应用数理统计

李光久　张金时　和　爱　编著

东南大学出版社

应用数理统计

李光久　张金时　和　爱　编著

东南大学出版社

(苏) 新登字第 012 号

内 容 简 介

本书介绍了数理统计的基本概念、估计理论、假设检验、方差分析、回归分析和正交试验设计等六章内容，取材广泛，条理清晰，深入浅出，注重应用．书后附有全部习题解答．阅读本书的读者需具有微积分、线性代数和概率论的基础知识．

本书可作为工科硕士研究生应用统计课程的教材和工科院校本科生教材，还可供师范院校师生及工程技术、经济管理人员参考．

责任编辑　黄英萍

应 用 数 理 统 计

李光久　张金时　和　爱　编著

*

东南大学出版社出版发行

(南京四牌楼 2 号　邮编 210018)

江苏工学院印刷厂印刷

*

开本 850×1168 毫米 1/32　印张 10.75　字数 270 千

1993 年 12 月第 1 版　1993 年 12 月第 1 次印刷

印数: 1-3000 册

ISBN　7—81023—828-0/O・75

定价：7.50 元

前　言

应用数理统计是通过对科学技术的试验数据和国民经济的统计资料的处理来研究随机现象的规律和特征，以便对被观测的随机现象作出合乎科学的统计推断。应用数理统计的方法已广泛应用于不同领域的各种问题，如人口素质，经济发展，股市行情，商品营销，产品质量，可靠性理论，工农业生产对比试验，医药疗效，安全工程，工程技术的试验研究等等问题。

本书力求体现如下特点：(1) 阐述基本概念、基本原理尽量联系客观实际背景，注意从实例引入理论；(2) 为适应不同层次的读者需要，在内容安排上，既有一定深度，如本书给出了方差分析理论赖以建立的柯赫仑定理的证明，介绍了近代贝叶斯估计理论；又有一定的广度，例如本书对非参数假设检验介绍了多种检验方法；(3) 为使读者容易理解应用数理统计的基本方法，在例题的选择上，既注意到典型性，又注意到广泛性与应用性；(4) 在习题的编选上，体现以计算为主，也兼顾必要的证明题，题目的类型注意多样性；(5) 条理清晰、深入浅出。

本书可作为工科硕士研究生应用统计课程教材，也可以适当选择本书有关内容作为工科本科生数理统计课程教材，还可供师范院校师生及工程技术、经济管理人员参考。

本书第一、二、三、四章以及全书各章的习题、习题答案由李光久编写，第五、六两章由张金时、和　爱编写。全书由李光久主编。

本书在编写过程中参阅了国内外有关专著和教材，得到了钟瑜荪老师的关心和支持. 在出版过程中得到江苏工学院王华冠、吴明新、王维华等同志的支持和帮助，在此一并表示谢意.

编著者

1993 年 6 月

目　录

第一章　数理统计的基本概念

§1　引　　言

数理统计和概率论一样，都是研究大量随机现象规律性的一门数学学科. 但两者研究方法是明显不同的：概率论的研究基于建立随机现象的数学模型，也就是把在实际中能观察到的随机现象的频率的概念，经科学的抽象形成概率的概念，并形成了概率的公理化结构. 在引入随机变量的定义之后，在已知分布函数的前提下，去研究随机变量函数的分布、数字特征以及极限理论，这是演绎推理的研究方法. 而数理统计是以概率论为理论基础，通过试验和收集数据资料，对随机现象所反映出的问题（如随机变量的分布函数、数字特征、相互关系等）进行估计和分析，作出统计推断，这是一种从感性（试验）到理性（推断），从局部（采样）到整体（论证），从特殊（数据）到一般（规律）的归纳推理的研究方法.

作为归纳推理的一个例子，我们来研究产品检验问题. 如能对产品进行逐个检验固然很好，但必然要耗费过多的人力、物力和时间. 如果检验将导致产品的破坏，逐个检验就根本不可能施行. 因此，作为替代检验全部产品的办法，就只有对其中少量产品（统计上叫做一个“样本”）进行检验，并由检验结果作出整批产品（统计上叫做“总体”）的评价. 比如从 10 000 只螺钉中抽出 100 只进行检验，发现 100 只中有 5 只是次品，只要螺钉是

“随机地”被抽取出来，就是说这批螺钉中的每一只是以同等的“机会”被抽到，我们就倾向于推断说这批产品差不多有5%是次品. 显然，这种结论不是绝对可靠的，我们不能说这批螺钉不多不少正好有5%的次品. 但是随着每次抽取100只进行检验的工作多次重复进行，这批螺钉次品的百分比将越来越准确地被显露出来.

类似地为了获得关于铁矿石中含铁量μ的信息，就可以随机挑选n个矿石样品，测定其铁的含量，得到一组n个数字x_1，x_2，…，x_n(n个测量数据)，这组数的平均值$\bar{x}=(x_1+x_2+\cdots+x_n)/n$就可以近似当作$\mu$值.

