高等学校教材

数理统计

西北工业大学概率统计教研室数理统计编写组　编

西北工业大学出版社

高等学校教材

数理统计

西北工业大学概率统计教研室数理统计编写组

朱燕堂　王朝杰
赵选民　秦超英　编

西北工业大学出版社

1990年6月　西安

内 容 简 介

本书比较系统地介绍了数理统计的基本概念、原理和方法。全书共分八章，内容包括数理统计的基本概念、抽样分布，参数估计，假设检验，方差分析，回归分析，贝叶斯估计，稳健估计和试验设计，多元分析，抽样调查简介等，书后附有概率论基础知识。

本书叙述清楚，推理严谨，内容较为丰富，适用面广，各章均配有适量例题和习题并附有习题答案，便于自学和教学。

本书可作为高等院校工科各专业研究生和高年级学生统计课程的教材，也可供理科、经济、管理，师范院校等专业的研究生和大学生及科学技术工作者参考。

高等学校教材

数理统计

编者　西北工业大学概率统计教研室数理统计编写组

责任编辑　雷　鹏

责任校对　钱伟峰

西北工业大学出版社出版

（西安市友谊西路127号）

陕西省新华书店发行

西北工业大学出版社印刷厂印装

ISBN 7-5612-0242-3/O·26(课)

开本 787×1092 毫米　1/16　15.5 印张　372 千字

1990 年 5 月第 1 版　1990 年 5 月第 1 次印刷

印数 1—1 400 册　定价：3.12 元

前　言

人们在日常生活中往往把统计理解为收集数据，画图，列表或对数据作一些简单运算，如求和、求平均数、求百分比、求方差等。其实，这是一种误解，这些仅仅是数理统计的非主要部分，这部分工作一般称为全局性的统计。如果研究的对象比较庞大，全局性的统计就比较困难。例如，要对某产品（比如螺丝钉）进行质量检查，若逐个检查，不仅费时费工，成本高，而且有时反而不准确。有时研究的对象虽不大，但属破坏性试验的质量检查，如灯泡的寿命试验，炮弹射程试验等，即使能得到全部试验数据，但产品被全部损坏掉了。这样，就需要考虑如何从整批产品中抽出一部分产品进行检查，然后据这些产品的质量情况，分析推断整批产品的质量情况。象这种研究从整体中如何抽出一部分来进行观察和测量，如何由部分的性质来推断整体的品质的问题，就是数理统计讨论的主要问题。

由于数理统计中局部数据是从整体中随机抽取的，所以数理统计是以概率论作为其理论基础的。建立在现代数学和概率论基础上的数理统计学是数学的一个分支学科。研究如何以有效的方式收集、整理和分析受到随机性影响的数据，以对所考察的问题作出推断或者预测，直至为采取决策和行动提供依据和建议。数理统计学可用于种种专门知识领域（物理、化学、工程、生物、经济、社会等等）中有关数据分析的问题，它不是以这些专门知识领域为研究对象，而只处理在数据的收集和分析推断中涉及的与随机性有关的数学问题。然而，用数理统计方法分析随机性数据所得的结论的恰当解释，离不开对所讨论问题的专门知识的了解。

数理统计的内容大致可分为收集数据和统计推断两部分，属于前者的分支学科有抽样技术，试验设计等；属于后者的分支学科有参数估计，假设检验，非参数统计，统计决策函数，相关分析，多元统计分析，贝叶斯统计，时间序列分析，概率统计计算，稳健统计，数据分析等。数理统计用于其它知识领域，又出现了如生物统计，人口统计，气象统计，体育统计，数学地质，计量经济学等分支学科。随着科学技术的发展，新的统计方法不断出现，如投影寻踪（Projection Pursuit，简它为 P.P），刀切法（Jackknife），自助法（Bootstrap）等。

数理统计的应用非常广泛，对工农业生产的发展，医药卫生，社会、经济和自然科学等领域都起了重要的作用。

本书是根据 1987 年在华中理工大学召开的全国工科院校研究生数理统计课程研讨会制定的基本内容和要求而编写的。全书共分八章，前六章介绍数理统计的理论基础和基本方法，内容包括：数理统计的基本概念，抽样分布，参数估计，假设检验，方差分析和回归分析。为了使读者对一些较新的近代统计方法有所了解，本书第七章简单介绍了贝叶斯估计和稳健估计。为使读者对数理统计的其它应用方法有所了解，第八章扼要介绍了试验设计的各种方法，多元分析方法和抽样调查方法。各章均配有适量习题，并附有习题答案。对概率论的基本概念和基本公式本书以附录的形式给出，便于读者查阅。

