# 统计推断导引

西安交通大学数学研究生教学丛书

范金城 吴可法 编著

科学出版社

西安交通大学数学研究生教学丛书

# 统 计 推 断 导 引

范金城　吴可法　编著

科 学 出 版 社

2001

# 内 容 简 介

作为研究生教材，本书从数理统计的基本概念出发，较系统地讲述了统计推断的原理、方法和应用，内容包括统计基本知识概述、参数估计、假设检验、区间估计、非参数统计推断、Bayes 统计和统计决策，本书注重统计基本思想的阐述与基本方法的介绍.

读者对象：数学系硕士博士研究生、教师及有关工程技术人员，具有大学本科水平和概率论基础知识的其他读者.

**图书在版编目(CIP)数据**

统计推断导引/范金城，吴可法编著.-北京：科学出版社，2001.8
(西安交通大学数学研究生教学丛书)
ISBN 7-03-008955-3

Ⅰ. 统…　Ⅱ. ①范…②吴…　Ⅲ. 统计推断-研究生-教材　Ⅳ. O212

中国版本图书馆 CIP 数据核字 (2000) 第 81791 号

科学出版社 出版
北京东黄城根北街16号
邮政编码：100717
http://www.sciencep.com
双青印刷厂 印刷

科学出版社发行　各地新华书店经销

*

2001 年 8 月第 一 版　开本：B5(720×1000)
2001 年 8 月第一次印刷　印张：20
印数：1—3 000　字数：367 000

**定价：30.00 元**

(如有印装质量问题，我社负责调换〈环伟〉)

# 序

《统计推断导引》一书是范金城、吴可法两位同志在西安交通大学多年教学的基础上写成的教材.本书既反映了他们的教学经验,也反映了当前工科院校教师、博士生等研究人员需要的统计知识和技术.

本书的特点有作者的自己介绍(见前言),但我认为值得强调的是以下两个方面:

1.系统而严谨地介绍了数理统计的基本概念、基础理论和基本方法.这从书的目录中就可看出,本书作者很周密地安排了估计理论、假设检验理论、置信区间、非参数方法等内容,这比国内一般常见的教材内容更为全面和系统,而且每一部分都能把主要的内容包括在内,所以比较全面地介绍了数理统计的内容.

2.特意介绍了不同的学派,这与常见的教材完全不同,本书将 Bayes 学派专列一章来系统介绍,在区间估计中又专列一节介绍了不同于经典学派的信任区间(Fisher 信任学派的内容),这使学生可以领略不同学派的观点、方法和结论,这对开阔视野、解放思想是很有帮助的.

我很高兴看到这本有特色的教材能够出版,这会促进教学改革,适应各种不同读者的需求.我更希望有更多的、各具特色的教材陆续出版,改变过去那种基本上都相同的书重复刊印的局面,真正达到“百花齐放、百家争鸣”.

张尧庭

2000 年 10 月 18 日

# 前　言

数理统计的任务是研究怎样用有效的方法去收集和使用带随机性影响的数据,从而对所考察的问题作出推断与预测,直至为采取某种决策提供依据与建议.数理统计学的核心内容为统计推断.英国著名统计学家 Fisher 认为,统计推断主要包括抽样分布、参数估计与假设检验.本书介绍统计推断的主要内容,因此定名为《统计推断导引》.本书可以作为概率论与数理统计专业的研究生教材,其基本内容可以作为外专业博士研究生的教材.

本书内容包括统计基本知识概述、参数点估计、参数假设检验、区间估计、非参数统计推断、Bayes 统计推断、统计决策共七章.

本书的主要特点如下:

1. 着重统计基本思想的阐述与统计基本方法的介绍.统计基本思想与方法的引入,是从广泛的应用背景与例题出发,由浅入深,循序渐进.尽力做到既论述严谨,又深入浅出,符合读者的认识规律,也符合教学规律.

2. 着重统计中新方法与不同学派的介绍.对于在现代统计学发展中起重大影响的 Bayes 学派的观点与理论,专设一章"Bayes 统计推断"进行较系统的介绍.又如对稳健性统计做了简单介绍.作为统计推断的概括与总结,最后一章为"统计决策".

3. 每章后进行小结,这对读者进行复习总结会有好处.每章后有适量习题,它是本书的有机组成部分.

本书的用语与符号与统计学国家标准一致.

本书是统计学的基础教程,在学习本书后,读者可以进一步学习统计学各个分支的理论与方法,如回归分析、方差分析、试验设计、抽样技术、可靠性统计、多元统计分析、时间序列分析等.

