医学专业研究生教学用书

# 医学统计与CHISS应用

YIXUE TONGJI YU
CHISS YINGYONG

主　编／童新元　王洪源　郭秀花

人民軍醫出版社
PEOPLE'S MILITARY MEDICAL PRESS

医学专业研究生教学用书

# 医学统计与 CHISS 应用

YIXUE TONGJI YU CHISS YINGYONG

主审　王　彤

主编　童新元　王洪源　郭秀花

人民军医出版社

People's Military Medical Press

北京

**图书在版编目(CIP)数据**

医学统计与 CHISS 应用/童新元等主编. —北京:人民军医出版社,2006.9
ISBN 7-5091-0423-8

Ⅰ.医… Ⅱ.童… Ⅲ.①医学统计-研究生-教材②医学统计-统计分析-应用软件,CHISS-研究生-教材 Ⅳ.R195.1

中国版本图书馆 CIP 数据核字(2006)第 069312 号

**策划编辑**:王 敏 杨德胜 **文字编辑**:余满松 **责任审读**:黄栩兵

**出 版 人**:齐学进

**出版发行**:人民军医出版社 **经销**:新华书店

**通信地址**:北京市 100036 信箱 188 分箱 **邮编**:100036

**电话**:(010)66882586(发行部)、51927290(总编室)

**传真**:(010)68222916(发行部)、66882583(办公室)

**网址**:www.pmmp.com.cn

**印刷**:北京国马印刷厂 **装订**:京兰装订有限公司

**开本**:787mm×1092mm 1/16

**印张**:17.25 **字数**:413 千字

**版、印次**:2006 年 9 月第 1 版第 1 次印刷

**印数**:0001～4500

**定价**:35.00 元

# 内 容 提 要

本教材是根据现代应用统计学发展的特点，考虑到在统计咨询和数据分析的实际工作中解决难点和重点问题的需要，配合医学院校医学统计学教学内容，结合中文统计软件 CHISS 而组织编写的，并力争做到方便教师教学和学生使用。分 22 章系统介绍了医学统计的基本概念、基本原理与基本方法，并突出统计设计、数据管理和数据质量控制的内容。本书内容丰富，实用性强，可作为医学研究生的教材使用；各医学专业研究生班和本科院校可根据教学对象和学时安排，对内容进行适当取舍；也可作为其他专业，如心理学、管理学、生物学的广大教师、科研人员、管理人员等学习统计学的参考书。

**责任编辑** 王 敏 杨德胜 余满松

# 编著人员名单


**主　审**　王　彤

**主　编**　童新元　中国人民解放军总医院

王洪源　北京大学公卫学院

郭秀花　首都医科大学公共卫生与家庭医学学院

**副主编**　王泓午　天津中医药学院

罗艳侠　首都医科大学公共卫生与家庭医学学院

赵　炜　中国人民解放军总医院

**编　委**　（以姓氏笔画为序）

万　霞　北京中医药大学

王泓午　天津中医药学院

王洪源　北京大学公卫学院

尹立群　天津中医药学院

冯　丹　中国人民解放军总医院

李　戈　天津中医药学院

杨新华　首都医科大学公共卫生与家庭医学学院

何丽云　中国中医研究院

陈　琦　首都医科大学卫生管理与教育学院

罗艳侠　首都医科大学公共卫生与家庭医学学院

赵　炜　中国人民解放军总医院

赵铁牛　天津中医药学院

胡镜清　中国中医研究院

夏　蕾　中国人民解放军总医院

徐　涛　中国协和医科大学

郭　静　中国疾病预防控制中心

郭秀花　首都医科大学公共卫生与家庭医学学院

曹红霞　首都医科大学公共卫生与家庭医学学院

曹秀堂　中国人民解放军总医院

童新元　中国人民解放军总医院

# 前　言

当今数字化信息时代，无论社会政治、军事、经济，还是生物医学、教育心理、工农业生产等各行各业都存在着大量的数据。有了数据不等于有了信息，数据只有通过科学的方法进行处理才能加工成为信息。统计学是一种非常有用的处理数据的方法，它可以帮助我们从数据中挖掘有用的证据，消除虚假的信息，找出哪些现象是必然的，哪些是偶然的，发现事物内部的规律性，透过现象发现事物内部的规律。