在编写此书的过程中，考虑到工科硕士学位研究生教学的特点，对数理统计的基本理论

和方法在工科研究生数学基础许可的范围内，力求做到推理严谨，讲述透彻。为了便于学生自学，在基本内容方面力求做到由浅入深，通俗易懂，问题提出明确清楚。

本书可作为工科研究生 40～60 学时的数理统计课程教材，也可作为本科生该课程的教材或教学参考书。

本书的第一章、第七章和第八章的第一节由朱燕堂同志编写，第二章和附录由秦超英同志编写，第三章、第四章和第八章的第二节、第三节由赵选民同志编写，第五章和第六章由王朝杰同志编写，朱燕堂同志设计了编写方案，赵选民同志选编了附表并对全稿进行了统一和整理。在本书的编写过程中，西北工业大学应用数学系概率统计教研室的许多同志给予了大力支持和帮助；本书作为讲义在西北工业大学88级和89级研究生教学中使用，曾提出了许多建设性意见；西北工业大学郑政谋教授担任了本书的审阅工作，提出了许多宝贵意见和建议，在此一并致以衷心的感谢！

由于作者水平有限，书中难免存在缺点和错误，敬请读者不吝赐教。

编　者

1989 年 3 月

目　录

第一章　数理统计的基本概念

§1-1　总体和样本

总体与样本是数理统计的两个重要概念，初学者必须对这两个概念的含义有较透彻的了解。

一、总体和表征总体的随机变量

1．总体

在实际应用中，把研究对象的全体叫做**总体**（或称**母体**）。把每个研究对象叫做**个体**，总体中个体的总数叫做**总体容量**。例如，在考察某批的灯泡的质量时，该批灯泡的全体就是一个总体，其中的每只灯泡就是个体。又如，在考察某个品种小麦的性状时，这块地上种植的小麦的所有植株就是总体，其中每一株小麦就是个体。

但是，在数理统计中，我们并不笼统地研究所关心的对象，而只是对它的数值指标感兴趣。例如，考察灯泡时，我们并不研究它的形状、式样等特征，而只是关心灯泡的寿命、亮度等数值指标的大小。当我们只考察灯泡寿命这项指标时，一批灯泡中的每一只灯泡均有一个确定的寿命值，因此，很自然地，我们应该把所有这些寿命值的全体当成总体，这时，每个灯泡的寿命值就是个体。又如，考察一批零件（如钢筋）的总体，当我们主要关心它的强度时，（当然也可以关心其长度，质量，某化学成分等）那么，这批零件的强度值的全体就是总体，每个零件的强度值就是个体。当考虑小麦植株总体时，主要考察株高、穗长等特征；当考察某批炮弹的质量时，主要考察该批炮弹的重量、初速、穿透率、射程等指标，其总体、个体也可同样表示出来。象这种所要考察的特征称为总体的统计特征。今后所说的总体，就是指总体的统计特征。并用X，Y，Z…表示之。

值得指出，当总体的统计特征不能直接用数值表示时，可以给以数量化，例如可以用“0”和“1”分别表示某产品是合格品和不合格品。其次，所考察的统计特征可以是一维的，也可以是多维的，例如，对于小麦植株总体来说，如果只考察株高X，那么，X是一维特征；如果把株高X、穗长Y共同作为小麦植株的特征来考察，那么所要考察的是二维统计特征(X,Y)。本书主要研究一维特征，而多维特征是多元统计分析的主要研究对象。

2．表征总体的随机变量

任何一个总体，都可用一个随机变量来描述它。总体中所包含的个体数目可以是有限的，也可以是无限的。例如就一批钢筋强度这个总体来说，具有各种强度值的钢筋的比例是按一定规律分布的，即任取一钢筋其强度为某可能值是有一定的概率的，也就是说，这批钢筋强度是一个随机变量。其实这个例子也可以这样来理解：从钢筋这个总体中随机地取出一根钢筋（即一个个体）并测定强度X的值，就可看作一个随机试验，于是强度X作为随机试验中被测量的量就是随机变量。又如灯泡的使用寿命；晶体管的直流放大系数；日平均气温的

度数等都可看成随机变量。象上面这种只把所要考察的那个统计特征叫作表征总体的随机变量。

今后总体的特征一指明，就把总体和表征它的随机变量等同起来，也就是说，凡提到总体就是指一个随机变量，凡提到随机变量就是一个总体，仍用大写字母X，Y，Z…表示总体。随机变量X的分布和数字特征就是总体的分布和数字特征。例如，如果随机变量X服从正态分布，则它所表征的总体就称为正态分布的总体。

二、样本

为了对总体X的某些特征进行研究，就必须对总体进行抽样观察，根据抽样观察所得的结果来对总体进行推断。这种从总体X中抽取若干个体来观察某种数量指标X的取值过程，称为**抽样**（又称**取样，采样**）。这种方法，称为**抽样法**。抽样法的基本思想是从要研究的对象的全体抽取一小部分进行观察和研究，从而对整体进行推断。