本书第二、四、六、七章由范金城编写,第一、三、五章由吴可法编写.范金城对全书进行了统编.本书初稿在数年前已完成,并在西安交通大学概率论与数理统计专业研究生中讲授过多次,教学效果颇好.

近年,我国有几本数理统计的著作或研究生教材出版,并在中国统计出版社出版了现代外国统计学优秀著作译丛,足见国家对统计学的重视.

在西安交通大学理学院与西安交通大学研究生院的大力支持下,编著者对初稿又进行了认真修改.

我们的老师张尧庭教授一直关心本书的出版,并为本书作序,在此我们表

示衷心感谢.

限于水平,缺点和错误在所难免,敬请指正.

编著者

2000 年 10 月于西安交通大学

# 目　录

# 第一章　统计基本知识概述

## §1.1　统计学与数理统计

在自然科学、社会科学、工程技术等领域中，人们经常要收集数据、积累资料，然后对其进行整理、计算和分析，这就需要统计学与数理统计方法.

谈到统计学，有人认为只是收集数据，列成图表，计算平均数、百分比等. 实际上，这些仅仅是统计学的初等部分. 统计学是研究数据的收集、整理与分析，并在此基础上作出推断和预测，直至为采取决策、行动提供参考、依据和建议.

人们用观察和实验的方法收集到的数据，大多数受到随机因素的影响. 数理统计学是统计学的一个分支，它的任务是研究怎样用有效的方法去收集和使用带随机性影响的数据，从而对所考察的问题作出一定的结论.

概率论是研究大量随机现象统计规律的数学学科. 数理统计的对象是处理带随机性影响的数据. 是否假定数据有随机性，是数理统计方法和其他数据处理方法的根本区别. 由此看来，概率论是数理统计的理论基础. 数理统计方法的依据是随机现象的统计规律性. 例如，从一大批产品中随机地抽一件检查为正品，并不能断言整批产品质量皆好. 若随机抽查 500 件，发现有 25 件次品，其余是正品，则可依据大数定律断言产品的正品率为 95% 左右. 再如，某工件长度真值为 $a$. 测量时有随机误差 $\varepsilon$，测量的结果为随机变量 $X$，$E(X) = a$，又 $\mathrm{Var}(X) = \sigma^2$. $\mathrm{Var}(X)$ 反映了一次测量的误差的波动性大. 现进行 $n$ 次独立测量，其结果为 $X_1, X_2, \cdots, X_n$，它们相互独立且具与 $X$ 有相同分布. 此时可用平均数 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ 估计 $a$，有

$$E(\bar{X}) = a, \qquad \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n},$$

表明用 $\bar{X}$ 估计 $a$ 无系统误差，而 $\bar{X}$ 的方差只有 $X$ 的方差的 $\frac{1}{n}$. 这表明，用 $\bar{X}$ 去估计 $a$ 确能大大减少随机误差的影响. 从理论上讲，只要对随机现象进行足够多次的观测，所研究的随机现象的统计规律就能清楚地呈现出来. 但是，实际上观测次数要受到限制，有时仅有少量观测. 因此，一个好的数理统计方法就在于能够有效地使用有限的观测数据，而作出尽可能精确、可靠的结论.

数理统计的内容及其分支学科主要有：收集数据方面，包括抽样方法和试验设计；整理数据方面，通常称为描述统计；数据分析与推断方面，这方面总称为统计推断，是数理统计研究的重点．本书定名为《统计推断导引》，其中包括最基本的统计推断内容：参数点估计、区间估计参数假设检验，非参数统计推断等．总之，本书主要阐述统计推断的基本思想与方法．Bayes 学派是数理统计的一大学派，以这一学派的观点构成了一整套 Bayes统计推断方法．Wald 的统计决策理论，使统计推断建立在严密的数学基础之上，且拓广了某些传统分支的研究领域和提出了一些有意义的统计问题．本书对 Bayes 统计推断与统计决策理论各专设一章作简要介绍．本书称为《导引》，是因为它包括数理统计基本理论与方法，是数理统计学的“入门”．在此基础上，可以进一步学习统计学的许多分支，如回归分析、方差分析、多元统计分析、时间序列分析等．

