研究生教学用书

# 应用数理统计

（第二版）

邰淑彩　孙韫玉　何娟娟　编著

WUHAN UNIVERSITY PRESS
武汉大学出版社

研究生教学用书

# 应用数理统计

（第二版）

郜淑彩　孙韫玉　何娟娟　编著

WUHAN UNIVERSITY PRESS
武汉大学出版社

**图书在版编目(CIP)数据**

应用数理统计/郜淑彩，孙韫玉，何娟娟编著．—2版．—武汉：武汉大学出版社，2005.7

研究生教学用书

ISBN 7-307-04488-9

Ⅰ.应… Ⅱ.①郜… ②孙… ③何… Ⅲ.数理统计 Ⅳ.O212

中国版本图书馆CIP数据核字(2005)第033758号

责任编辑：李汉保　　责任校对：刘　欣　　版式设计：支　笛

出版发行：**武汉大学出版社**　(430072　武昌　珞珈山)

(电子邮件：wdp4@whu.edu.cn 网址：www.wdp.com.cn)

印刷：湖北恒泰印务有限公司

开本：787×980　1/16　印张：23.875　字数：437千字

版次：2000年5月第1版　2005年7月第2版

2005年7月第2版第1次印刷

ISBN 7-307-04488-9/O·322　定价：30.00元

# 内容简介

本书是由武汉大学研究生院资助的为高等院校非数学类专业硕士研究生编写的教材。其主要内容包括:数理统计的基本概念,参数估计,假设检验,回归分析,方差分析及正交试验设计,多元统计分析等。书中除了介绍了数理统计的经典理论外,还适当地介绍了一些近代数理统计的概念与方法,同时还介绍了目前国际上流行的应用统计软件SAS和SPSS等。书中各章节均配有适量的习题,书末附有习题答案。

本书可以作为高等院校工科、经济类、农学、医学、师范、财经、统计、管理等非数学类专业硕士和博士研究生以及高年级本科生学习应用数理统计课程的教科书,也可以作为相关工程技术人员、高等学校教师的应用数理统计参考书。

# 前　言

《应用数理统计》是武汉大学工科研究生必修的学位课程之一，该课程以工科院校本科生通用的工科数学《概率论与数理统计》作为预备知识。

数理统计一直是应用数学中最重要、最活跃的学科之一，该学科在自然科学和社会科学中的应用越来越广泛深入，在国民经济和科学技术中的作用越来越重要。作为工科研究生，理应具备数理统计的基础知识、掌握其思想方法。为此，我们根据多年的教学实践经验和我校研究生相关学科的特点，参照工科研究生《应用数理统计》课程教学的基本要求，考虑到面向21世纪工科研究生数理统计课程教学改革和实际应用的需要，编写了这本教材。

编写这本书的主要指导思想有以下四点：

第一，注重思想方法的介绍。数理统计不仅是一门数学理论，而且还是一种“看世界”的方法。具备正确的数理统计的思想方法是工科研究生观察现实世界所必备的文化修养。为此，本书特别注重介绍各种统计方法的统计思想、问题的背景、统计方法产生的历史和不同统计学派的特点，使学生能够对数理统计的思想方法有一个系统的、全面的了解。

第二，注重应用性。数理统计是一门应用性很强的应用数学学科，该学科的应用几乎遍及社会科学、自然科学与工程技术的各个领域，成为解决国民经济和科学技术中实际问题的重要工具。因此，本书除了突出各种统计方法的应用外，还编入了应用面较广的多元统计分析和贝叶斯估计方法，充实了一些应用性较强的内容，如：极大似然估计的数值解法及渐近分布；样本容量 $n$ 的确定；单参数、多参数模型下的似然比检验；非参数检验的若干方法，自变量的选择与岭估计等。并介绍了目前世界上流行的统计软件，以适应工科研究生解决实际问题的需要。

第三，突出数理统计这门学科的特点。统计推断是由样本去推断总体，由部分去推断整体，这就不可避免地会产生误差，数理统计的任务就是要充分利用样本的信息，做出尽可能精确、尽可能可靠的推断。书中通过方法的介绍、实例的分析来突出这个特点。