在一个总体（例如一批钢筋强度）X中，抽取n个个体$X_1, X_2, \cdots, X_n$（实际上，$X_1, X_2, \cdots, X_n$是所取的n个钢筋强度），这n个个体$X_1, X_2 \cdots, X_n$，称为总体X的一个容量为n的样本（又称子样），通常记为$(X_1, X_2, \cdots, X_n)$。

由于每个$X_i (i=1,2,\cdots,n)$是由总体X中随机取出的，它的取值就在总体可能取值范围中随机取得，所以每个X_i都是一个随机变量，而样本$(X_1, X_2, \cdots X_n)$，则是一个n维随机变量。在一次抽取之后，它们都是具体的数值，记作$(x_1, x_2, \cdots x_n)$，称为$(X_1, X_2, \cdots X_n)$的一个观察值，简称样本值。两次不同抽取得到的样本值（两批n个数据）一般是不相同的。

样本$(X_1, X_2, \cdots, X_n)$所可能取值的全体，称为样本空间，一个样本观察值$(x_1, x_2, \cdots, x_n)$就是样本空间中的一个点。

我们的目的是要根据观察到的样本值$(x_1, x_2, \cdots, x_n)$，对总体X的某些特征进行估计、推断。这就需要对样本的抽取提出一些要求，使之能更好地反映总体的特性，一般提出下列两条：

（1） 独立性。因为独立观察是一种最简单的观察方法，所以自然要求$X_1, X_2, \cdots X_n$是相互独立的随机变量，这就是说每个观察结果既不影响其它观察结果，也不受其它观察结果的影响。

（2） 代表性。因抽取的样本要能代表总体的特征，所以要求每个$X_i (i=1,2,\cdots,n)$必须与总体X有相同的分布。

凡满足相互独立且与总体有相同分布这两个条件的样本称为**简单随机样本**。今后如不特别声明，凡提到样本，都是指简单随机样本。获得简单随机样本的方法称为**简单随机抽样**。

处理简单随机样本，可以应用概率论中对独立随机变量所建立的许多重要定理。所以，概率论为数理统计的研究提供了必要的理论基础。

怎样才能得到简单随机样本呢？办法很简单，当抽取的样本容量n相对于总体来说是很小时(例如总体为10 000件，抽取$n=50$件)，则连续抽取n个个体就可近似地认为是一个简单随机样本，这是因为抽取的个数很少时，可认为对总体不产生影响或影响很小的缘故。如果是有放回地抽取，则不必要求n相对很小，这样抽得的n个个体也是一个简单随机样本。又如对一物体测量其长度，测量值是一个随机变量，现在重复测量n次得到$X_1, X_2, \cdots, X_n$也是简单随机样本。

综上所述，所谓总体就是指一个随机变量X，所谓总体$F(x)$，就是指分布函数为$F(x)$的一个随机变量X。所谓样本就是n个相互独立且与总体X有相同分布的随机变量$X_i(i=1,2,\cdots,n)$所组成的n维随机变量$(X_1,X_2,\cdots,X_n)$，每一次具体的抽样，所得到的数据就是n维随机变量的值。

还必须注意，样本具有两重性，在具体抽样之前，它是n维随机变量，但一经抽取便是一组确定的具体数值。有时为方便，把样本和样本值统称样本。

若总体的X分布密度和分布函数分别为$\varphi(x)$和$F(x)$，则可推得样本$(X_1,X_2,\cdots X_n)$的分布密度为$\prod_{i=1}^{n}\varphi(x_i)$，分布函数为$\prod_{i=1}^{n}F(x_i)$，其中$\varphi(x_i)$与$F(x_i)$分别是$X_i(i=1,2,\cdots,n)$的分布密度与分布函数。

最后，我们把上面关于总体与样本的讨论用定义和定理的形式表达出来，便于读者有一个更明确的数学概念。

定义 1.1 若随机变量$X_1,X_2,\cdots,X_n$相互独立且每个$X_i(i=1,2,\cdots,n)$与总体X具有相同的概率分布，则称随机变量$X_1,X_2,\cdots,X_n$为来自总体X的容量为n的简单随机样本。（单个X_i称为来自总体X的样品）。若X有分布密度$\varphi(x)$，（或分布函数$F(x)$），也称$(X_1,X_2,\cdots,X_n)$为来自总体$\varphi(x)$（或$F(x)$）的样本。

