医学统计学及SAS应用

主编／王炳顺

（修订版）

上海交通大学出版社

医学统计学及 SAS 应用

（修订版）

主　编　王炳顺
副主编　宋艳艳

上海交通大学出版社

内 容 提 要

本书基于医学资料实例，介绍了常见的统计学分析方法，着重于统计学基本理论的领悟和统计学思维训练。目标是促进读者理解医学研究资料的数据处理过程和统计学技术的应用，将统计学原理和技术运用到医学科研工作中。本书在介绍常用医学统计方法的基础上，淡化统计计算的复杂过程，使用SAS统计软件包实现统计分析，其中包括运用SAS软件包组织数据，输入数据，建立数据文件，进行统计分析，并正确阅读、解释软件包的输出结果。

本书面向医学生、医生、医学研究者及医药学相关工作人员，体现了实用性的特点，最终目的是促进读者能够对医学科研的实际资料进行综合的统计分析。

图书在版编目(CIP)数据

医学统计学及SAS应用/王炳顺主编. —修订版.
—上海：上海交通大学出版社，2009
ISBN 978-7-313-04844-8

Ⅰ. 医… Ⅱ. 王… Ⅲ. 医学统计—统计分析—应用软件，SAS Ⅳ. R195.1—39

中国版本图书馆CIP数据核字(2007)第083961号

医学统计学及SAS应用
(修订版)
王炳顺 主编
上海交通大学出版社出版发行
(上海市番禺路951号 邮政编码200030)
电话：64071208 出版人：韩建民
上海交大印务有限公司 全国新华书店经销
开本：787mm×1092mm 1/16 印张：29 字数：719千字
2007年8月第1版 2009年9月第2版 2009年9月第2次印刷
印数：3 030
ISBN978-7-313-04844-8/R 定价：49.50元

主　编　王炳顺

副主编　宋艳艳

审　阅　苏炳华 何清波

编　者(按汉语拼音顺序排列)

刘丹萍(四川大学华西公共卫生学院)

罗剑锋(复旦大学公共卫生学院)

宋艳艳(上海交通大学基础医学院)

王柏松(上海交通大学基础医学院)

王炳顺(上海交通大学基础医学院)

王筱金(上海交通大学基础医学院)

张莉娜(上海交通大学基础医学院)

学术秘书　王筱金

前　言

生命现象最突出的特征是它几乎无限的多样性，在有性繁殖的种群中没有两个个体完全相同。由于生物体都存在个体差异，生物医学研究中变异无处不在，而在外在因素的影响下，医学现象更加变化万端，相互关系错综复杂。例如，某病的发生或流行是什么因素所致？可能涉及的多种因素中哪些是确实无关的？哪些是真正相关的？其中又以何者为主？何者为次？又如用某种临床新疗法治疗某种疾病，有的患者治愈了而有的患者却无效？如何客观判断该疗法究竟是否有效？或者与常规疗法相比较时新疗法是否有优势？对于诸如此类医学问题如何正确地开展研究？如何去获取确切而必需的资料？如何对这些资料科学地进行分析，从而得出可靠的判断和结论？医学统计学就是帮助解决这类问题的一个强有力工具。

数理统计学作为数学的一个分支，是公理和定理紧密结合的一个完整的数学系统，涉及概率论、微积分和高等代数等领域。为了使这些理论也适用于医学研究工作者，加强概率论和数理统计方法在各种医学具体问题中的应用，将统计学理论和方法进行简化，在相对简单的水平与医学研究相结合就产生了医学统计学。为此，笔者试图以直观的风格编写此教材，强调的是统计学概念而不是数学细节，不注重这些方法的理论根据、数学论证，不从数学上讨论统计概念和方法，而尽量从直觉水平进行表述，尽可能形象直观地展示统计理论方法。本书编者多为生物医学统计工作者，书中的表述自然反映了我们的医学专业背景。