第四，注重与本科《概率论与数理统计》教材的有机衔接。由于本科生在

工科数学《概率论与数理统计》中已学习过了分位数、参数的区间估计的概念及正态总体参数的区间估计和假设检验等，本书不再赘述。而在内容的广度和深度方面作了较大的调整和提高，例如增加了常用分布族、极大似然估计的数值解法及其渐近分布、似然比检验及变量的选择与岭估计等。以便让学生开拓思路，掌握更多的方法去解决生产和科研中的实际问题。

本书由邰淑彩组织编写，大纲和体系由集体讨论而定。全书共分六章。第一章、第六章及第二章中的§2.4由孙韫玉编写；第二章中的§2.1～§2.3及第三章由何娟娟编写；其余章节由邰淑彩编写。全书最后由邰淑彩统稿。

武汉大学数学与统计学院刘禄勤教授逐字逐句地审阅了全书，提出了许多宝贵的意见，作者根据刘教授的意见，对全书作了修改和补充。本书的出版自始至终得到了武汉大学研究生院、武汉大学数学与统计学院、武汉大学出版社的大力支持，尤其是李汉保编辑给予了很多帮助。在此向所有协助本书出版的同志表示衷心的感谢！

在编写本书过程中，我们参考了较多的相关文献，在参考文献中未一一列出，在此对相关文献作者表示衷心感谢！

限于作者的水平，书中难免存在缺点和错误，恳请专家和读者批评指正。

作　者

2004年9月于武汉大学

# 目　录

# 第一章 基础知识

## §1.1 引 言

数理统计学是数学的一个分支.该学科的任务是研究怎样用有效的方法去收集、分析和使用受随机性影响的数据.

数理统计学研究的对象是受随机性影响的数据.是否假定数据有随机性,这是区别数理统计方法和其他数据处理方法的根本点.数据的随机性来源有二:一是抽样的随机性,出于经济原因的考虑或时间限制或问题性质决定,不可能或没有必要得到研究对象的全部资料,而只能用"一定的方式"抽取其中一部分进行考察,这样所得数据的随机性就是来自抽样的随机性;二是试验过程中的随机误差,即在试验过程中未加控制,或无法控制,或不便控制,甚至是不了解的因素所引起的误差.在实际问题中这两类随机性常常交织在一起.例如某工厂生产出大量的电视机显像管,为了检测显像管的寿命,推断寿命的分布类型、相关参数的具体数值以及是否达到生产要求,等等,必须对显像管的寿命进行测试,由于寿命试验具有破坏性,所以只能抽取少量显像管以一定的方式进行加速老化试验而得到部分数据,这里,抽样的随机性对数据便有影响.另外产品即使是在同一条件下生产出来的,但各台显像管的寿命仍会有差异,这就是随机误差对数据的影响.

数理统计学研究的内容随着科学技术和生产实践的不断进步而逐步扩大,概括起来可以分为两大类:(1)用有效的方法去收集数据.这里"有效"一词有两方面的含义,一是可以建立一个在数学上便于处理的模型来描述所得数据;二是数据中要包含尽可能多的与所研究的问题有关的信息.对该问题的研究构成了数理统计学中的两个分支,即抽样理论和试验设计,这些不是本书的主要内容.(2)有效地使用数据.获取数据以后,必须使用有效的方法去集中和提取数据中的相关信息,以对所研究的问题作出尽可能精确和可靠的结论,这种"结论"在统计学中叫做"推断".有效地使用数据是比有效地收集数据更为复杂的问题,这一问题的研究构成了数理统计学的中心内容——统计推断.上面提到的推断显像

管寿命的分布类型、相关参数的具体数值以及是否达到生产要求等都是统计推断所要解决的问题.本书将主要讨论统计推断.数理统计学中除了上面所提到的一些分支外,还有不少内容,如质量控制,可靠性理论,统计决策理论,时间序列分析等,限于篇幅,本书未作介绍.