定理 1.1 若$(X_1,X_2,\cdots,X_n)$为来自总体$\varphi(x)$(或$F(x)$)的样本，则$(X_1,X_2,\cdots,X_n)$具有联合分布密度$\prod_{i=1}^{n}\varphi(x_i)$（或分布函数$\prod_{i=1}^{n}F(x_i)$）。

【例 1】 若总体X服从参数为p的两点分布，即$P(X=1)=p,P(X=0)=q,(p+q=1)$，设$(X_1,X_2,X_3)$是由$X$中抽取的一个简单随机样本。

（1） 写出它的样本空间。

（2） 写出(X_1,X_2,X_3)的概率分布。

解 （1）样本(X_1,X_2,X_3)的观察值(x_1,x_2,x_3)是一个三维向量，其中$x_i=0$或$1(i=1,2,3)$。所以样本空间由下列$2^3=8$个点组成：(0,0,0)，(0,0,1)，(0,1,0)，(0,1,1)，(1,0,0)，(1,0,1)，(1,1,0)，(1,1,1)。

（2） (X_1,X_2,X_3)的概率分布为

$$P\{(X_1,X_2,X_3)=(x_1,x_2,x_3)\}=p^kq^{3-k}$$

其中k为观察值(x_1,x_2,x_3)中1的个数。

【例 2】 设总体X服从正态$N(\mu,\sigma^2)$分布，$(X_1,X_2,\cdots,X_5)$是一样本。

（1） 写出它的样本空间。

（2） 写出$(X_1,X_2,\cdots,X_5)$的分布密度。

解 （1） 样本$(X_1,X_2,\cdots,X_5)$的观察值$(x_1,x_2,\cdots,x_5)$是五维向量，而$x_i\in(-\infty,\infty)$ $i=1,2,\cdots,5$，所以样本空间是所有可能的点$(x_1,x_2,\cdots,x_5)$的全体，即五维欧氏空间。

（2） $(X_1,X_2,\cdots,X_5)$的分布密度为

$$\prod_{i=1}^{5}\varphi(x_i)=\left(\frac{1}{\sqrt{2\pi}\sigma}\right)^5 e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{5}(x_i-\mu)^2}$$

§1-2 统计量和样本矩

一、统计量

样本是总体的代表及反映，但我们抽取样本之后，并不是直接利用样本的 n 个观察值进行推断，而是需要对这些值进行加工、提炼。把样本中所包含的有关我们所关心的事物的信息都集中起来，这便是针对不同的问题构造出样本的某种函数，这种函数在数理统计中称为统计量。我们引入下列定义：

定义 1.2 设 $(X_1, X_2, \cdots, X_n)$ 为总体 X 的一个样本，若 $f(X_1, X_2, \cdots, X_n)$ 为 n 元连续函数，且 f 中不包含任何未知参数，则称 $f(X_1, X_2, \cdots, X)_n$ 为一个**统计量**。如果 $(x_1, x_2, \cdots, x_n)$ 是 $(X_1, X_2, \cdots, X_n)$ 的观察值，则 $f(x_1, x_2, \cdots, x_n)$ 是 $f(X_1, X_2, \cdots, X_n)$ 的一个**观察值**。

如果某个统计量用来估计分布中的未知参数，则称该统计量为一个**估计量**，估计量的观察值称为**估计值**。

例如，设 $(X_1, X_2, \cdots, X_n)$ 是从具有分布密度为 $N(\mu, \sigma^2)$ 的正态总体中抽取的一个样本，其中 μ，σ^2 是未知参数，则 $\frac{1}{n}\sum_{i=1}^{n} X_i - \mu$，$\frac{1}{\sigma}\sum_{i=1}^{n} X_i$ 都不是统计量，因为它们含有未知参数 μ 或 σ^2，而 $\sum_{i=1}^{n} X_i$，$\sum_{i=1}^{n} X_i^2$，$X_1^2 - X_2^2 - \cdots - X_n^2$ 等都是统计量。

从统计量的定义可以看出，由于样本 $(X_1, X_2, \cdots, X_n)$ 是随机变量，作为样本的连续函数的统计量 $f(X_1, X_2, \cdots, X_n)$ 也是随机变量，它应有确定的概率分布，因此统计量也具有两重性。

二、常用统计量——样本矩

为了推断总体的性质，往往从某些数字特征入手，用样本的数字特征去推断总体的相应数字特征，而样本的数字特征或样本矩是如何定义的？它们与总体相应的数字特征有何关系呢？下面介绍之。

定义 1.3 设 $(X_1, X_2, \cdots, X_n)$ 是从总体 X 中抽取的样本，称统计量。

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{为样本均值}$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2 \qquad \text{为样本方差}$$

$$S^{*2} = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \overline{X})^2 \qquad \text{为修正样本方差（简称样本方差）}$$

$$S = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2} \qquad \text{为样本标准差}$$

$$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \quad (k = 1, 2, \cdots) \qquad \text{为样本 } k \text{ 阶原点矩}$$