数理统计方法的应用十分广泛．由于随机因素影响无所不在，因而在自然科学、社会科学、工程技术以至工业、农业、国防、经济等各个领域都需要用统计方法．反过来，这些应用又推动了数理统计学的发展．应该注意，用数理统计方法分析随机性数据所得结论的解释，离不开所涉及问题的专门知识，例如计量经济学用到数理统计方法，就要涉及经济学知识．而本书主要阐述统计推断的基本理论与方法．

## §1.2 样本与样本分布

### 1.2.1 样本与总体

通过观察或试验得到的数据，称为**样本**．从某厂生产的某批日光灯管中随机抽取 $n$ 只进行寿命试验，得到数据 $X_1, X_2, \cdots, X_n$（单位：小时），则它们全体的 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^{\mathrm{T}}$ 就称为样本．取得样本是提供所研究问题的有关信息．样本就是所研究对象的一批观测值．观测值的数目 $n$ 称为样本量，或称样本大小．

在统计学中，所研究对象的全体称为**总体**，其中每一个称为**个体**．例如，要研究某块玉米地单株玉米的产量，各单株玉米产量的全体构成总体，而每一单株玉米的产量是一个体．要考察某批日光灯的寿命情况，该批日光灯寿命的全体构成总体，每只日光灯的寿命是一个体．统计学所研究的对象，往往是一种数量指标．如上述我们所考察的是单株玉米的产量，而不是一株株玉米本身；是每只日光灯的寿命，而不是一只只日光灯本身．有时所研究的是某种属性特征．例如，考察某厂生产的某种电子元件，检查其中每一只为合格品，还是次品．这时，可以把这种属性特征数量化．如可用“0”表示合格品，“1”表示次品．

这样,样本就是由“0”和“1”组成的 $n$ 个数据.如果研究的问题不同,那么总体也会不同.但在每一研究问题中,总体必须明确.

样本就是按一定程序从总体中随机抽取的一组(一个或多个)个体.我们要根据样本的情况,来推断总体的情况.从哲学观点看来,这就是从局部的情况,来推断整体的情况.这正是统计推断的基本思想.从总体中抽取样本的过程称为抽样.统计学中,抽取每一个体必须具有随机性,因此,统计学中的抽样皆指随机抽样.

从实用的观点看来,样本就是一批已知的数据.但是,数理统计所处理的是含随机性影响的数据.由于样本 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 由总体中随机抽样得到,其取值具有随机性.从概率论观点看来,样本是 $n$ 维随机变量;而表现为已知数字的具体样本,则是这 $n$ 维随机变量的观测值.样本的这种两重性虽然是一件平凡的事情,但有很大的重要性.

**例 2.1** 为考察一物体的重量 $a$,用一架天平重复称量 $n$ 次,记称量结果为 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$,这就是样本.实际称量出来的 $n$ 个数,则是样本 $\boldsymbol{X}$ 的观测值.此例中总体理解为“一切可能出现的称量结果的集合”.这里,总体并非现实存在的对象的集合,而是用思维抽象出来的集合.但其总体也是明确的.

### 1.2.2 样本分布与总体分布

样本 $\boldsymbol{X}$ 既然是 $n$ 维随机变量,就有一定的概率分布.这一概率分布称为**样本分布**.样本分布是样本所受随机性影响的最完整的统计描述.它既与总体有关,又与抽样方式有关.抽样方式主要有不放回抽样、放回抽样等.从总体中一次同时抽取所需个数的抽样单元,或逐个抽取、且已抽到的单元不放回到总体中去,这种抽样方法称为不放回抽样.逐个抽取且每个被抽取到并经观测后的抽样单元,在抽取下一个抽样单元前都必须放回总体中去,这种抽样方法称为放回抽样.

**例 2.2** 抽样检查 设有一批产品共 $N$ 件,其中有 $M$ 件废品,$N$ 已知而 $M$ 未知.现从其中不放回地抽出 $n$ 件进行检验,每次抽取时未被抽过的产品被抽出的概率相同.要根据抽样检验结果对 $M$ 或废品率$\dfrac{M}{N}$ 进行统计推断.

