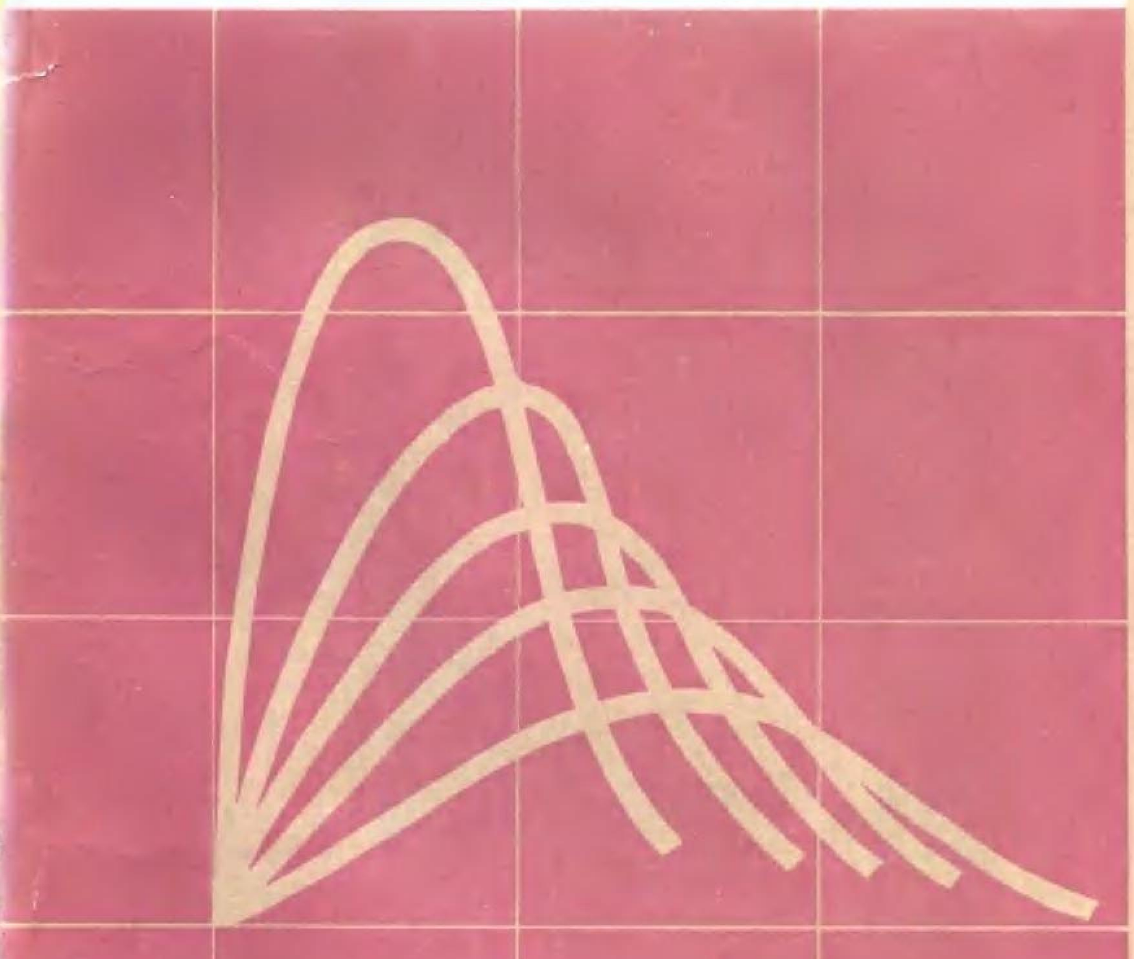
观测与最小二乘法

E.M.Mikhail 等



测绘出版社

观测与最小二乘法

E.M.Mikhail 等著
唐昌先 邹笃醇译
胡明城 校

测绘出版社

本书包括三部分：第一部分介绍一些基本概念；第二部分介绍几种最小二乘平差的通用方法 第三部分论述一些新的课题，包括最小二乘联合平差法、序贯数据处理、最小二乘内插、滤波及拟合推估等内容。

本书可供高等院校测绘专业教学参考，也适合有关人员自学用。

Edward M. Mikhail
With Contributions by
F. Ackermann
OBSERVATIONS AND LEAST SQUARES
IEP—A Dun-Donnelley Publisher, New York, 1976