特别当 $k=1$ 时，$A_1=\bar{X}$。

$B_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^k$ $(k=1,2,\cdots)$ 为**样本 k 阶中心矩**

特别当 $k=2$ 时，$B_2=S^2$。

用 $\bar{x}$，s^2，a_k，b_k 分别表示 $\bar{X}$，S^2，A_k，B_k 的观察值，此时只要把公式中的 X_i 改为 x_i 即可。

由大数定律可知，只要总体的 k 阶矩存在，则样本的 k 阶矩依概率收敛于总体的 k 阶矩。

【例 1】 从一批机器零件毛坯中随机抽取 8 件，测得其重量（单位：kg）为：

230，243，185，240，228，196，246，200；

（1）写出总体，样本，样本值，样本容量；

（2）求样本均值，样本方差，样本二阶原点矩（到小数第二位）的值。

解

（1）总体：本批机器零件毛坯重量 X；

样本：$(X_1,X_2,\cdots,X_8)$；

样本值：（230，243，185，240，228，196，246，200）；

样本容量：$n=8$。

（2）样本均值：$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i=\frac{1}{8}(230+243+\cdots+200)=221\text{kg}$

样本方差：$s^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{8}\sum_{i=1}^{n}(x_i-221)^2$

$$=\frac{1}{8}[9^2+22^2+(-36)^2+19^2+7^2+(-25)^2+25^2+(-21)^2]$$

$$=495(\text{kg})^2$$

样本二阶原点矩：$a_2=\frac{1}{n}\sum_{i=1}^{n}x_i^2=\frac{1}{8}(230^2+243^2+\cdots+200^2)$

$$=49373.75$$

上例中的数据较大，为了简化计算和避免除法舍去误差，常用下式

$$\sum_{i=1}^{n}(x_i-\bar{x})^2=\sum_{i=1}^{n}x_i^2-\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2=\sum_{i=1}^{n}x_i^2-n\bar{x}^2$$

容易看出，当 x_i 皆减去某常数 a 时，$\sum_{i=1}^{n}(x_i-\bar{x})^2$ 值不变，在计算中常用此法，使数据变小些。例如本例可都减去 200，再计算就简单多了。

三、样本矩的期望与方差

在下面的讨论中，我们用 μ 表示总体 X 的数学期望，σ^2 表示总体的方差，a_k 表示总体的 k 阶原点矩，μ_k 表示总体的 k 阶中心矩，即记

$$EX=\mu,\quad DX=E(X-\mu)^2=\sigma^2$$

$$EX^k=a_k,\quad E(X-\mu)^k=\mu_k$$

并且约定，当我们用到它们时，假定这些矩都是存在的。

定理 1.2 设$(X_1,X_2,\cdots,X_n)$是来自总体X的样本，则有

（1） $$E\bar{X}=\mu,\quad D\bar{X}=\frac{1}{n}\sigma^2$$

（2）$\bar{X}$ 的三阶、四阶原点矩分别为

$$E\bar{X}^3=\frac{1}{n^2}[\alpha_3+3(n-1)\alpha_2\mu+(n-1)(n-2)\mu^3]$$

$$E\bar{X}^4=\frac{1}{n^3}[\alpha_4+4(n-1)\alpha_3\mu+6(n-1)(n-2)\alpha_2\mu^2$$

$$+3(n-1)\alpha_2^2+(n-1)(n-2)(n-3)\mu^4]$$

（3）$\bar{X}$ 的三阶、四阶中心矩分别为

$$\mu_3(\bar{X})\triangleq E(\bar{X}-\mu)^3=\frac{1}{n^2}\mu_3=\frac{1}{n^2}E(X-\mu)^3$$

$$\mu_4(\bar{X})\triangleq E(\bar{X}-\mu)^4=\frac{1}{n^3}\mu_4+\frac{3(n-1)\mu_2^2}{n^3}$$

（4）样本方差 S^2 的期望，方差分别为

$$ES^2=\frac{n-1}{n}\sigma^2$$

$$DS^2=\frac{\mu_4-\mu_2^2}{n}-\frac{2(\mu_4-2\mu_2^2)}{n^2}+\frac{\mu_4-3\mu_2^2}{n^3}$$

其中 $\mu_2=\sigma^2$

证明 （1） $$E\bar{X}=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}EX_i=\frac{1}{n}\sum_{i=1}^{n}\mu=\mu$$

$$D(\bar{X})=E(\bar{X}-\mu)^2=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right)^2=\frac{1}{n^2}E\left[\sum_{i=1}^{n}(X_i-\mu)\right]^2$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}E(X_i-\mu)^2=\frac{1}{n}\sigma^2$$