需要强调的是由于收集和使用的仅仅是部分数据,且带有随机性,要利用部分去推断总体,其结论只能做到尽可能而非绝对的精确和可靠,而结论的正确性程度显然可以用概率来度量,因此概率论是数理统计的基础,这是毫无疑义的.不过统计方法的具体使用并不需要很高深的数学知识,但这些方法的理论依据,不具备较多、较深的数学知识就说不清楚.本书着重介绍数理统计方法,也给出一些必要的数学推导,但不追求其严密性及完整性.

由于随机性影响无所不在,所以数理统计方法的应用十分广泛,几乎在人类活动的一切领域中都能程度不同地找到它的应用.例如,在工农业生产中最佳生产工艺的安排,产品质量的控制管理,技术革新前后产品质量的鉴定,元、器件寿命的计算,工程设计中安全系数的统计分析,等等;在医学卫生领域中药物疗效的检验,某种疾病与某特定因素的关系大小的确定,大气污染的诸有害成分的主成分分析,等等;在社会、经济领域中的抽样调查、民意测验、市场预测等都要用到数理统计方法.此外,在测量、通信、气象、水文、地震预报、地质探矿、考古研究、刑事鉴别等各方面数理统计方法都得到愈来愈多的应用.

数理统计学是一门较年轻的学科.该学科正式诞生于19世纪后期,至20世纪20年代这门学科已稳稳地站住了脚跟,到40年代已形成为一个成熟的数学分支.第二次世界大战后工农业和科技等方面迅速发展,对数理统计不断提出新的课题,使其不断地向纵深发展.特别是由于电子计算机具有大批量、高速度处理数理的能力,使得数理统计方法得到广泛的应用,各种使用方便的统计程序包的出现使得数理统计在各领域中发挥着越来越大的作用.

## §1.2 数理统计中的几个基本概念

### 1.2.1 总体与样本

直观地说,把研究对象的全体称为总体(又称母体),而组成总体的每个元素称为个体.总体包含的个体数可以是有限的,也可以是无限的.对每个个体来说,它有各个方面的特性,而人们关心的往往不是个体的一切方面,只是它的某个(或某几个)数量指标以及该指标在总体中的概率分布情况.例如,在研究一批电视机组成的总体时,可能关心的是电视机显像管寿命的概率分布情况,由于任何

一台显像管的寿命事先是不能确定的，而每一台显像管都确实对应有一个寿命值，所以可以认为显像管寿命是一个随机变量，也就是说，把总体与一个随机变量（显像管寿命）联系起来，对总体的研究也就转化为对表示总体的随机变量的统计规律的研究．这样，就可以用精确的语言来描述总体与个体．总体就是一个具有确定概率分布的随机变量（一维或多维），而一个个体则是随机变量的一次观测值．以后常用大写字母 $X$、$Y$ 等表示总体．

总体 $X$ 的概率分布是确定的，但又是未知的，或至少分布的某些参数是未知的．为了研究总体的情况，必须在总体中抽取一定数量的个体进行观测，这个过程称为抽样（也称取样、采样）．从一个总体 $X$ 中抽取 $n$ 个个体为观测值（$x_1$，$x_2$，$\cdots$，$x_n$），这样取得的（$x_1$，$x_2$，$\cdots$，$x_n$）称为取自总体 $X$ 的一个样本（又称子样）．样本中个体的数目 $n$ 称为样本容量．注意到数理统计学的对象是受随机性影响的数据，对于某一次抽样来说，所得观测数据（$x_1$，$x_2$，$\cdots$，$x_n$）是完全确定的一组数，但由于抽样的随机性，观测值（$x_1$，$x_2$，$\cdots$，$x_n$）也是随机变化的，为了强调随机性，用（$X_1$，$X_2$，$\cdots$，$X_n$）表示样本，也就是说，应从两个角度来看待样本，在抽样之前，样本是一个（$n$ 维）随机变量，在进行一次具体的抽样之后它就是一个数组，这就是样本的二重性．认识样本的二重性是十分重要的，一般在理论推导中总把样本视为随机变量，而在用理论推导所得出的结论进行具体推断时，样本就成了具体数字了．把样本（$X_1$，$X_2$，$\cdots$，$X_n$）可能取值的全体称为样本空间，记为 $X$，它通常为 $n$ 维空间或其中的某一个子集，一个具体样本值（$x_1$，$x_2$，$\cdots$，$x_n$）则是样本空间中的一个点．