在工程技术的统计学问题中，我们的要求是合理地安排试验，科学地收集、有效地利用有限的数据资料，最大限度地排除由于资料的不足引起的随机干扰，在概率理论的指导下尽可能地作出准确而可靠的结论.

不同性质的问题要求有不同的方法和技巧，但从收集数据到归纳推理得出问题的结论的步骤多数场合几乎是一样的，这些步骤是：

1. 问题的提出 一开始必需明确问题的性质和研究的目的，在此基础上建立起适当的数学模型.

2. 试验的设计 这一步结合最后一步统计方法的选择共同确定样本的大小n(抽取并检验产品的数目或做试验的次数等等)、试验的物理方法以及数据收集的技巧，目的在于用最少的费用和时间获得最大限度的信息.

3. 试验的施行和数据的收集 这一步应严格按照设计的方案办事.

4. 数据的初步处理 将获得的数据资料进行初步加工，整理成表格形式，绘制成图形(曲线图、条形图、直方图等等)，

也可以计算出一些样本数字特征，如样本值的平均值和散度.

5. 统计推断 最后一步就是在掌握试验提供的数据资料的基础上应用适当的数理统计的理论和方法，对研究对象的未知性质作出合乎科学的论断，使提出的问题得以圆满解决.

§2 总体和样本

数理统计研究的对象是不确定的随机现象，如果考察的对象在类似的情况下可以反复观测，或可以反复进行试验，所能得到的观测值（数据）的全体称为**总体**. 总体中的每个数值称为**个体**. 例如，一批零件的强度的全体是一个总体，而每只零件的强度就是一个个体. 又例如，连续生产出的电子元件的寿命的全体是一个总体，每只元件寿命是一个个体，而且这个总体既包含已测得的寿命值，还包括将来可能测得的寿命值. 零件强度数值全体是有限的，这种总体称为**有限总体**. 元件寿命的数值全体是无限的，称为**无限总体**.

如同概率论中的情形，我们所关心的不只是总体中能有哪些数值，而且更关心这些数值（强度、寿命等）的分布规律，或者说观测到某个可能值的概率或概率密度. 这表明我们可以把总体与一个随机变量 X 可能取值的全体等同起来. 如果表征总体的随机变量 X 的分布函数为 $F(x)$，以后我们就称总体 X 的分布为 $F(x)$.

从总体中抽取若干个个体的过程称为**抽样**（又称取样或采样）. 通过 n 次独立的观测或试验，反复抽样可获得 n 个个体 $X_1, X_2, \cdots, X_n$，这样取得的 $(X_1, X_2, \cdots, X_n)$ 称为总体 X 的一个**样本**（又称**子样**）. 样本中个体的数目 n 称为**样本容量**. 注意到

在重复取样中每个 $X_i(i=1,2,\cdots,n)$ 都是一个随机变量，因此我们可以把样本 $(X_1,X_2,\cdots,X_n)$ 看作为一个 n 维随机变量．在一次抽样中观测到的 $(X_1,X_2,\cdots,X_n)$ 的一组确定的值 $(x_1,x_2,\cdots,x_n)$ 称为**样本观测值**（或**数据**）．样本 $(X_1,X_2,\cdots,X_n)$ 所有可能取值的全体（它是 n 维空间或其中的一个子集）称为**样本空间**．一个样本观测值 $(x_1,x_2,\cdots,x_n)$ 就是样本空间中的一个点．

例如，一批只有正、次品之分的产品，如果用数字 0、1 分别表示次品、正品，那末总体 X 是仅含 0 与 1 两个数值的有限总体．现在用有放回抽样的方法，随机抽取 $n=5$ 个产品，样本 $(X_1,X_2,\cdots,X_5)$ 的所有可能取值的全体是由 5 维向量 $(x_1,x_2,\cdots,x_5)$ 组成的样本空间，由于 x_i 只能取 0 或 1，因此样本空间由 $2^5=32$ 个点组成．比如一次抽样得到的观测值为 (1, 1, 0, 1, 0)，它就是样本空间中的一个点．

我们抽取样本的目的是为了对总体的分布进行各种分析推断，因而希望抽取的样本能很好地反映总体的特性，这就需要对随机抽样方法提出一些要求．通常提出下面两点要求：(1) 代表性：要求样本的每个分量 X_i 与总体 X 具有相同的分布 $F(x)$；(2) 独立性：$X_1,X_2,\cdots,X_n$ 为相互独立的随机变量，即每个 X_i 取得什么观测结果，彼此之间互不影响．满足这样两点要求的样本称为**简单随机样本**．获得简单随机样本的抽样方法称为简单随机抽样．在后面提到样本总理解为简单随机样本．显然，如果总体 X 的分布函数为 $F(x)$，那末简单随机样本 $(X_1,X_2,\cdots,X_n)$ 的联合分布函数就是 $\prod_{i=1}^{n}F(x_i)$.