医学统计学是统计学在医学领域的应用，被我国医学院校和研究机构列为医学专业研究生的必修课。同时，根据现代统计学的迅速发展，不少大专院校还开设了现代医学统计选修课。可见医学统计学在医学研究中地位之重要。

在国内外医学统计学家的长期努力下，已编写了不少经典的统计学教材，每一本书都是在前辈们工作的基础上不断更新发展而产生的，没有一本是凭空写出来的书。我们在教学工作中觉得适合于临床医学研究生教学和使用的教材并不多。为了适应现代医学统计学的发展，特编写这本医学统计学教材。本教材是在参阅国内外多本著名教材的基础上，结合编写人员的教学工作编写而成的，本教材有如下的特点：

(1)在内容安排上注意与医学科研实际相结合，注意统计知识的整体性与连贯性，将科研统计设计、数据管理、质量控制与数据统计有机结合，将部分设计内容前置，突出统计设计的重要性，并强调数据管理与数据质量的必要性。

(2)不需要太多的数学理论知识，没有公式的推导与证明。全书系统介绍医学统计学的“三基”，即基本概念、基本原理与基本方法。其特点是实用性强：重点在于什么样的问题采用什么样的统计设计；什么样的实际数据，采用什么统计分析方法；以及如何对统计学结果进行合理的解释。

(3)计算机技术的产生被认为是统计学的第二次技术革命，已成为现代统计教育和统计分析的一个非常重要的工具。本书结合中文统计软件 CHISS 而编写，CHISS 软件方便好用，为学习者节省了大量统计计算工作时间，从而将学习重点转移到对统计的“三基”的理解，而非数据公式的具体使用与计算。

(4)注重统计学方法的适用性与通用性，并将它与现代统计学的理论相结合，如介绍率的多重比较、非参数的多重比较、列联表的统计学分析及 Meta 分析等内容。

我们衷心地感谢山西医科大学王彤教授对全书进行审阅和指导，感谢中国人民解放军总医院研究生处的大力支持，也感谢人民军医出版社王敏主任的帮助。同时，还要感谢各位作者们及采用本教材进行教学工作所在单位的鼎力支持。

由于时间仓促和学识所限，本书中还有许多不足之处，我们将虚心听取各位专家、同仁及本书使用者的批评与建议，争取在再版中给予弥补。

童新元　谨识于北京

2006 年 6 月

“按照现代理论，自然规律的基础不是因果性，相反，本质上具有统计性质。……人们断言，一切自然规律‘在原则上’都是统计性的，只是我们观察操作不完善，我们才受骗去信仰严格的因果性。”

Albert Einstein（德国科学家，1879—1955）

摘自：爱因斯坦 1928 年关于《物理学的基本概念及其最近的变化》的演说

# 目　录

“当人类科学探索者在问题的丛林中遇到难以逾越的障碍时，惟有统计工具可为其开辟一条前进的通道”。

——英国著名遗传学家 Galton（1822—1911）

# 第一章 绪 论

## 第一节 统计学概论

21 世纪，人类进入信息化社会，统计对我们每个人来说并不陌生，报刊杂志、电视广播、网络媒体等每时每刻都传递着很多统计数据和信息，我们常听到很多关于“统计”的词汇。例如，据统计，今天平均气温 20℃，降水概率 70%；去年国民生产总值 GDP 增长 8%；北京市人均寿命 72 岁；SARS 病人中医务人员占 20%等等。可以说统计学的知识已经渗透到自然科学、社会科学及人类生活的各个领域。在现代社会中，大到国家重大政策的制定，小到人们的日常生活，几乎都离不开统计。

1983 年 12 月 8 日，第六届全国人民代表大会常务委员会第三次会议通过了《中华人民共和国统计法》，为发挥统计信息、咨询、监督的作用提供了重要的法律保障，对发展我国统计事业具有重要意义。1998 年 9 月，国家教育部颁布的《普通高等学校本科专业目录和专业介绍》将统计学列为理学类一级学科。2001 年，国家教育部为推进基础教育改革而推出新课程标准，将统计学纳入新的小学数学课程。教学大纲要求小学生要“经历运用数据描述信息、做出推理的过程，发展统计观念”。美国国立卫生研究院（National Institutes of Health，NIH）的基金申请明确要求基金合作者中有统计学家，并且在所立项目中有统计学方面的思考。美国国家药品食品管理局（Food and Drug Administration，FDA）要求新药的研发试验中，必须有统计学家指导研究设计、数据分析和报告的呈递等。