观测与最小二乘法
E.M.Mikhail 等著
唐昌先 邹笃醇 译
胡明城 校
*
测绘出版社出版
测绘出版社印刷厂印刷
新华书店北京发行所发行·各地新华书店经售
*
开本 787×1092 1/16·印张 23·字数 528千字
1984年6月第一版·1984年6月第一次印刷
印数 1—10,400册·定价 2.40元
统一书号：15039·新276

前　言

本书是作者和F·Ackerman教授多年来对大学生和研究生讲授最小二乘法的教材经不断修订而成。本书阐述的内容能使初学者对基本概念有完整的了解。也能为研究生提供充分的学习资料。因此，本书既适用于观测和最小二乘法方面的初级课程，也适用于一、二门高级课程。此外，本书详细地列举了大量例题，并作了完全推导，使得本书也可以作为自学的参考书。

本书分三部分：第一部分阐述基本概念；第二部分详述了最小二乘平差的通用方法；第三部分讨论一些高级课题。正文后的附录提供了必需的补充资料。

第一章介绍了观测和数学模型的概念，这是应该首先阅读的。第二章是与本书主题有关的各种统计概念的摘要。由于作者不打算用这一章代替统计文献，所以只作了很少的理论推导。但是，就学习概要来说，它包括了一维和多维两种情况，已经充分完备了。初学者在开始时不需要阅读这一章的全部。第三、四章探讨误差及其传播。“误差”是一个通用的术语，它是与在统计概念上更为适宜的术语“观测”发生联系的。关于误差传播，读者需要熟悉矩阵运算。因此在阅读第四章之前，必须学习附录 A。

作为大学生的初级课程，可以包括第一部分中的第一章、第二章的一部分和第三、四章；第二部分中的第五、七章、第八和十一章的一部分。如果初级课程还需兼顾研究生时，则应加入第六章。这样，学生就能领会基本统计学、误差传播和二、三种用条件方程的最小二乘平差方法。本书把“条件方程”这一术语广义化了，意味着含有观测的一切方程。

中级课程可以包括第二章；阅读第四章和第二部分中的第六、九、十、十一章，以及第三部分中第十二和十三章的一部分。在中级课程的后期，学生将学习强制平差法(只包含参数的方程)，把模型中全部变量作为观测结果来处理的联合平差法；逐次平差法；以及序贯数据处理法。

为了更高级的学习，在第三部分中的第十二、十三章的后半部分以及第十四章为读者提供了进一步的资料。特别在第十四章中，介绍了内插(推估)、滤波和拟合的概念。阅读这一章之后，学生就更好地具备了条件，通过有关技术文献的学习，来加深对这些高级课程的认识。

各课题的叙述是循序展开的，每一种方法都是其前面较为简单的方法的逐步延伸。为了跟随这种循序展开，必须熟悉矩阵代数和学习附录 A。虽然可以改用广义矩阵代数，但就本书所包括的课题来说，作者认为无此必要。

本书叙述的内容是完整的。但在书末还附有参考文献，为读者提供补充读物。读者需要寻找特定课题的参考资料时，应当查阅这些文献。

本书所用的术语在第一次出现时，通常都加以定义。但也在不加定义的情况下，引用了某些通用的统计学术语和平差术语。读者如果不熟悉这些通用术语，可参考统计学词汇或 1961 年出版的国际摄影测量协会的多种语言词典。

目　　录

第一部分

第二部分

第三部分

第一部分

第一章　观测和模型的概念

1.1 引　言

只有在可供利用的数据的个数超过了唯一的确定未知量所需要的最少个数（即有多余数据）时，平差才是有意义的。由于这些数据通常是从受到随机波动（通常称为误差）影响的观测（或测量）取得的，因而在一个观测组中，每一充分的子集所得的结果不同于另一子集，在这个意义上，多余数据通常是互相不一致的。为了取得一个唯一的解，要应用一个附加的准则(例如最小二乘)。

为了便于了解平差方法的细节，本章介绍了若干基本概念和它们的互相关系。

1.2 观测（测量）

观测和测量这两个术语，今后将交替使用，它们具有相同的意义。观测（或测量）实际上常常用来既指操作或操作过程，又指这种操作的实际结果。平差中把这种结果（特别是数值结果）称为观测。这种观测数据是科学和工程技术的基础，因为它们为分析和处理提供了材料。