对于有限总体，采用有放回抽样方法得到的样本就是简单随机样本．如果有限总体由 N 个值组成，一次抽取容量为 n 的样本的取法有 C_N^n 种，每种被取到的概率都一样时，这种样本也称为简单随机样本．

抽取这种样本的方法可用随机数表（见附表 Ⅰ）．例如有 500 袋包装食品，为了调查包装食品的平均重量，只取出 30 袋测它们的重量，用它们的平均重量来代替全体的平均值，可将 500 袋先标上 1～500 的编号，然后从随机数表中每 3 个数字组成一个数，直到取的数有 30 个小于等于 500 的不同的数为止，对应这 30 个编号的袋子就是要取的样本．随机数表构造不同，选用的方法也是多种多样的．

对于一个总体，通过随机抽样得到的数据总带有总体特性的信息，而数据一般是按试验的先后次序被记录下来．例如，对热镀铅合金冷轧碳素薄钢板 08A1 试样进行抗拉强度试验，采集到容量 $n=100$ 的样本观测数据如表 1-1 所示(单位 MPa)．

表 1-1

320	380	340	410	380	340	360	350	320	370
350	340	350	360	370	350	380	370	300	420
370	390	390	440	330	390	330	360	400	370
320	350	360	340	340	350	350	390	380	340
400	360	350	390	400	350	360	340	370	420
420	400	350	370	330	320	390	380	400	370
390	330	360	380	350	330	360	300	360	360
360	390	350	370	370	350	390	370	370	340
370	400	360	350	380	380	360	340	330	370
340	360	390	400	370	410	360	400	340	360

从表中容易找出数据最小值是 300，最大值是 440. 将数据按大小次序排列，并将各个数值出现的频数与频率整理出来，就得到样本的频数、频率表如表 1-2.

表 1-2

σ_b (MPa)	频数		频率	积累频数	积累频率
	记码	频数			
300	丅	2	0.002	2	0.02
310		0	0.00	2	0.02
320	正	4	0.04	6	0.06
330	正 一	6	0.06	12	0.12
340	正正 一	11	0.11	23	0.23
350	正正正	14	0.14	37	0.37
360	正正正 一	16	0.16	53	0.53
370	正正正	15	0.15	68	0.68
380	正丅	8	0.08	76	0.76
390	正正	10	0.10	86	0.86
400	正丅	8	0.08	94	0.94
410	丅	2	0.02	96	0.96
420	下	3	0.03	99	0.99
430		0	0.00	99	0.99
440		1	0.01	100	1.00

表 1-2 中情况就比较清晰地显示出样本的统计特性，数据分布中间多、两头少，比如有 37 个样本值小于等于 350，有 76%的样本值小于等于 380.

一般地，设容量为 n 的样本值中总共有 m 个不同的数值，

将它们按由小到大的次序排列为

$$x_{(1)}, x_{(2)}, \cdots, x_{(m)} \qquad (m \leqslant n)$$

对应的频率为

$$f_1^*, f_2^*, \cdots, f_m^*$$

引入函数

$$f^*(x) = \begin{cases} f_j^* & \text{，当} x = x_{(j)} \ (j = 1,2,\cdots,m) \\ 0 & \text{，对不出现在样本值中的其它的} x. \end{cases} \tag{1.1}$$

这个函数称为**样本频率函数**. $f^*(x)$ 决定了样本的频率分布.

定义函数 $F_n^*(x)$: 对任何实数 x,

$$F_n^*(x) = \sum_{t \leqslant x} f^*(t) \tag{1.2}$$

这个函数称为**样本分布函数**，大多数统计著作称 $F_n^*(x)$ 为总体 X 的**经验分布函数**.

可以证明，经验分布函数具有如下性质：

(1) $0 \leqslant F_n^*(x) \leqslant 1$;

(2) $F_n^*(x)$ 是单调不减函数;

(3) $F_n^*(x)$ 处处右连续.