什么是统计？《现代汉语词典》中一种定义为“总括地计算”。在不同的领域，不同的教科书上对统计的定义不尽相同。实际上统计有两种含义。广义上讲统计是通常人们所遇见的任何用数字、表格与图形所表达的一个事实。统计学的英语单词是 statistics，最早来源于词 state，即指政府的信息，现在统计不限于此。现代社会中各行各业都存在大量的统计信息和指标，如经济学中的物价指数，社会学中的人口失业率，商业中的产品占有率，金融学中的投资机会，医学中的发病率等；统计学作为一种学科，有其自身独有的知识体系和方法论。著名 Webster 国际大辞典中定义，统计学是“a science dealing with the collection，analysis，interpretation，and presentation of masses of numerical data”。即统计学是一门关于收集、分析、解释和表达数据的科学。

现代统计学是面对不确定性问题如何做决策的科学。作为医务工作者常常遇见很多实际问题，如在全民健康教育中，我们宣传吸烟、过量饮酒会损害健康，而运动、积极乐观的心情有

利于人的健康。然而，吸烟的危害到底有多大？运动有多大好处？人们常常要问癌症病人做手术后能生存多久？一种新药的用量、用法如何，疗效怎样判定？这些都是同统计有关的不确定性命题。统计学就是把统计学的语言引入具体的科学领域，把具体科学领域中遇到的问题抽象为统计学问题，最终用统计学知识解决具体的科学领域问题的过程。从哲学角度看，统计学是从个性中寻找共性，透过偶然现象看事物内部本质规律的一种方法和手段。可以说统计学既是一门科学，也是一门哲学。

统计学的理论是随着人类社会生产的需要而产生，随着人类社会生产的发展而更新的，特别是近 20 年来，统计学在理论方法和应用方面得到迅速发展。统计学与各个专业结合而形成数十个学科分支，如人口统计学、心理统计学、遗传统计学、社会统计学、经济统计学等，而且，新的领域与统计学结合形成的新的分支将继续不断出现。

医学统计学是用统计学的原理和方法，研究医学领域中不确定性现象的规律性的一门学科，是统计学与医学相结合形成的一个交叉科学，是现在及未来一个世纪中最活跃、最有生命力的学科之一。

## 第二节　医学统计的基本内容

传统的观念认为，医学统计学只是对医学问题做数据列表和数据分析，这是很片面的。现代医学统计学主要包括试验设计、数据管理与质量控制、数据统计分析三个大方面的内容。

### 一、试 验 设 计

试验设计(design)是对整个研究过程的总体设想和安排，是医学科研和统计工作的基础。研究设计直接影响着试验结果的准确性、可靠性、严密性和代表性，一旦出现设计上的失误或缺陷，可能会导致整个研究的失败。

试验设计包括试验因素、调查项目、样本含量估计及研究对象选择、随机抽样方法、观测指标选取、误差控制等等。研究设计就是要解决这些问题，并通过周密的考虑和妥善的安排，用尽可能少的人力、物力和财力获取准确可靠的科学结论。在一项科学研究中，试验设计占整个研究工作量和投入的比例约为 5%，但是，一项完美的试验设计的完成预示着该项研究已经至少完成了 75%。

### 二、数据管理与质量控制

数据管理与质量控制包括如何搜集和获取试验与观察数据、如何整理数据、如何避免产生错误的数据、如何保证数据的质量、如何评价数据的质量，它对整个研究过程进行管理、控制和监督。数据收集是科学试验很重要的方面，其原则是要确保收集数据的时效性、完整性、准确性和真实性。

1. 时效性　医疗档案只有做到及时地记录病人的信息，才能动态地显示病情变化，根据需要及时调整治疗方案。在医学研究中，临床数据是对病人当时情况的真实观察与记录，因此必须按照观测记录表格中规定的时点填写各项目的数据。

2. 完整性　包括两方面，一方面是要按临床观测记录表的要求，收集和填写所有项目的数据，形成完整的数据记录表。另一方面是收集规定的全部研究对象的资料，研究人员无权对