随便地思考一下测量的概念，不会产生什么问题，因为我们能够容易地接受例如测量距离、角度、温度和速度这种表面上似乎简单的观念。但是，当我们进一步考察称为“测量”的这种操作时，事情就不是那样容易了。事实上，即使是用带尺测量距离，也是一种相当复杂的操作，分析这种操作，就可以导出如下的测量的基本性质：

1. 测量常常意味着实施一种实际操作；测量过程通常包括若干个基本操作，如准备，仪器的整置或校正(或兼有两者)、照准、匹配和比较。

2. 由过程所得的读数或结果被认为是代表该测量过程的。因此，作为测量结果所得的数字，含有产生该数字所在的环境和过去的有关数据。

3. 无论怎样简单的测量，几乎总是用仪器或者是借助于自然事件来实施的(简单的计算某一事件的次数是例外情况)。

4. 各种测量是以一些标准为参考来实施的，而这些标准是根据协定建立起来的，有些任意性。这样，所谓测量，本质上就是与标准进行比较。因此，由于测量结果总是具有

某种量纲的，所以人们就要关心单位和量纲。

5. 对测量问题作深入的思考，就会发现测量还涉及到一些理论概念，例如用于距离和角度的几何抽象，在自然界就没有真实的和直接的等同物。为了描述某些自然要素，我们选择了与我们有关的一些概念，如位置，面积或长宽。

6. 虽然实施测量是一种操作或过程，但由这种过程所得出的结果，只有与测量所根据的和所涉及的理论概念相联系时，才能取得测量的意义。

测量所涉及的理论抽象称为模型。在本书所探讨的科学和工程技术中，所谓模型，几乎总是指数学模型。虽然这样的模型概念好象是理论性的，但对平差课题来说，事实上它具有根本的重要性。彻底了解模型概念将大有助于推进平差工作。

1.3 数学模型

由于本书中所探讨的大部分内容属于数量问题，因而所考虑的总是数学模型。这里把数学模型定义为人们在描述一种实存状况或一组事件时所采用的理论体系或抽象概念。这样的描述不一定意味着是完全的或彻底的，只是关系到所要思考的那些方面或性质。由于每一种模型是为特定目的服务的，所以从一个观点到另一个观点，建立的模型可能是各种各样的。这样，同样的实存体系就可用一个以上模型来描述。为了估计实存状况，就用模型代替实存状况。

数学模型通常认为是由两部分组成：函数模型和随机模型。函数模型一般描述实存状况的确定的性质或者描述正在研究中的事件；而随机模型则表征和描述所含的变数（特别是表示观测的变数）的非确定性或随机性。任何时候都要一起考虑函数模型和随机模型，因为有几种可能的组合，每一种组合代表一种可能的数学模型。

1.3.1 函数模型 无论什么时候规划测量工作，通常都要选择某种函数模型，用来代表与测量相联系的实存体系统或虚拟体系。事实上，测量的目的，通常都是为了估计函数模型的某些或全部参数值。在大地测量学、摄影测量学和普通测量学中，我们一般是探讨与时间无关的几何模型，有时探讨动力模型，例如以下的一些模型。

1. 普通测量中的几何模型：在欧几里得空间中用三个角度、三个顶点或三条边表征的平面三角形，或者还要相对于一个坐标系定向。

2. 摄影测量学中的几何模型：航摄照片看成是（几何）地面点的透视影像。

3. 大地测量学中的动力模型：地球重力场。

4. 摄影测量学中的动力模型：轨道摄影和时间的扫描摄影。

因为所用的几何模型比较简单，而且易于形象化，模型元素以及与其有关的物理元素，通常不明显地区分。但必须承认，自然界并不存在诸如点、角、距离或坐标这一类客体。这些只是函数模型的元素，用来描述自然客体的对应特征或它的位置关系。

函数模型的陈述往往不明显，因为它们可能指的是惯用模型，而这种模型大半是以隐含方式定义的。例如，如果一个测量员说他测量了一段距离，那么他指的是抽象了的两个客体并且被认为是两几何点。甚至他可能不是指直接几何意义上的距离，而是指距离在平