为了使样本能很好地反映总体，除了要求抽样具有随机性以外，较为自然也最有实用价值的要求是：(1)独立性．因为独立观测是一种最简单的观测方法，对所得数据进行处理也较方便，所以要求在抽取 $n$ 个个体时，每一次抽样都是独立进行的，即各次抽取的结果彼此互不影响，用概率论语言叙述，即要求 $X_1$，$X_2$，$\cdots$，$X_n$ 是相互独立的随机变量．(2)代表性．要求抽取 $n$ 个个体时每一次抽样都是在完全相同的条件下进行的，这样就能保证每一个 $X_i$（$i=1,2,\cdots,n$）都具有总体的特征，即要求每一个 $X_i$（$i=1,2,\cdots,n$）都与总体有相同的分布．上述两点要求在多数情况下是容易得到满足的．凡满足这两个要求所得的样本称为简单随机样本．为明确起见下面用定义形式表述出来．

**定义 1.2.1** 设随机变量 $X_1$，$X_2$，$\cdots$，$X_n$ 相互独立，且每一个 $X_i$（$i=1,2,\cdots,n$）与总体 $X$ 有相同的概率分布，则称（$X_1$，$X_2$，$\cdots$，$X_n$）为来自总体 $X$ 的容量为 $n$ 的简单随机样本．

今后若无特别声明，凡提到样本，都是简单随机样本．

引入了简单随机样本的概念，就可以利用概率论中对独立同分布的随机变

量序列所建立的许多重要定理,这些重要定理的结论为数理统计提供了必要的理论基础.下面的定理以后会多次用到,该定理由多维随机变量的分布性质推出.

**定理1.2.1** 若$(X_1,X_2,\cdots,X_n)$是来自总体$X$的样本,设$X$的分布函数为$F(x)$,则样本$(X_1,X_2,\cdots,X_n)$的联合分布函数为$\prod_{i=1}^{n}F(x_i)$.

**例1.2.1** 设$(X_1,X_2,\cdots,X_n)$是取自总体$X$的样本,$X$服从参数为$\lambda$的指数分布,则$(X_1,X_2,\cdots,X_n)$的联合分布函数为

$$F(x_1,x_2,\cdots,x_n)=\begin{cases}\prod_{i=1}^{n}(1-\mathrm{e}^{-\lambda x_i}), x_i>0,(i=1,2,\cdots,n)\\ 0, \quad \text{其他}\end{cases}$$

联合概率密度函数为

$$f(x_1,x_2,\cdots,x_n)=\begin{cases}\prod_{i=1}^{n}\lambda\mathrm{e}^{-\lambda x_i}, x_i>0,(i=1,2,\cdots,n)\\ 0, \quad \text{其他}\end{cases}$$

### 1.2.2 统计量

样本是对总体进行统计分析和推断的依据,虽然样本含有总体的信息,但比较分散,必须经过一定的加工、提炼,把分散在样本中有用的信息集中起来.具体地说,就是针对不同问题构造样本的各种函数,再利用这些函数去推断总体的性质,在数理统计学中称这种函数为统计量.

**定义1.2.2** 设$(X_1,X_2,\cdots,X_n)$为取自总体$X$的一个样本,$T(X_1,X_2,\cdots,X_n)$为$(X_1,X_2,\cdots,X_n)$的一个实值函数,且$T$中不包含任何未知参数,则称$T=T(X_1,X_2,\cdots,X_n)$为一个统计量.

作为统计量必须不含任何未知参数,这一点是非常重要的.因为在有些情形,统计量$T$是作为未知参数$\theta$的估计量而构造的,若$T$中含有未知参数$\theta$,就无法作为$\theta$的估计了.注意到样本的二重性,作为样本的函数的统计量也就具有二重性,即统计量$T(X_1,X_2,\cdots,X_n)$为随机变量,它应有确定的概率分布,称之为抽样分布;而对于样本的一个观测值$(x_1,x_2,\cdots,x_n)$,统计量$T(X_1,X_2,\cdots,X_n)$也有一个相应的值$T(x_1,x_2,\cdots,x_n)$.