入组后的病例随意进行取舍。

3. 准确性 医学研究数据记录的准确性是反映客观情况的关键，临床观测记录表的记录人员应认真学习研究方案中各项目的定义，认真领会其含义，加上本人的专业知识，掌握测量和评价方法，尽量准确地填写，无法把握的问题应及时与研究者沟通信息。

4. 真实性 真实地反映病情是研究的重要原则之一，医学研究的结论是建立在数据真实的基础上的，因此研究记录应做到能够反映当时病人的真实情况。

数据的质量是科学研究工作的生命，在整个研究过程中必须投入相应的人力和物力来保证数据质量，其投入应占科研工作总投入的10%左右。

### 三、数据统计分析

数据统计分析(analysis of data)，是指计算有关的统计指标，以反映数据的综合特征，阐明事物的内在联系和规律。统计分析包括了统计描述和统计推断两方面的内容。

1. 统计描述(descriptive statistics) 是根据研究设计的要求，选用适当的统计指标、统计表、统计图等方法，对资料的数量特征及分布规律进行测定和描述。

2. 统计推断(inferential statistics) 是根据概率分布和抽样的原理，在随机变量的样本信息基础上推断总体特征。统计推断是统计分析的主要内容，包括参数估计和 $t$-检验、方差分析、卡方检验等假设检验的方法。在进行统计分析时，要根据统计设计的目的和要求，选用正确的统计推断方法，对样本资料进行准确的描述和推断，才能得到真实可靠的分析结果。对于一项较复杂的科研，应由比较专业的统计学家来对数据进行统计分析或进行指导，统计分析人员应花费一定的时间研究专业背景资料，对试验设计和研究过程进行了解，对分析方法的选择应事先进行探索和比较，统计分析的投入应占科研工作总投入的5%左右。

## 第三节 怎样学习医学统计学

我们在科研咨询和教学工作中发现，不少医学工作者遇到统计学问题就产生恐惧心理，认为统计学很抽象，难学难懂。有的曾经多次学习统计学，但对统计学仍不得要领。有的学生统计学考试成绩优良，但在从事科研工作时，遇到很简单的实际问题也无法独立解决。

分析其原因，一方面是学科本身的特点所造成的，医学比较注意形象直观思维的培养，注重对“看得见，摸得着”的现象与原因进行关联分析，而统计学比较注意抽象逻辑推理的训练，医学生学习统计学开始时往往不适应。但更主要的原因是学习统计学的方法存在问题，不少学生采用死记硬背公式，花大量的时间重复计算，而对统计学知识的系统性和准确性掌握不够。诚然，我们必须付出一定的努力，花费一定的时间和精力来学习统计计算，但是，在学习方法上我们提出如下建议：

(1)抓住“三基”，即基本概念、基本原理和基本方法。对复杂的公式的推导及公式本身只须要了解一下其作用，而不必死记硬背其具体的形式。在医学科学研究中所应用的统计学知识中约70%是最基本的概念和经典的统计方法，其余则是较为复杂的、近代发展起来的统计理论和技术，而出现错误最多的却偏偏是前一部分。

(2)重视统计应用，使知识转化为智慧和能力。学习统计学一定要结合实例，最好从问题的原形入手，将其转化成统计问题，这是正确使用统计学的关键一步。然后根据设计类型、资

料性质和分析目的，选择合适的统计分析方法进行处理。要经过从理论到实践，再从实践到理论的反复过程，循序渐进，才能逐渐掌握统计学，运用统计学解决实际问题时，才能得心应手。

(3)注重对统计学知识的系统性和准确性掌握。要掌握可供选用的设计方法种类，可供选用的统计分析方法的种类，学习完一个章节、一个阶段时要及时进行归纳和小结。

(4)重视各种检验方法适用的前提条件及应用场合，不必拘泥于具体的计算推导过程。

(5)掌握一种统计学软件，学会正确使用统计软件和正确选择统计方法。

(6)学会对软件输出结果及统计学结果做正确解释。

## 第四节　统计学的基本概念

统计学作为一门独立的学科，有许多专用的术语和概念，本节将介绍统计学中广泛应用的几个基本概念和基础知识，包括同质与变异，变量与随机变量，总体、个体和样本，参数和统计量，参数估计和假设检验，误差及频率、概率等。