面上或者甚至在椭球面上的投影。同样，在A点上的B点和C点之间的角度CAB，通常不是在平面ABC上的角度，而是投影到另一个平面上的角度，例如在水平面（XY）上的角度。

在一定的领域内，教育的任务，也许是阐明用于各种不同任务的各不同工作模型。知道在某些情况下用某些模型来工作，在另一些情况下要建立新的模型，这是科学家或工程师的专门技术的一部分。

就予定要达到的目的来说，函数模型应当以足够的精度与客观实际一致。

1.3.2 模型的有关观测 一个函数模型是一个完全虚拟的结构，用来以适合于分析的明确方式描述实际事件。函数模型通过测量或观测与客观实际发生联系，而测量或观测的本身是实际操作。在比较简单的情况下，测量至少是直接指函数模型的某些元素。但是，模型的全部元素没有必要都是可观测的，而且这样的要求往往是不实际的。在牵涉更广的情况下，测量实际上并不与所讨论的模型的元素直接发生关系。例如，电磁波测距实际上是测定时间或时间差，而不是测量距离。这里，当测量结果和模型联系时，就牵涉到更多的理论要与模型相结合。事实上，附加的变量必须包括进去，而且增加了新的函数关系，这样，就把模型概念扩展到显然很简单的测距任务之外了。因此，为了使测量结果与模型元素发生联系，就必须扩展模型。

由于测量及其性质所引起的函数模型的扩大，具有根本的重要性。观测结果的评价，取决于这些结果是用什么仪器和方法取得的，又是如何取得的。测量的线长在很大程度上取决于测量过程，取决于是否知道了仪器的校准情况，还取决于加入哪些改正，等等。这也适用于角度方向观测的情况。例如，我们必须考虑零方向，并确定它在模型中的位置，把它作为已知量还是作为未知量。

因此，在测量结果能够与模型元素相结合之前的过程，也许是一个相当长的过程。由测量所得的直接读数在可以认为适用之前，往往需要归算或预处理(这就要补充一个改正模型)。

在观测结果与模型发生联系时，为了简单起见，观测结果的很多特征往往被舍弃了。例如，获取实际观测结果所在的特殊情况很少得到考虑。按照函数模型的扩大，这一扩大部分模型的建立简化了。但是，当函数模型变化时，随机模型也必须相应地改变。

1.3.3 随机模型——测量结果的统计性质 根据作业经验，熟练的观测员知道测量结果经常处于某些不可理解的影响的支配下。它们可能处于不能完全控制的一些自然影响的支配之下，当观测重复进行时，导致测量结果的某些变化。测量结果的变化可以归因于某些明显的自然原因，也可能不是如此。这种统计性的变化，不论是起因于忽略了自然影响或随时间变化的影响，或者是起因于自然过程更多的固有特性，都是测量结果的基本性质。

过去，这些变化被说成是起因于观测误差。现在我们则认为测量结果的可变性或随机性是观测结果的主要特性，为了解释起见，我们提到一些统计学概念。

从实用观点来说，评定观测结果的统计性质是有些困难的。一种方法是进行重复观测，由此导出所需要的特性，但通常是不够的。另一种方法是常用的，就是以过去在同样

条件下得出的类似观测结果作为一般性参考，来取得统计特性。因此，当实施测量时，应当记录测量期间有关的全部自然条件和外界环境，以便适当地判断测量结果。事实上，在实践中我们采纳的观测结果的统计特性往往是有些粗略的。例如，在大地测量中，观测通常被认为在统计上是独立的，而且是等精度（等权）的。在摄影测量中，对于像点坐标也是这样认为的，虽然我们知道，底片的收缩会引起像点之间的（物理）相关性。我们假定象点坐标独立，主要是因为这种假定简单，同时也是由于迄今在决定相关性方面遇到了实际的困难。

所涉及到的各变量的统计特性的全部假定，称为随机模型。它包括全部模型变量，并指定被认为是固定的那些量（平差时的常数，是已知的或先前确定的）以及被认为是自由的量(即平差中的待定参数)。