（2）为了求 $E\bar{X}^3$，$E\bar{X}^4$ 只需注意

$$\left(\sum_{j=1}^{n}X_j\right)^3=\sum_{j=1}^{n}X_j^3+3\sum_{\substack{j=1\\j\neq k}}^{n}\sum_{k=1}^{n}X_j^2X_k+\sum_{\substack{j=1\\j\neq k,\,k\neq l}}^{n}\sum_{\substack{k=1\\j\neq l}}^{n}\sum_{l=1}^{n}X_jX_kX_l$$

$$\left(\sum_{i=1}^{n}X_i\right)^4=\left(\sum_{i=1}^{n}X_i\right)\left(\sum_{j=1}^{n}X_j^3+3\sum_{j\neq k}\sum X_j^2X_k+\sum_{j\neq k\neq l}\sum\sum X_jX_kX_l\right.$$

$$=\sum_{i=1}^{n}X_i^4+4\sum_{j\neq k}X_jX_k^3+3\sum_{j\neq k}X_j^2X_k^2+6\sum_{i\neq j\neq k}X_i^2X_jX_k$$

$$+\sum_{i\neq j\neq k\neq l}X_iX_jX_kX_l$$

(3) $$\mu_3(\bar{X})=\frac{1}{n^3}E\Big[\sum_{i=1}^{n}(X_i-\mu)\Big]^3$$

$$=\frac{1}{n^3}\sum_{i=1}^{n}E(X_i-\mu)^3=\frac{1}{n^2}\mu_3$$

$$\mu_4(\bar{X})=\frac{1}{n^4}E\Big[\sum_{i=1}^{n}(X_i-\mu)\Big]^4$$

$$=\frac{1}{n^4}\sum_{i=1}^{n}E(X_i-\mu)^4+C_4^2\frac{1}{n^4}\sum_{i<j}\sum E\{(X_i-\mu)^2(X_j-\mu)^2\}$$

$$=\frac{\mu_4}{n^3}+\frac{3(n-1)}{n^3}\mu_2^2$$

说明：上式中 $E(X_i-\mu)^2E(X_j-\mu)^2=\mu_2\cdot\mu_2=\mu_2^2$

而$\sum_{i=1}^{n}\sum_{j=1}^{n}$有$(n-1)+(n-2)+\cdots+2+1=\frac{n(n-1)}{2}$项，$C_4^2\frac{1}{n^4}\frac{n(n-1)}{2}\mu_2^2=\frac{3(n-1)}{n^3}\mu_2^2$。

(4) $$ES^2=E\Big[\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\Big]=E\Big[\frac{1}{n}\sum_{i=1}^{n}X_i^2-\frac{1}{n^2}\Big(\sum_{i=1}^{n}X_i\Big)^2\Big]$$

$$=\alpha_2-\frac{1}{n^2}E\Big[\sum_{i=1}^{n}X_i^2+\sum_{i\neq j}\sum X_iX_j\Big]$$

$$=\alpha_2-\frac{1}{n^2}[n\alpha_2+n(n-1)\mu^2]$$

$$=\frac{n-1}{n}(\alpha_2-\mu^2)$$

$$=\frac{n-1}{n}\sigma^2$$

至于 DS^2 的证明较繁，从略。有兴趣的读者可参看主要参考书目[1]第8页。

§1-3 顺序统计量和经验分布函数

一、顺序统计量

顺序统计量以及由顺序统计量得来的其它统计量（例如，样本极差和样本中位数）是极为重要的一类统计量。它在近代统计推断中起着重要的作用。因为顺序统计量有一些性质不依赖于总体的分布，并且计算量很小，使用起来较方便。因此，在质量管理、可靠性等方面得到了广泛的应用。我们在此仅扼要地介绍有关顺序统计量的内容。

设$(X_1,X_2,\cdots,X_n)$是从总体X中抽取的容量为n的样本；记$(x_1,x_2,\cdots,x_n)$，是样本的一个观察值，将观察值由小到大按顺序重新编号排列，并记作$x_{(1)},x_{(2)},\cdots x_{(n)}$，其中$x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n)}$，如果任何两个$x_i$，$x_j$相等，则它们的次序任意。当$(X_1,X_2,\cdots,X_n)$取值为$(x_1,x_2,\cdots,x_n)$时，我们定义$X_{(k)}$取值为$x_{(k)}(k=,1,2,\cdots,n)$。由此得到的$(X_{(1)},X_{(2)},\cdots,$