**解** 先将属性指标数量化.正品对应于“0”,废品对应于“1”.这样,样本 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 中每一 $X_i$ 只能以“0”或“1”为值.

为掌握推理的思想方法,先就 $n=3$ 讨论其样本分布.先计算概率 $P[X_1=x_1,X_2=x_2,X_3=x_3]$,$x_i$ 取 0 或 1.按古典概型,总试验结果数为 $\mathrm{A}_N^3=N(N-1)(N-2)$.事件 $\{X_1=0,X_2=0,X_3=0\}$ 就是“取出的全是正品”,

即 3 件全从 $N-M$ 件正品中取得,所含试验结果数为 $A_{N-M}^3=(N-M)(N-M-1)(N-M-2)$.因此

$$\begin{aligned}&P[X_1=0,X_2=0,X_3=0]\\&=\frac{(N-M)(N-M-1)(N-M-2)}{N(N-1)(N-2)}.\end{aligned}$$

同样可得

$$P[X_1=1,X_2=1,X_3=1]=\frac{M(M-1)(M-2)}{N(N-1)(N-2)}.$$

事件$\{X_1=0,X_2=1,X_3=0\}$就是"第二件为废品,而第一、三件为正品",可考虑先从 $M$ 件废品中取一件放在第二位置上,再从 $N-M$ 件正品中取两件排列,放在第一、三位置上,所含试验结果数为 $MA_{N-M}^2=M(N-M)(N-M-1)$.因此,

$$P[X_1=0,X_2=1,X_3=0]=\frac{M(N-M)(N-M-1)}{N(N-1)(N-2)}.$$

同样可得

$$\begin{aligned}P[X_1=1,X_2=0,X_3=0]&=P[X_1=0,X_2=0,X_3=1]\\&=\frac{M(N-M)(N-M-1)}{N(N-1)(N-2)},\end{aligned}$$

$$\begin{aligned}&P[X_1=0,X_2=1,X_3=1]=P[X_1=1,X_2=0,X_3=1]\\&=P[X_1=1,X_2=1,X_3=0]=\frac{M(M-1)(N-M)}{N(N-1)(N-2)}.\end{aligned}$$

综合以上的讨论,就得到 $n=3$ 时的样本分布.

以下,讨论样本量为 $n$ 时的样本分布$P[X_1=x_1,X_2=x_2,\cdots,X_n=x_n]$,$x_i$ 取 0 或 1. 总的试验结果数为 $A_N^n=N(N-1)\cdots(N-n+1)$.事件$\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}$就是共有 $T=\sum_{i=1}^n x_i$ 件废品出现在特定的序位上,其余 $n-T$ 件是正品.这可以通过下列步骤实现:先从 $M$ 件废品中取出 $T$ 件进行排列,依次放到那特定的 $T$ 个位置上;然后从 $N-M$ 件正品中取出$n-T$ 件进行排列,依次放到其余 $n-T$ 个位置上.所含试验结果数为

$$\begin{aligned}A_M^TA_{N-M}^{n-T}&=M(M-1)\cdots(M-T+1)(N-M)\\&\quad\cdot(N-M-1)\cdots(N-M-n+T+1).\end{aligned}$$

由此可知

$$\begin{aligned}&P[X_1=x_1,X_2=x_2,\cdots,X_n=x_n]\\&=[N(N-1)\cdots(N-n+1)]^{-1}M(M-1)\cdots(M-T+1)\\&\quad\cdot(N-M)(N-M+1)\cdots(N-M-n+T+1),\end{aligned}\tag{2.1}$$

其中 $T=\sum_{i=1}^n x_i$,即样本中所含废品数.

**例 2.3** 在例 **2.2** 的假定条件下,若改为放回抽样,求样本分布 $P[X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n]$.

**解** 计算概率 $P[X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n]$.在放回抽样情况下,第 $i$ 次抽取时不论前$i-1$次抽取的结果如何,总有 $N$ 个产品,其中 $M$ 个废品.从而$\{X_i = x_i\}$, $i = 1,2,\cdots,n$,这 $n$ 个事件独立,且 $P[X_i = 1] = \dfrac{M}{N}$, $P[X_i = 0] = \dfrac{N-M}{N}$.且

$$P[X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n] = \prod_{i=1}^{n} P[X_i = x_i],$$

而在上式右端乘积的 $n$ 个因子中,$x_i$ 为 1 的有 $T = \sum_{i=1}^{n} x_i$ 个,其余为 0. 因此

$$P[X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n] = \left(\frac{M}{N}\right)^T \left(\frac{N-M}{N}\right)^{n-T}. \tag{2.2}$$

由此可见,放回抽样比不放回抽样计算简单.