最小二乘平差的理论经典没有明显地确立随机模型的概念，而是使用观测误差或观测误差的特性这些术语。对现在成熟的科技工作者来说，“测量”或“观测”这样的术语是不含糊的，并被用于很广泛的意义上。在本书中，凡是有先验估计的任何随机变量，都称为“观测”。这样的估值可根据直接的或以前的测量结果或其它的方式推导而得。

变量选定的例子：

1. 在三角高程测量中，折射系数可以作为常数来处理，也可以作为由平差来求定的自由参数，或者作为随机变量(具有“标准差”)。在后一情况下，它被认为是一个观测，因而作为观测来处理。

2. 在与一组数据相拟合的多项式中，一个或更多个系数可以作为是观测（具有“期望”或均值零和给定的标准差）来处理。这种作法往往可以减小不定解的可能性。

如上所述，在观测和平差理论中，我们所探讨的随机变量的处理，需要有统计概念的知识。因此，在第一部分的末尾，特别是在第二章，专门讨论对平差有用的统计学方面的问题。这些问题的探讨必然是扼要的，读者如要详细学习和加深的理解，可以参考统计学课程(见本书后的文献目录)。

第一部分的第三章对观测结果的特性作了详细讨论。并试图把经典误差理论和近代观测理论联系起来。第四章将包括从一组随机变量到另一组随机变量的传播原理和方法。

第二章 统计学概念摘要

2.1 引　　言

为了理解本书将要探讨的观测理论，这一章将简要地叙述所必须的一些统计学概念。就最小二乘法的入门基础以及其评价和实际应用来说，这些概念也是必要的。所要介绍的这些概念是摘自概率论和统计学这两个有关数学领域的。简单地说，概率包含关于实验结果机遇的各种定律。概率论则探讨相继发生的或同时发生的大量现象的平均情况。概率论试图描述和推估当试验重复许多次时将会产生的平均情况。另一方面，统计学所探讨的是应用概率的定律，根据已知的一组观测数据，以求得估值或作出推断。因此，统计学是从事收集，分析和解释数据的各种方法的研究。这两个学科用来帮助试验者在面临不定性时作出适当的判断。这两个学科得出了一些准则，根据这些准则可以进行判断。但在实际上，它们从来不直接导致作出判断。

2.2 概率、分布和密度函数

2.2.1 概率　概率的经典定义把一个事件的概率的概念与重复试验时该事件出现的频率联系起来；这种重复也许是真实的，也许是假设的。由这样的联系导出了概率的定义，就是当重复试验次数趋近于无限时，事件出现的频率的极根。例如，如果用一个骰子投掷 6 的相对频率倾向于 1/6（在投掷的次数 n 趋近∞时），那么用骰子投掷 6 的概率就认为是 1/6 或 0.166。

在现代统计学中，把概率作为是期望的频率的这个概念，已不再通用了。现在把概率认为是与统计事件相联系的独立基本概念，它的性质是建立在公理上的。为了理解这个概念，我们首先介绍随机变量的意义。

2.2.2 随机变量　概率是与统计事件相联系的，无论是真实事件或假设事件，一个事件是一次统计试验的结果（例如投掷骰子，观测角度或计算不合格的产品件数）。如果一个统计事件有几种可能的结果，我们就把该事件与随机变量 $\tilde{x}$ 联系起来。随机变量可以定义为这样的一个变量，它取几个可能的值，每一个值都结合一个概率。

在概率中，我们通常是根据一个具有特定参数的已知数学模型来求一个系统的特性。需要研究的各元素的全体称为总体，而我们所探求的正是有关该总体特性的信息。在理论意义上，总体被假定为含有无限多的观测，但实际上总体包含的观测数虽然可能很多，却总是有限的。总体包括所研究的随机变量全部可能的取值。换言之，总体就是与随机变量结合的统计事件的一切可能结果的总和。由于总体所包括的元素很多，研究每一元素来评

定总体特征．这是不可能的，或者是完全不现实的。因此，为了研究总体，我们从其中抽出一定数目的观测，这些抽出的观测称为样本。根据样本研究的结果，我们就可以对抽取样本的总体作出推断和陈述。

抽样方法及其大小将会影响所要作出的结论。显然，样本越大，我们就在一定程度上对所得的结果更加置信。就抽样方法来说，我们必须注意，抽样不要遵循某种规则的模式。当发生这种情况时，样本元素有呈现系统影响的危险，而样本的扩大会导致总体可能不是完全正确的。为了减小这样的困难，必须确保样本是随机抽出来的，亦即总体的每一个元素被抽选到样本中的机会是相等的。也就是说，样本每一个元素的抽选与每一其它的抽选是互相独立的。