### 一、同质和变异

1. 同质(homogeneity)　就是性质相同，它是进行统计分析的前提。统计分析是在一定数量的观察对象的基础上进行的，这一定数量的观察对象在研究的主要方面必须具有相同的性质。比如，研究某地高血压病人的生活质量，研究对象必须是同质的，即都是同一地区的高血压病人。如果不能满足同质的要求，那么研究数据就是杂乱无章的，便不能得出有用的信息和结论。

2. 变异　同质是相对的，研究对象只是在某一方面是性质相同的，同类的观察对象之间往往也存在着变异。比如，同一地区、同一年龄的男童的身高并非完全相同，而是千差万别、参差不齐的，这种参差不齐的情况就是变异(variance)。与同质的相对性不同，变异是绝对的、客观存在的，这在生物学和医学界是非常普遍的，患同类疾病的病人，病情会有轻有重，相同病情的病人用同样的治疗方法治疗，病人的预后也不尽一致。正是因为变异的普遍存在，统计学才得以不断的发展。统计学就是处理变异性的科学，如果生物学界的个体都是完全一样、千篇一律的，统计学就没有存在的必要了。

### 二、个体、总体与样本

1. 个体(individual)　可以是一个人、一个动物、一个家庭、一个地区、一份样品等，是科学研究的基本观察对象(observation unit)。

2. 总体(population)　是性质相同的所有观察对象的某种变量值的集合。如调查某地2004年正常成年男子的血红蛋白含量，观察对象是该地2004年的正常成年男子，观察对象是每个成年男子，研究总体是该地2004年正常成年男子的血红蛋白含量，同质的基础是同一地区、同一年份、同为正常成年男子。总体所包含的范围是随着研究目的的不同而变化的，根据总体中观察对象数是否已知可将总体分为有限总体和无限总体。有限总体(finite population)包括有限个观察对象单位，它是有时间和空间限制的，某地2004年正常成年男子的血红蛋白含量就是个有限总体，因为这个总体在确定的时间和空间范围内包括了有限个观察对象。无限总体(infinite population)是指没有时间、空间限制的无限个观察对象组成的总体，如研究

贫血患者用某种药物治疗的疗效，总体包括了所有用该药治疗的贫血患者的疗效，是没有时间和空间限制的，因而观察对象的数量是无限的，这个总体为无限总体。

3. 样本 医学研究中，很多总体都是无限总体，即使对于有限总体而言，如果该总体所包含的观察对象数过多，要直接研究总体也是不可能和不必要的。所以在研究时经常是从总体中抽取样本，用样本信息来推断总体特征。样本(sample)是从总体中随机抽取的部分观察对象所组成的集合。比如，从北京地区正常成年男子中随机抽取 7 000 人组成样本。抽样的目的是用样本信息来推断总体特征，因此要保证样本的可靠性和代表性，使样本能够充分地反映总体的真实情况。这就要求抽样要遵守随机化的原则，并保证足够的样本含量。随机抽样(random sampling)是指按照随机化的原则抽取观察对象组成样本，以避免研究者和研究对象给样本带来的偏倚。样本量(sample size)是指样本中所包含的观察对象数。

## 三、变量和随机变量

### (一)变量

统计学是研究变异性的，变异性是通过各观察对象的某项特征或指标来反映的，因此我们要研究生物界个体的变异性，就要先确定观察对象，然后对每个观察对象的某项特征或指标进行观察和测量，这种观察对象的特征或指标就是变量(variable)，观察对象中各个变量的观察结果称为变量值(value of variable)或观察值(observed value，observation)，因为测量不同的观察对象会得到不同的观察结果，所以称之为变量。例如以老年人为观察对象，调查研究某地某年老年人的疾病和健康状况，年龄变量的观察结果有大有小，性别变量的观察结果有男有女，血压变量的观察结果有高有低，病情变量的观察结果有轻有重。

根据变量的测量结果不同可以将变量分为数值变量、定性变量和等级变量三大类。一组变量值统称为数据(data)。研究数据根据其性质可以分为定量数据、定性数据和等级数据。

1. 定量变量(quantitative variable) 也叫数值变量(numerical variable)，是用仪器、工具或其他定量方法进行测定或衡量所取得的数据。其变量值是定量的，表现为大小不等的数值，一般带有度量衡单位。如，身高(cm)、体重(kg)、白细胞计数($10^9$/L)、血压(kPa)、龋齿个数等。由一组同质的定量变量值所组成的数据称为定量数据(quantitative data)，定量数据的各个观察值之间有量的区别，没有性质的不同。