设样本观测值为 $(x_1, x_2, \cdots, x_n)$，还可以这样来定义 $F_n^*(x)$: 对任何实数 x

$$F_n^*(x) = \frac{1}{n} \sum_{i=1}^{n} \varepsilon(x - x_i) \tag{1.3}$$

这里函数 $\varepsilon(x)$ 是

$$\varepsilon(x) = \begin{cases} 0, & \text{当} x < 0 \\ 1, & \text{当} x \geqslant 0 \end{cases}$$

对于上面给出的例子，根据作出的频数、频率表，即可绘出样本频率函数 $f^*(x)$ 与样本分布函数 $F_n^*(x)$ 的图形，如图 1－1 所示．

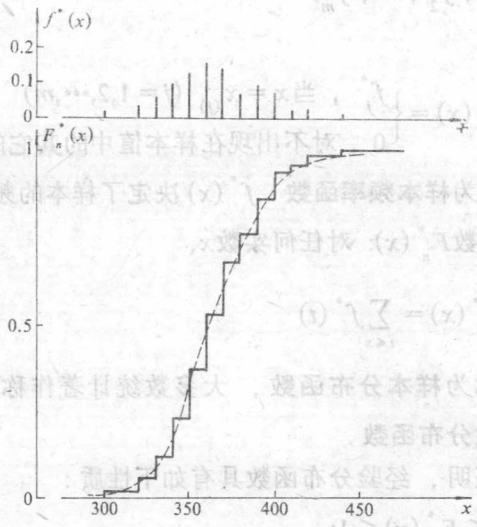

图 1-1

容易看出，$F_n^*(x)$ 是一个离散型分布函数，它只有有限个间断点，因此样本的各阶矩都是有限的．设样本观测值为 $(x_1, x_2, \cdots, x_n)$，则有样本均值

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \int_{-\infty}^{+\infty} x \mathrm{d}F_n^*(x)$$

样本方差

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \int_{-\infty}^{+\infty}(x-\overline{x})^2 \mathrm{d}F_n^*(x)$$

以及样本k阶原点矩与k阶中心矩

$$a_k = \frac{1}{n}\sum_{i=1}^{n}x_i^k = \int_{-\infty}^{+\infty}x^k \mathrm{d}F_n^*(x) \quad (k=2,3,\cdots)$$

$$b_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^k = \int_{-\infty}^{+\infty}(x-\overline{x})^k \mathrm{d}F_n^*(x)$$

$$(k=3,4,\cdots)$$

最后，我们来研究样本分布函数$F_n^*(x)$与总体X的分布函数$F(x)$的关系．设样本为$(X_1,X_2,\cdots,X_n)$，对任何实数x，实际上

$$P\{F_n^*(x)=k/n\}=P\{\text{恰有}k\text{个}X_i,\ X_i\leqslant x\} \tag{1.4}$$

注意到$X_i(i=1,2,\cdots,n)$与X有相同的分布函数，因此事件“X_i落在$(-\infty,x]$中”的概率为$F(x)$，而n次独立抽样可看作n重贝努里试验，从而

$$P\{F_n^*(x)=k/n\}=C_n^k[F(x)]^k[1-F(x)]^{n-k} \tag{1.5}$$

根据贝努里大数定律，当$n\to\infty$时，$F_n^*(x)$依概率收敛于$F(x)$，即对任意给定的$\varepsilon>0$，有

$$\lim_{n\to\infty}P\left(\left|F_n^*(x)-F(x)\right|<\varepsilon\right)=1 \tag{1.6}$$

更深刻的结果由格列汶科(Гливенко)于1933年作出的．

格列汶科定理　设总体X的分布函数为$F(x)$，样本分布函数为$F_n^*(x)$，对任何实数x，有

$$P\left(\lim_{n\to\infty}\sup_{-\infty<x<+\infty}\left|F_n^*(x)-F(x)\right|=0\right)=1 \tag{1.7}$$

由此可见，当 n 充分大时，样本分布函数 $F_n^*(x)$ 是总体分布函数 $F(x)$ 的一个很好的近似．数理统计学中一切都以样本为依据，其理由就在于此．图 1－1 中除了画出样本分布函数 $F_{100}^*(x)$ 外，还用虚线画出与其相应的正态分布函数 $F(x)$.

§3 统计量及其分布

通过随机抽样所得到的样本包含有总体的各种信息，样本是总体的代表和反映．但是样本的 n 个分量所携带的信息是零星和分散的，还不能直接用于对总体的各种特性的研究，这就需要对样本所含的信息进行数学上的加工和制作，使得反映总体某一方面的信息更加集中和凝炼，这在数理统计学中常常是针对不同问题通过构造合适的样本函数——统计量来达到我们的目的．

定义 1 设 $(X_1,X_2,\cdots,X_n)$ 为总体 X 的一个样本，若 $\Phi(x_1,x_2,\cdots,x_n)$ 是一个不含有任何未知参数的 n 元实值函数，则称 $\Phi(X_1,X_2,\cdots,X_n)$ 是一个**统计量**．

由定义可以看到，统计量是随机变量 $X_1,X_2,\cdots,X_n$ 的函数，因此统计量 $\Phi(X_1,X_2,\cdots,X_n)$ 也是随机变量，它应有确定的概率分布，统计量的分布称为**抽样分布**．

从后面的统计推断过程就能体会到统计量不含有任何未知参数的重要性．值得注意的是，虽然一个统计量不含有任何未知参数，但是它的分布也有可能含有未知参数．