随机变量$\tilde{x}$的全组可能值，连同这些值的概率，组成与该随机变量相联系的所谓概率分布。因此，一种概率分布就是描述与一随机变量的各个可能值有关的各种概率。这些分布有两种典型：累积分布函数和概率密度函数。

2.2.3 累积分布函数 随机变量$\tilde{x}$的累积分布函数$F(x)$由下列关系式来定义：

$$P(\tilde{x} \leqslant x) = F(x) \tag{2-1}$$

亦即随机变量$\tilde{x}$的取值小于或等于x（x是一个流动变量）的概率定义为$F(x)$或累积分布函数。因为按公认的定义，概率限制在值0和+1之间(即：$0 \leqslant P \leqslant +1$)，故分布函数$F(x)$满足下面的边缘条件：

$$\lim_{x\to\infty} F(x) = 0, \qquad \lim_{x\to\infty} F(x) = +1 \tag{2-2}$$

虽然方程（2-1）中的基本定义对于离散函数和连续函数都是成立的，但本书主要探讨连续分布函数。

2.2.4 概率密度函数 为了阐明概率密度函数$f(x)$，采取与物理学中密度概念类似的方式，这个函数可以定义为一个间隔Δx的概率。取连续的可微函数，$F(x)$和$f(x)$之间的关系是：

$$F(x) = \int_{-\infty}^{x} f(r)dr \tag{2-3}$$

和

$$f(x) = \frac{\partial F(x)}{\partial x} \tag{2-4}$$

根据这两个关系式，我们易于导出如下的陈述：

1. 一个随机变量$\tilde{x}$在间隔x_1和x_2（此处$x_2 > x_1$）之间取值的概率由下式给定：

$$P(x_1 < \tilde{x} < x_2) = F(x_2) - F(x_1) = \int_{x_1}^{x_2} f(r)dr \tag{2-5}$$

这个概率在图2-1中用晕线所示的面积表示。

2. 一个随机变量$\tilde{x}$取值小于x_1的概率由下式给定：

$$P(-\infty < \tilde{x} < x_1) = F(x_1) - F(-\infty) = F(x_1) = \int_{-\infty}^{x_1} f(r)dr \tag{2-6}$$

注意：按照方程(2-2)，$F(-\infty) = 0$．

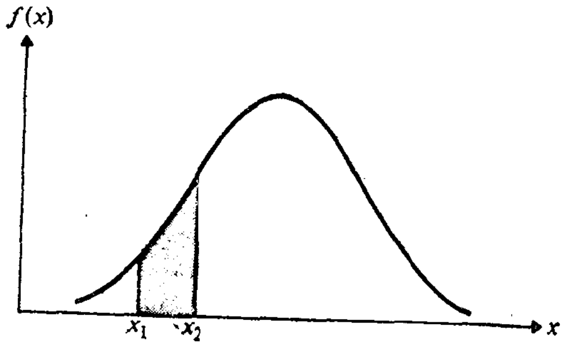


图 2-1

3. 与（2）类似，$\tilde{x}$取值大于x_1的概率为

$$P(x_1 < \tilde{x} < \infty) = F(\infty) - F(x_1) = 1 - F(x_1) = \int_{x_1}^{\infty} f(r)dr. \tag{2-7}$$

注意：按照方程(2-2)，$F(\infty) = 1$.

根据以上的探讨，$f(x)$表征为概率密度函数的两个条件为：

1. 就所有的x值，有$f(x) \geq 0$.