$X_{(n)}$)称为样本($X_1, X_2, \cdots, X_n$)的一组顺序统计量。$x_{(1)}, x_{(2)}, \cdots, x_{(n)}$ 称为顺序统计量的值。显然

$$X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(n)}$$

其中 $X_{(1)} = \min\limits_{1 \leqslant i \leqslant 1} X_i$ 称为极小顺序统计量，即无论样本取得怎样一组观察值，它的观察值总是取其中最小的一个值 $x_{(1)}$。

$X_{(n)} = \max\limits_{1 \leqslant i \leqslant n} X_i$ 称为极大顺序统计量，即它的观察值是无论样本取得怎样一组观察值，总是取其中最大的一个值 $x_{(n)}$。

$X_{(k)}$ 称为第 k 个顺序统计量，它的观察值是无论样本取得怎样一组观察值，总是取其中的 $x_{(k)}$，$x_{(k)}$ 为($X_1, X_2, \cdots, X_n$)的观察值($x_1, x_2, \cdots, x_n$)中按大小顺序排列后的第 k 个数值($k = 1, 2, \cdots, n$)。

由上可见，$X_{(k)}$ 是样本 ($X_1, X_2, \cdots, X_n$)的函数，所以 $X_{(1)}, X_{(2)} \cdots X_{(n)}$ 也都是随机变量。

如果 $X_1, X_2, \cdots, X_n$ 是来自同一总体的 n 个相互独立随机变量，那么顺序统计量 $X_{(1)}$，$X_{(2)}, \cdots, X_{(n)}$ 是否也相互独立呢？这可从下例看出。

【例1】 设(X_1, X_2, X_3)为来自总体X的一个容量为 3 的样本，且设X的分布列为

X	0	1	2
P	1/3	1/3	1/3

现在把样本 X_1, X_2, X_3 与由它们所构成的顺序统计量 $X_{(1)}, X_{(2)}, X_{(3)}$ 的一切可能观测值列于表中。

X_1	X_2	X_3	$X_{(1)}$	$X_{(2)}$	$X_{(3)}$	X_1	X_2	X_3	$X_{(1)}$	$X_{(2)}$	$X_{(3)}$
0	0	0	0	0	0	2	0	1	0	1	2
0	0	1	0	0	1	2	1	0	0	1	2
0	1	0	0	0	1	0	2	2	0	2	2
1	0	0	0	0	1	2	0	2	0	2	2
0	0	2	0	0	2	2	2	0	0	2	2
0	2	0	0	0	2	1	1	1	1	1	1
2	0	0	0	0	2	1	1	2	1	1	2
1	1	0	0	1	1	1	2	1	1	1	2
1	0	1	0	1	1	2	1	1	1	1	2
0	1	1	0	1	1	1	2	2	1	2	2
0	1	2	0	1	2	2	1	2	1	2	2
0	2	1	0	1	2	2	2	1	1	2	2
1	0	2	0	1	2	2	2	2	2	2	2
1	2	0	0	1	2						

由于样本(X_1, X_2, X_3)取到每一组观测值的概率都等于 $\frac{1}{27}$，容易从表中看出

$$P(X_{(1)} = 0) = \frac{19}{27} \neq \frac{9}{13} = P(X_{(1)} = 0 \mid X_{(2)} = 1)$$

因此，一般说来顺序统计量之间是相互不独立的。

下面我们不加证明地写出第 k 个顺序统计量的分布，极大顺序统计量的分布，极小顺序统计量的分布与顺序统计量的联合分布，为方便起见，只写出总体 X 是连续型的情形。

定理 1.3 设总体 X 的分布密度为 $\varphi(x)$，$(X_1, X_2, \cdots, X_n)$ 是 X 的样本，$(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 是它的顺序统计量，那么

（1）第 k 个顺序统计量 $X_{(k)}$ 的分布密度为

$$g_{X_{(k)}}(x) = nC_{n-1}^{k-1}[F(x)]^{k-1}[1-F(x)]^{n-k}\varphi(x)$$

（2）极大顺序统计量 $X_{(n)}$ 的分布密度为

$$g_{X_{(n)}}(x) = n[F(x)]^{n-1}\varphi(x)$$

（3）极小顺序统计量 $X_{(1)}$ 的分布密度为

$$g_{X_{(1)}}(x) = n[1-F(x)]^{n-1}\varphi(x)$$

（4）$(X_{(1)}, X_{(2)}, \cdots, X_{(n)})$ 的联合分布密度为

$$g_{(X_{(1)}, \cdots, X_{(n)})}(x_{(1)}, x_{(2)}, \cdots x_{(n)}) = \begin{cases} n!\prod_{k=1}^{n} f(x_{(k)}) & \text{若 } x_{(1)} < x_{(2)} < \cdots < x_{(n)} \\ 0 & \text{其它} \end{cases}$$