下面介绍几个常用的重要统计量.

**定义1.2.3** 设$(X_1,X_2,\cdots,X_n)$是从总体$X$中抽取的一个样本,我们定义下列统计量:

样本均值 $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$

样本方差 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right)$

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

样本 $k$ 阶原点矩 $M_k = \frac{1}{n}\sum_{i=1}^{n}X_i^k, (k = 1,2,\cdots)$

样本 $k$ 阶中心矩 $M'_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^k, (k = 2,3,\cdots)$

这些统计量统称为总体的样本矩.

显然 $M_1 = \bar{X}$, $\bar{X}$ 是样本的算术平均值, $M'_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$. 本书中常将 $M'_2$ 用 $\tilde{S}^2$ 表示, $\tilde{S}^2$ 与 $S^2$ 略有不同, 但它们都是样本平均偏差平方和. $\bar{X}$ 和 $S^2$ 是以后用得最多的统计量, 由下面定理可以看出, $\bar{X}$ 集中反映了总体均值的信息, $S^2$ 集中反映了总体方差的信息.

**定理 1.2.2** 设 $(X_1, X_2, \cdots, X_n)$ 是取自总体 $X$ 的一个样本, 若 $X$ 的二阶矩存在, 并记 $E(X) = \mu, D(X) = \sigma^2$, 则有

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

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

$$D(\bar{X}) = D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}D(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}D(X) = \frac{\sigma^2}{n}$$

$$\begin{aligned}E(S^2) &= E\left[\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2 - n\bar{X}^2\right)\right]\\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n}E(X_i^2) - nE(\bar{X}^2)\right]\\ &= \frac{1}{n-1}\left[\sum_{i=1}^{n}(D(X_i) + (E(X_i))^2) - n(D(\bar{X}) + (E(\bar{X})^2)\right]\\ &= \frac{1}{n-1}\left[n(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right] = \sigma^2\end{aligned}$$

还可以推出 $D(S^2)$ 与总体矩的关系式, 因推导较繁琐, 故而略去.

需要指出的是, 若总体 $X$ 的 $k$ 阶矩存在, 则样本的 $k$ 阶矩必依概率收敛于总体的 $k$ 阶矩. 例如, $M_k = \frac{1}{n}\sum_{i=1}^{n}X_i^k$ 为样本 $k$ 阶原点矩, $\mu_k = E(X^k)$ 为总体 $k$ 阶原点矩, 因为 $X_1, X_2, \cdots, X_n$ 相互独立且与 $X$ 同分布, 所以 $X_1^k, X_2^k, \cdots, X_n^k$ 相互独立且与 $X^k$ 同分布, 再注意到

$$E(M_k) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i^k\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i^k) = \frac{1}{n}\sum_{i=1}^{n} E(X^k) = \mu_k$$

故由独立同分布的辛钦(Хинчин)大数定律可知,当 $n \to \infty$ 时,$M_k$ 依概率收敛于 $\mu_k$.

**定义 1.2.4** 设 $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ 是从二维总体 $(X, Y)$ 中抽取的一个样本,记

$$S_{12} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(T_i - \bar{Y})$$

$$R = \frac{S_{12}}{\tilde{S}_1 \tilde{S}_2}$$

其中 $\tilde{S}_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$, $\tilde{S}_2^2 = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$,则称 $S_{12}$, $R$ 分别为二维样本的协方差和二维样本的相关系数. $S_{12}$ 及 $R$ 集中反映二维总体中 $X$ 与 $Y$ 的协方差及相关系数的信息.

### 1.2.3 理论分布与经验分布

总体 $X$ 的分布函数称为理论分布函数,样本的分布函数则称为经验分布函数,其具体定义如下.

**定义 1.2.5** 设 $(X_1, X_2, \cdots, X_n)$ 为取自总体 $X$ 的一个样本,$(x_1, x_2, \cdots, x_n)$ 是样本的一个观测值,将这些值按大小递增顺序排列成 $x_{(1)} \leqslant x_{(2)} \leqslant \cdots \leqslant x_{(n)}$,并作函数

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

称 $F_n(x)$ 为总体 $X$ 的经验分布函数,如图 1-1 所示.