2. $\int_{-\infty}^{\infty} f(x)dx = 1$，由于$x$已经隐含有无限值，又由于概率的总值在这种情况下是密度曲线下面的总面积，故必须等于一。

一个函数必须严格地满足上面给出的两个准则（1）和（2），才可以认为是一个概率密度函数。

探讨一个以上随机变量的情况之前，必须提到，概率密度函数表达整个的总体。这个函数通常用若干个变量来表示，这些变量在统计学术语中称为参数。知道这些参数，就全部确定了密度函数。与此相反，从总体抽得的样本值通常用来计算所谓统计量。每一个统计量通常表示相应的总体参数的估值。

2.3 多维分布、边缘分布和条件分布、独立性

2.2节中所介绍的概念，只包含一个随机变量，在许多实践中，我们常常遇到多于一个随机变量的情况。例如，在两个随机变量$\tilde{x}$和$\tilde{y}$的情形下，就有二维累积分布函数$F(x, y)$。$\tilde{x}$取值小于x和$\tilde{y}$取值小于y（x和y都是流动变量）联合概率由下式给定：

$$F(x, y) = P(\tilde{x} < x, \tilde{y} < y). \tag{2-8}$$

按类似于一维情形的方式，如果$f(x,y)$指的是二维概率密度函数，取连续可微函数，则有

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} \tag{2-9}$$

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)dudv \tag{2-10}$$

如前所述，由于 $0 \leqslant P \leqslant 1$ 是概率值的容许范围，故有

$$\lim_{\substack{x \to -\infty \\ y \to -\infty}} F(x,y) = 0 \qquad \lim_{\substack{x \to +\infty \\ y \to +\infty}} F(x,y) = +1 \tag{2-11}$$

最后，两个随机变量 $\tilde{x}$ 和 $\tilde{y}$ 分别在 x_1 与 x_2 和 y_1 与 y_2 之间取值的概率由下式给定：

$$P(x_1 < \tilde{x} < x_2, \quad y_1 < \tilde{y} < y_2) = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(u,v)dudv \tag{2-12}$$

二维联合概率分布的几何表达，可按类似于一维的方式（图 2-2）来实现。$\tilde{x}$ 和 $\tilde{y}$ 分别位于区间 $x_1 \to x_2$ 和 $y_1 \to y_2$ 内的概率，等于由曲面 $f(x,y)$，平面 x-y 和四个垂直平面 $x = x_1$，$x = x_2$，$y = y_1$，$y = y_2$ 所构成的体积。

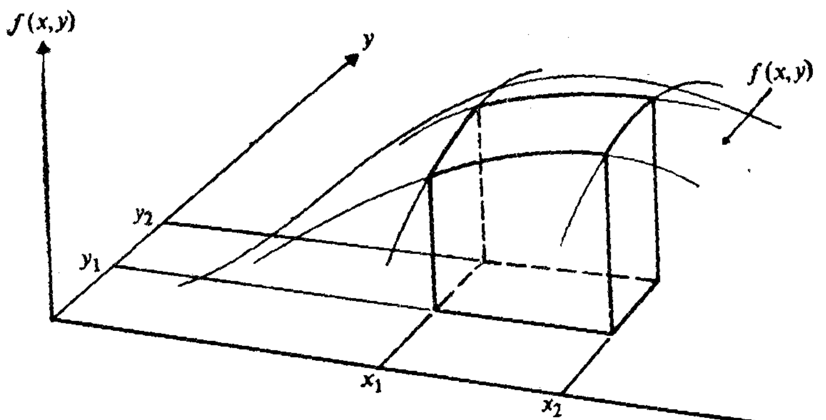


图 2-2

就 n 维的情形，几个随机变量集合成一个随机向量 $\tilde{\mathbf{x}}$，或

$$\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2 \cdots \tilde{x}_n]' \tag{2-13}$$

这个随机向量的 n 维联合密度分布函数 $f(\mathbf{x})$ 为

$$f(\mathbf{x}) = f(x_1, x_2, \cdots x_n), \tag{2-14}$$

相应的累积分布函数 $F(\mathbf{x})$ 为

$$F(\mathbf{x}) = F(x_1, x_2, \cdots x_n). \tag{2-15}$$

随机向量 $\tilde{\mathbf{x}}$ 取值小于 $\mathbf{x}$ 的概率由下式给定：

$$\begin{aligned} F(\mathbf{x}) &= P(\tilde{\mathbf{x}} < \mathbf{x}) = P(\tilde{x}_1 < x_1, \tilde{x}_2 < x_2, \cdots \tilde{x}_n < x_n) \\ &= \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(u_1, u_2 \cdots u_n) du_1 . du_2 \cdots du_n \end{aligned} \tag{2-16}$$