【例 2】 设 $(X_1, X_2, \cdots, X_5)$ 是从具有均匀分布为

$$\varphi(x) = \begin{cases} 1 & \text{当} 0 < x < 1 \\ 0 & \text{其 它} \end{cases}$$

的总体中抽取的样本，则由定理 1.3 得到，顺序统计量 $(X_{(1)}, X_{(2)}, \cdots, X_{(5)})$ 的联合分布密度是

$$g(y_1, y_2, \cdots, y_5) = \begin{cases} 5! & \text{当} 0 < y_1 < y_2 < \cdots < y_5 < 1 \\ 0 & \text{其 它} \end{cases}$$

下面介绍样本中位数和样本极差的定义。

样本中位数定义为

$$\widetilde{X} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{当 } n \text{ 是奇数时} \\ \frac{1}{2}\left[X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right] & \text{当 } n \text{ 是偶数时} \end{cases}$$

它的观察值为

$$\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & \text{当 } n \text{ 是奇数时} \\ \frac{1}{2}\left[x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right] & \text{当 } n \text{ 是偶数时} \end{cases}$$

由定义可知，当 n 为奇数时，正中间的数只有一个，这个数就是中位数；当 n 为偶数时，正中间的数有两个，中位数等于这两个数的算术平均值。因此可以简单地说，样本中位数就是正中间的那一个数值。它是刻划样本的位置特征的量。

样本中位数计算方便，有时比样本均值更具有代表性，因为样本中位数不受极端值的影响。

【例 3】 从总体中抽取容量为 5 的样本，并得样本值为32，55，28，35，30，试求样本中位数和样本均值。

解 将样本值按大小递增次序排列如下：

28, 30, 32, 35, 55

由于 $n=5$，所以

样本中位数 $x_{(\frac{n+1}{2})}=x_{(3)}=32$

样本均值 $\bar{x}=\frac{1}{5}(28+30+32+35+55)=36$

可见，样本均值大于5个数中的4个数，这是因为有一个特别大的数55的缘故。样本均值对极端值或异常值较为敏感，因此，有时估计总体均值用样本中位数比用样本均值效果好。

样本极差定义为

$$R=X_{(n)}-X_{(1)}=\max_{1\leqslant i\leqslant 1}X_i-\min_{1\leqslant i\leqslant 1}X_i$$

它的观察值为

$$r=x_{(n)}-x_{(1)}=\max_{1\leqslant i\leqslant n}x_i-\min_{1\leqslant i\leqslant n}x_i$$

样本极差是极大顺序统计量与极小顺序统计量之差，它与样本方差一样，也是反映观察值离散程度的数字特征，不过计算简单，所以在实际问题中有广泛的应用。特别在质量管理方面，往往用样本极差代替样本方差。

二、经验分布函数

定义 1.4 设 $(X_1,X_2,\cdots,X_n)$ 是来自总体 X 的样本，样本的顺序统计量为 $X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n)}$，当固定一组顺序统计量的观察值 $x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n)}$ 时，对任何实数 x，称下式

$$F_{(n)}(x)=\begin{cases}0 & x\leqslant x_{(1)}\\ \dfrac{k}{n} & x_{(k)}<x\leqslant x_{(k+1)}\quad k=1,2,\cdots,n-1\\ 1 & x>x_{(n)}\end{cases}$$

为总体 X 的经验分布函数。换句话说，对任何实数 x，$F_{(n)}(x)$ 等于诸 x_i 中不超过 x 的个数再除以 n。

经验分布函数的性质：

（1）当 $x_{(1)},x_{(2)},\cdots,x_{(n)}$ 的值固定时，$F_{(n)}(x)$ 是一个分布函数，只能在 $x_{(k)}(k=1,2,\cdots,n)$ 处有间断点，跃度是 $\frac{1}{n}$ 的倍数（如有 l 个元相重，$x_{(k-1)}<x_{(k)}=x_{(k+1)}=\cdots=x_{(k+l-1)}<x_{(k+l)}$，则在间断点 $x_{(k)}$ 上的跃度为 $\frac{l}{n}$）。事实上，当 $x_{(1)},x_{(2)},\cdots,x_{(n)}$ 的值固定时，$F_{(n)}(x)$ 是 x 的函数并且满足分布函数的性质：

① $0\leqslant F_{(n)}(x)\leqslant 1$；

② $F_{(n)}(-\infty)=0$，$F_{(n)}(+\infty)=1$

③ 非减左连续。

（2）$F_{(n)}(x)$ 是随机变量，且服从二项分布。事实上，由于给定一组样本值，就可作一个经验分布函数 $F_{(n)}(x)$，因此 $F_{(n)}(x)$ 是顺序统计量的观察值 $x_{(1)},x_{(2)},\cdots,x_{(n)}$ 的函数，所以，