从总体中随机抽取一个个体,结果为一随机变量.这个随机变量所有可能取值的集合就是总体各种不同数值全体所构成的集合.这个随机变量的概率分布反映了总体数量指标的分布状况.例如,某一批日光灯中,寿命在 1500 小时以上所占比率即寿命随机变量 $X$ 超过 1500 小时的概率,即 $P[X > 1500]$.因此,人们把这样抽样得到的随机变量的概率分布称为总体分布.再考虑样本量为 1 的样本 $X_1$,它就是上述抽样得到的随机变量.因此,我们把**总体分布**定义为样本量为 1 时的样本分布.总体的统计特性完全由总体分布刻画,因此在统计学中,一个总体必然对应着一个完全确定的总体分布.

知道样本分布,可以完全确定总体分布;但知道总体分布,并不能完全确定样本分布.样本分布不仅与总体分布有关,而且与随机抽样的方式有关.如例 2.2,例 2.3 中,总体分布都是

$$P[X_1 = x_1] = \frac{M^{x_1}(N-M)^{1-x_1}}{N}, \quad x_1 = 0,1. \tag{2.3}$$

但因抽样方式不同,其样本分布各为(2.1)、(2.2),是不同的.

抽样的目的是由取得的样本对总体分布中某些未知因素(如例 2.2、例 2.3 中的废品率 $p = M/N$)作出统计推断.为了使样本能更好地反映总体信息,必须进一步探讨抽样方法.最常用的抽样方法叫**简单随机抽样**,要求满足下列两点:

(1) 代表性 样本 $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)^{\mathrm{T}}$ 每一 $X_i$ 都能同样反映总体的特性.从概率论的观点讲,每一 $X_i$ 的概率分布均与总体分布相同.

(2) 独立性 样本 $\boldsymbol{X}$ 中,每一 $X_i$ 的取值既不受样本中其他$n-1$分量的影响,也不影响其他 $n-1$ 分量的取值.从概率论观点讲,$X_1, X_2, \cdots, X_n$ 是相

互独立的随机变量.

由简单随机抽样得到的样本 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 称为**简单随机样本**，或称独立同分布样本. 本书主要讨论简单随机样本，并简称为样本. 例 2.3 中，抽样方式为放回抽样，其样本为简单随机样本. 例 2.2 中，抽样方式为不放回抽样，其样本为非简单随机样本.

**定义 2.1**　设随机变量 $X$ 的分布函数为 $F(x)$，$n$ 维随机变量 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 中各分量 $X_1,X_2,\cdots,X_n$ 相互独立且与 $X$ 有相同分布. 我们称 $\boldsymbol{X}$ 为总体(也可称总体 $F(x)$)，$F(x)$ 为总体分布函数；并称 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 为来自总体 $\boldsymbol{X}$(或 $F(x)$)的样本量为 $n$ 的简单随机样本，或称独立同分布样本，简记为 i.i.d. 样本.

定义 2.1 反映了统计基本概念与概率论基本概念的联系. 若 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 是从分布函数为 $F(x)$ 的总体 $\boldsymbol{X}$ 中抽得的 i.i.d. 样本，则 $X_1,X_2,\cdots,X_n$ 是独立同分布随机变量，且其分布函数皆为 $F(x)$. 记为

$$X_1,X_2,\cdots,X_n \sim X \tag{2.4}$$

或

$$X_1,X_2,\cdots,X_n \sim F. \tag{2.5}$$

若分布 $F$ 有密度函数 $f(x)$，则可记为

$$X_1,X_2,\cdots,X_n \sim f. \tag{2.6}$$

应该注意，对于 i.i.d. 样本，由总体分布可以决定样本分布. 这时，样本 $\boldsymbol{X}=(X_1,X_2,\cdots,X_n)^{\mathrm{T}}$ 的 $n$ 维分布函数为

$$F(x_1)F(x_2)\cdots F(x_n)=\prod_{j=1}^{n}F(x_j), \tag{2.7}$$

在有密度函数的情况下，样本 $\boldsymbol{X}$ 的 $n$ 维密度函数为

$$f(x_1)f(x_2)\cdots f(x_n)=\prod_{j=1}^{n}f(x_j), \tag{2.8}$$

在总体 $X$ 为离散随机变量时，具有概率函数 $p(x)=P[X=x]$，而样本 $\boldsymbol{X}$ 的 $n$ 维概率函数为

$$p(x_1)p(x_2)\cdots p(x_n)=\prod_{j=1}^{n}p(x_j). \tag{2.9}$$

例如，例 2.3 的总体分布的概率函数为(2.3)，而其样本分布的 $n$ 维概率函数为(2.2).

### 1.2.3　样本空间与分布族

**定义 2.2**　样本所有可能取值构成的集合($n$ 维空间或其中一个子集)，