虽然这些关系式所表示的一维情形的推广是比较简捷的，但二维和多维情形所涉及的一些新的概念是很不相同的。这些概念包含边缘分布、条件分布、独立性(和相关性)，所有这些概念只适用于一维以上的情形。

2.3.1 边缘分布 边缘分布是从 n 维分布（$n \geqslant 2$）求得的，但不考虑随机向量 $\tilde{\mathbf{x}}$ 的一个或更多个分量的分布。例如，$\tilde{x}$ 和 $\tilde{y}$ 的二维联合分布，如不考虑其与 $\tilde{y}$ 的关系，就可以把它化为 $\tilde{x}$ 的一维边缘分布。化为 $\tilde{x}$ 的一维边缘分布的方法，是在方程（2-8）或（2-10）累积分布函数中，把 $\tilde{y}$ 的限值选取为 $y = +\infty$。因此，随机变量 $\tilde{x}$ 的边缘累积

分布函数就给定为$\tilde{x}$取值小于x❶和$\tilde{y}$取值小于$+\infty$的联合概率，或

$$F_m(x)=F(x_1\infty)=P(\tilde{x}<x,\tilde{y}<\infty)=\int_{-\infty}^{x}\int_{-\infty}^{\infty}f(u,v)\,du\,dv \tag{2-17}$$

联合密度分布函数$f(x,y)$化为边缘形式

$$f_m(x)=\int_{-\infty}^{\infty}f(u,v)\,dv \tag{2-18}$$

由此得出$\tilde{x}$的边缘累积分布函数为

$$F_m(x)=\int_{-\infty}^{x}f_m(x)\,dx=P(\tilde{x}<x). \tag{2-19}$$

推广到n维的情形，易于看出，对于随机变量全部可能的组合$1,2,\cdots$和$(n-1)$，都存在边缘分布。

2.3.2 独立性　设$F(x,y)$表示两个随机变量$\tilde{x}$和$\tilde{y}$的联合二维累积分布函数，$F(x)$和$F(y)$分别表示$\tilde{x}$和$\tilde{y}$的边缘累积分布函数。如果

$$F(x,y)=F(x)\cdot F(y), \tag{2-20}$$

则两个随机变量$\tilde{x}$和$\tilde{y}$被说成是独立的。倘若边缘累积分布函数是连续的，（2-20）式也就导出了密度函数的关系。这样，就两个独立的随机变量而言。则有

$$f(x,y)=f(x)\cdot f(y). \tag{2-21}$$

事实上，方程（2-21）中的关系可以直接推广到n维密度函数的情形，或推广到

$$f(x_1,x_2,\cdots x_n)=f(x_1)\cdot f(x_2)\cdots f(x_n). \tag{2-22}$$

两个或更多个随机变量独立的概念，是统计事件独立概念的推论。可以回想，如果两事件的联合事件的概率等于各单独事件概率的乘积，即

$$P(ab)=P(a)\cdot P(b), \tag{2-23}$$

则该两事件被说成是统计独立的。必须区别变量的统计独立（或随机独立）与函数独立。两个变量可以是函数独立，但未必是随机独立。

2.3.3 条件分布　条件分布的概念，可以类似于条件概率的方式来定义。例如，$P(b|a)$读为已知a时b的概率，它是在统计事件a已发生的条件下，统计事件b的概率。仿此，条件分布（$\tilde{y}|\tilde{x}$）（读为已知$\tilde{x}$时$\tilde{y}$的分布）是在随机变量$\tilde{x}$和$\tilde{y}$的二维分布中已知$\tilde{x}$的值时$\tilde{y}$的分布。

已知$\tilde{x}$时$\tilde{y}$的条件密度函数为

$$f(y|x)=\frac{f(x,y)}{f_m(x)} \tag{2-24}$$

独立性和条件分布这两个概念是互相直接联系的。如果两个随机变量$\tilde{x}$和$\tilde{y}$是独立的，则已知$\tilde{x}$时$\tilde{y}$的条件分布对$\tilde{x}$的任意值是相同的，反之也是一样。根据已知的关系，则有

$$f(y|x)=\frac{f(x_1y)}{f_m(x)}=\frac{f_m(x)\cdot f_m(y)}{f_m(x)}=f_m(y)=f(y) \tag{2-25}$$

❶ 原书为x_1——校者注。