前面强调过样本具有二重性,那么反映在经验分布函数 $F_n(x)$ 上又该如何体现呢?理论分布与经验分布又有什么关系呢?这就要看经验分布函数所具有的独特性质.

(1) 对于给定的样本值 $(x_1, x_2, \cdots, x_n)$,经验分布函数 $F_n(x)$ 是单调不减、右连续的跳跃函数(或称阶梯函数),且 $0 \leqslant F_n(x) \leqslant 1$, $F_n(-\infty) = 0$, $F_n(+\infty) = 1$,故 $F_n(x)$ 满足分布函数的三个条件.

(2) 对于 $x$ 的每一个固定值而言,$F_n(x)$ 又是样本 $(X_1, X_2, \cdots, X_n)$ 的函数,因而 $F_n(x)$ 是一个统计量,作为随机变量的 $F_n(x)$,其可能取的值为 $0, \frac{1}{n}$,

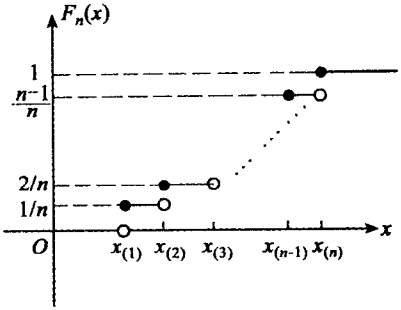


图 1-1

$\frac{2}{n},\cdots,\frac{n-1}{n},1$，且 $P\left\{F_n(x)=\frac{k}{n}\right\}=C_n^k[F(x)]^k[1-F(x)]^{n-k}$，这是因为 $X_1,X_2,\cdots,X_n$ 相互独立且与 $X$ 同分布，事件 $\left\{F_n(x)=\frac{k}{n}\right\}$ 发生等价于事件{$(X_1,X_2,\cdots,X_n)$ 的一个观测值中恰有 $k$ 个分量小于或等于 $x$} 发生，又因为抽取一个样本等价于对总体 $X$ 进行 $n$ 次独立重复试验，即 $n$ 重贝努利(J. Bernoulli)试验，而

$$P\{X_i\leqslant x\}=P\{X\leqslant x\}=F(x)$$

以 $A$ 表示事件{观测值某分量落入$(-\infty,x]$中}，那么

$$\begin{aligned}P\left\{F_n(x)=\frac{k}{n}\right\}&=P\{\text{恰有 }k\text{ 个 }X_i\text{ 满足 }X_i\leqslant x\}\\&=P\{n\text{ 次独立重复试验中 }A\text{ 恰好出现 }k\text{ 次}\}\\&=C_n^k[F(x)]^k[1-F(x)]^{n-k}.\end{aligned}$$

(3) 由上述讨论可知随机变量 $nF_n(x)$ 服从二项分布 $B(n,F(x))$，于是由贝努利大数定律知，当 $n\to\infty$ 时，经验分布函数 $F_n(x)$ 依概率收敛于总体 $X$ 的分布函数 $F(x)$. 还有比这更深刻的结果，这就是格列汶科定理.

**定理 1.2.3** (格列汶科 Гливенко 定理) 设总体 $X$ 的分布函数为 $F(x)$，经验分布函数为 $F_n(x)$，则当 $n\to\infty$ 时有

$$P\{\lim_{n\to\infty}\sup_{-\infty<x<+\infty}|F_n(x)-F(x)|=0\}=1.$$

定理 1.2.3 的证明从略. 从这个定理大致可以看到，当 $n$ 充分大时，事件{对所有 $x$ 值，$F_n(x)$ 与 $F(x)$ 最大差异非常小} 发生的概率近似等于 1.

这一性质说明，当 $n$ 很大时，可以用经验分布函数 $F_n(x)$ 去估计总体 $X$ 的理论分布函数 $F(x)$. 而这正是数理统计中用样本进行估计和推断总体的理论依据.