2. 定性变量(qualitative variable) 也称为分类变量(categorical variable)，其变量值是定性的，表现为互不相容的类别或属性。各观察对象之间一般没有量的区别，但有质的不同。

如果变量只有相互独立的两种属性，称为二分类变量(binary variable)，如人的性别有男或女，检查乙肝表面抗原的携带情况有阳性或阴性，癌症患者结局有生存或死亡等。如果变量的观察结果表现为相互独立的多种属性，称为多分类变量(polytomous variable)，比如血型分为相互独立的四类：A 型、B 型、O 型和 AB 型，肺癌可分为腺癌、鳞癌、腺鳞癌、未分化癌、类癌和支气管腺癌等，各类之间只有性质的不同，没有大小和程度上的差别。

由一组同质的定性变量值所组成的数据称为定性数据(qualitative data)，定性数据也可以由按照定性变量值的属性分组，然后清点各组的观察对象个数得到，亦称为计数数据(count data)。

3. 等级变量(ranked variable) 也叫顺序变量(ordinal variable)，等级变量可以体现程度上的不同，但是不能精确地测量相邻的两个变量值之间的差别，通常等级变量有两个以上的等

级。比如患者的预后情况可分为治愈、显效、好转、无效、恶化五级，医师对病人的总体疗效评价可分为很好、好、一般、差四级，癌症的病理分级为Ⅰ级、Ⅱ级、Ⅲ级。

由同质的顺序变量值组成的数据称为等级数据(ranked data)，它是介于定量数据和定性数据之间的半定量观察结果。等级数据也可先将观察对象按照各个等级分组，然后清点各组观察对象的数目得到。

**(二)随机变量**

变异是生物个体的共有特征，反映了生物个体的不确定性。在测量观察对象的变量值之前，只知道变量值所在的可能范围，并不知道其具体取值情况，比如只知道 8 岁男童的身高可能在 50～150cm 的范围内，但不能确定某个男童的具体身高，正是因为变量的这种不确定性，概率论中将其称为随机变量(random variable)。

## 四、统计量和参数

**(一)统计量**

在科研工作中，通过对样本中的观察对象的变量值进行统计分析所得到的统计指标称为统计量(statistic)。例如为了调查了解某地 2004 年正常成年男子的血红蛋白含量，随机抽取 240 人组成样本，它们的血红蛋白含量的平均值就是一个统计量，样本回归系数、样本标准差、样本率等也是统计量。

**(二)参数(parameter)**

参数是反映总体特征的统计指标。如果样本的代表性好，那么统计量与相应的参数的数值就非常接近，就可以用样本统计量来估计总体参数，所以样本的统计量也称为参数的估计值。例如用样本均数、样本回归系数、样本标准差和样本率来估计总体均数、总体回归系数、总体标准差和总体率。

参数估计和假设检验是统计推断的两个重要领域。

1. *参数估计*(parameter estimation)　是在总体参数未知时，用样本统计量来估计总体参数，它包括点估计和区间估计。点估计(point estimation)是给出被估计参数一个适当的估计值，即样本统计量；区间估计(interval estimation)是按照预先给定的概率，给出未知参数可能的数值范围。

2. *假设检验*(test of hypothesis)　是先对总体参数或总体分布做出某种假设，如假设两个总体率相等，或总体服从某种分布等，然后用适当的检验方法根据样本信息，推断应当拒绝或不拒绝此假设。根据其假设是针对参数还是分布，假设检验可分为参数检验和非参数检验，参数检验如 $t$ 检验、方差分析等；非参数检验如秩和检验、$\chi^2$ 检验、游程检验等。

## 五、频率和概率

**(一)频率**

1. *定义*　在相同的条件下，进行了 $n$ 次试验，在这 $n$ 次试验中，事件 A 发生的次数 $n_A$ 称为事件 A 发生的频数，$f_n(A)=n_A/n$ 称为事件 A 发生的频率(frequency)，$0\leqslant f_n(A)\leqslant 1$。频率的大小反映了事件 A 发生的频繁程度，频率大，则事件 A 发生就频繁，这意味着 A 在一次试验中发生的可能性就大，反之亦然。

2. *稳定性*　大量试验表明，当重复试验的次数 $n$ 逐渐增大时，频率 $f_n(A)$ 将呈现出稳定