确定这些数据的准确度。下面我们将介绍一种很简单的检验数据的方法——$t$检验——来解决这个问题。

统计学中的许多传统方法，如$t$检验、$F$检验以及$\chi^2$检验等，在许多研究领域中，例如在心理学中，都得到了成功的应用。但它们在光学中几乎没有得到应用。例如为分析未知物理现象打下基础的线性回归这样有用的近似法也没有得到应用。本书的主要目的就是要说明这个方法是怎样从概率论中引出来的，和它们适用于光学的哪些方面。

我们将用完美而简单的考尔莫高洛夫(Kolmogoroff)公理从头开始阐述概率论。并将通过湍流理论、激光散斑以及成象理论等例子来说明概率论是怎样在光学中得到应用的。重点将放在以下几个方面：建立“统计模型”，其目的是把问题简化到能够求解的程度；讨论“蒙特卡洛（Monte Carlo）计算”，它几乎可以成为一门独立的计算机学科，它在遥感理论中起着日益重要的作用；以及各种近似计算方法。本书还要介绍抽样定理、快速傅里叶变换法，以及其他简捷算法。

下面将要讨论的几种分析方法，例如卷积定理、传递定理、雅可比变换等等，所有这些都是从事光学研究的人们所熟悉的，不难看出它们都是线性理论的内容。我们认为，通过讲授概率论，那些修完一学期线性理论课程的学生是能很容易地学会这些内容的。

我们将首先应用考尔莫高洛夫的三个公理，阐明基础概率论。对于那些时间有限或者凭直觉能掌握基础概率论的读者来说，可以跃过前面几节，直接阅读第2节，这样效果比较好。第2节论述了统计方法和线性理论之间的重要而有趣的联系。

# 1. 基 本 概 念

为了使本节简明扼要，我们避免采用严格的集合论论证方法，而采用一种启发式的处理问题的方法。集合论仅给术语“或”和“与”以严密的数学含义。而我们只运用它的直观意义。如果读者对严格的方法感兴趣，可参阅帕波里斯（Papoul s）的著作〔1〕。

## 1.1 实验的概念；事件

“实验”一词可以简要地表述成——为获得结果而进行的任何一个确定的过程。实验的每一次重复称为一次“试验”。实验可以是实际存在的（如在实验室里进行的）也可以是想象的。但是，在任何情况下，它们必须是可以多次重复的。这些试验产生一系列由实验可以直接观测到的结果。实验的可能结果称为“事件”。

以掷骰子的实验为例。实验的结果是一个数字（在其它实验中可能是许多个数字），即骰子顶面上的点数。投掷四次后，掷出点数可能是4、2、1和5。每次的掷出点数是1到6之间的任一个数。可以用掷出点数是“偶数”或“奇数”作为描述实验的事件$\{An\}$。

**事件“空间”** 我们把一个实验可能出现的所有的$N$个事件的集合定义为事件“空间”。例如在掷骰子的实验中，事件$\{An\}=(1, 2, \cdots, 6)$是空间元素。而事件空间就是事件

$\bar{C}=(A_1或A_2或\cdots或An)$。

然而，空间元素不是唯一的。同样一个实验可以用许多不同的可能出现的空间来描述。例如在掷骰子的实验中，空间元素可以用下述事件描述。

$\{An\}$ = （掷出点数全部小于4；掷出点数全部等于 4；掷出点数全部大于4），

$n=1，2，3$　　(1a)①

根据待估计的概率来确定使用什么样的空间(详见1.4节)。

**不相容事件**　归根到底正是应用了这个概念我们才能把概率计算出来。如果事件$A$的出现和事件$B$的出现是互相排斥的，称事件$A$和事件$B$为不相容事件或独立事件。例如掷骰子掷出了“2点”，则它就排斥了“5点”。因此，事件“2点”和事件“5点”是独立的。

“小于4点”与“大于2点”是两个非独立事件。如果结果是“3点”，很明显，事件“小于4点”和事件“大于2点”两者同时出现。因此这两个事件不是互相排斥的。

**必然事件**　我们来考察掷骰子实验中的事件“1点到6点之间的任何一个数”。在重复每次试验时是必然出现这个事件的，因此称它为“必然事件”。通常任何一个实验都有一个“必然”事件与它相联系，这就是它的全部事件空间$(A_1或A_2或\cdots An)$。

## 1.2　概率的定义

与实验中每一个可能出现的事件$A$相联系的是事件$A$出

① 此式原文有误，已改正. ——校注

现的概率$P(A)$。根据定义，这个量遵循以下公理（见参考文献〔3〕）。

公理Ⅰ　$P(A) \geq 0$。即概率决不会是负数。

公理Ⅱ　$P(C)=1$，式中$C$是必然事件（见上）。这一性质确定了概率数值的变化范围。

公理Ⅲ　如果事件$A$与$B$是互不相容的，则$P(A$或$B)=P(A)+P(B)$。例如在掷骰子试验中，$P$（1点或2点）$=P$（1点）$+P$（2点）。这里的“或”字要按一般的意义来理解；集合论从数学上对它下了严格的定义。

根据这三条公理以及由公理Ⅲ推广到无数个互不相容事件的推论（详见后文），可以导出概率论中所有的定理，因此也可导出全部统计方法。本书目的之一就是要说明这种说法是正确的。

## 1.3　一些基本推论

在上述公理的基础上作以下说明：假如$\overline{A}$表示“非$A$”，它是$A$的一个对立事件。显然$A$与$\overline{A}$是互不相容事件（见1.1节）。因此，根据公理Ⅲ，$P(A$或$\overline{A})=P(A)+P(\overline{A})$。此外，事件$A$与事件$\overline{A}$一起构成一个必然事件（见1.1节）。因此，根据公理Ⅱ，$P(A$或$\overline{A})=P(A)+P(\overline{A})=1$。将此等式与公理Ⅰ一起用于$A$，必然得到$P(A) \leq 1$。把此式与公理Ⅰ相结合，可得到

$$0 \leq P(A) \leq 1 \tag{1b}$$

所有的概率位于0和1之间。概率为零的事件称为“不可能”事件。因此，所有事件都处于必然与不可能事件之间（这点并不深奥，然而是必不可少的）。

**可加性** 考察两个不相容事件$A_1$和$A_2$。假设事件$A_2$与事件$A_3$也是互不相容的。暂且命$B_1$ = （$A_2$或$A_3$）。则$A_1$与$B_1$是互不相容的（可直接判定），根据公理Ⅲ，$P$($A_1$或$B_1$）= $P$（$A_1$）+ $P$（$B_1$）。然后把最后两个关系式结合在一起，可得到

$$P(A_1\text{或}A_2\text{或}A_3)=P(A_1)+P(A_2+A_3)$$

再次应用公理Ⅲ，可得

$$P(A_1\text{或}A_2\text{或}A_3)=P(A_1)+P(A_2)+P(A_3)$$

显然，这个方法可以推广应用于$n$个互不相容的事件中。因此，如果事件$A_1$，$A_2\cdots An$是互不相容的，则有

$$P(A_1\text{或}A_2\text{或}\cdots\text{或}An)=P(A_1)+P(A_2)+\cdots+P(An) \tag{2}$$

概率的这个性质称为$n$个互不相容事件的“可加”性。应该指出，严格地说上式不能推广到$n=\infty$的情况中。我们把$n=\infty$时的相加性称为公理Ⅳ。

**归一化性质** 假设事件$A_1$，$A_2$，$\cdots A_N$是互不相容的，并组成一个事件空间（见1.1节），则

$$(A_1\text{或}A_2\text{或}\cdots\text{或}A_N)=C$$

$C$是必然事件。于是，结合公理Ⅱ和（2）式可得

$$P(A_1)+P(A_2)+\cdots+P(A_N)=1 \tag{3}$$

这就是事件空间的归一化性质。由以上可见，（3）式实际上说明了在给定的实验中必然会发生可能事件$A_1$，$A_2$，…，$A_N$中的某一事件。

**边缘概率** 假设$P$（$A_mB_n$）是联合事件$A_m$和$B_n$的概率，其中$\{B_n\}$是互不相容的，并形成一个空间。如果$P$（$A_mB_n$）是已知的，能否设法计算出$P$（$A_m$）呢？

事件$A_m$与联合事件$A_m$和（$B_1$或$B_2$或…或$B_N$）是等价

的，由于后者是必然事件。这个联合事件也可以表述为（$A_m$和$B_1$）或（$A_m$和$B_2$）或…或（$A_m$和$B_N$）事件，这些事件是互不相容的，由于$\{Bn\}$是互不相容的。于是，应用（2）式得到

$$P(A_m) = \sum_{n=1}^{N} P(A_m B_n) \tag{3a}$$

$P(A_m)$称为边缘概率。

## 1.4 概率的“传统”定义

假设（3）式中所有的概率$P(A_n)$（$n=1$，…，$N$）都是相等的。于是，根据（3）式，$P(A_n)=1/N$。

再假定将事件$B$定义为事件空间中的一个子集（$A_1$或$A_2$或…或$A_n$），$n \leqslant N$，则根据（2）式

$P$（$A_1$或…或$A_n$）$\equiv P(B) = n\,(1/N)$

或

$$P(B)\ n/N \tag{4}$$

这就是大约30年前对概率下的“传统”定义。（4）式中的$n$是不相容事件的数目，其中每一个事件的发生都会引起事件$B$的发生。此外，$N$是可能发生的不相容事件的总数（4）式表明，$P(B)$就是使事件$B$能出现的独立方式数与可能事件总数的比值。学校里的作业就有排列未知数$n$和$N$的计算练习。至今这还能唤起人们对从缸中模红球与白球并计算其出现概率的回忆。

（4）式是我们遇到的第一个可用来计算概率的表达式。在许多普通的实验中，$n$和$N$可以直接计算出来（即使并不容易）。这一点可用下述例子来说明。