统计学不是干巴巴的学术理论，不是冷冰冰的复杂公式和数据处理，而是应该渗入到医学研究等各方面的一种思维方式。本教材不是为了使读者成为专业统计学工作者，而是在于促进读者如何从不确定性或概率的角度来思考问题，在开展医学相关研究设计，进行数据的搜集、整理、分析时具备清晰的思路。编者尽其所能帮助读者排除形成这种思维方式时的障碍、减少学习医学统计学时的困难，目的是给读者提供如何面对不确定性的一种思考方法，建立以科学方法开展实验与分析的逻辑观念。

可以说正确使用统计学方法可以使医学研究的结果更真实可信，而且统计学思维至少有助于医学工作者批判性地阅读医学文献，更好地理解文献资料中的数据。

在试图掌握复杂的统计学方法以前，必须先理解简单的方法。我们并不试图囊括全部统计学方法（事实上也不可能），而是介绍一些医学研究中常用且经典的统计学方法与技术。假若将那些受人尊敬的前辈所著的大部头名作比作统计学习的“大餐”，那么这本小书仅仅是一碟“素菜”，编者本着让医学专业读者开胃、易消化的原则，着重介绍统计学基本概念及其思想原理，掌握常用统计方法的实质、特点及应用条件，强调的是实际问题的处理，使统计分析切合问题的重心，进行有效、足够的分析，而不是去追求统计分析的复杂性。

医学专业读者如何将学到的统计学方法应用到实际问题当中？统计学方法的应用需要理解医学研究资料的数据处理过程和统计学相关软件的实际应用。我们不主张死记公式，应淡化统计方法的推导和计算。统计学繁杂的计算不应成为统计学方法应用的障碍，为此掌握必

要的统计软件的使用技能已成为必不可少的重要环节。本教材各章节内容的统计计算都交付给权威的SAS统计软件去完成，本教材所用SAS参考程序可以发Email到sas4biostart@gmail.com获取下载链接。

本教材是在前辈史秉璋、苏炳华、何清波等教授所编写的《医学统计学及其软件包》等系列教材基础上，结合各位青年教师“医学统计学”相关课程教学实践和实际教学需要汇编而成。参加编写的教师参阅了大量中外文书籍、借鉴了许多(医学)统计学界前辈、同道学者出版的有关文献以及网络共享的一些学术资源，并直接引用了一些经典性论述和例子。我们尽量标注参考文献出处，然而难免挂一漏万，所标注的参考文献只是其中的一小部分，编者向本教材引用材料的所有作者、编者、译者表示诚挚谢意。

本教材编写(包括SAS软件版权的购买)得到了上海交通大学医学院研究生课程建设项目(编号：YKC0506，“211工程”建设经费)的大力资助，同时这项工作也得到了上海高校选拔培养优秀青年教师科研专项基金、上海市教育委员会科研项目(编号：06BZ007)、上海交通大学医学院基金(编号：05XJ21003)的支持，特此致谢。

衷心感谢所有参与编写的各位教师，大家为本教材的编写付出了辛勤的劳动、贡献了智慧和经验。教材各章节的后面都附了编者姓名，其余由主编统筹完成，其中张莉娜老师核对验证了全部SAS程序，王筱金老师作为学术秘书做了大量细致工作。

还要特别感谢苏炳华教授、何清波教授，两位前辈在百忙中为审阅本书稿付出了大量心血。

可以说我们是怀抱着编好一本教材的良好愿望，尽最大努力完成了书稿，然而学识所限，书中谬误在所难免(勘误表将会放在前述网站)，同行和读者的批评指正将是我们最大的礼物，由此将鞭策我们不断再版加以充实、完善。

编者　谨识

2009年8月

目　录

第一章 绪 论

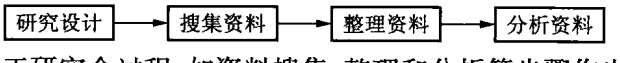
第一节 医学统计学概述

统计学(statistics)是研究如何有效地搜集、整理和分析带有随机性的数据,以对所考察的问题作出推断和预测,直至为采取一定的决策和行动提供依据和建议的科学。统计学方法已成为科学研究和管理工作的重要工具。医学统计学是结合医学实际需要,运用概率论和数理统计学的原理和方法,开展医学研究设计,进行数据资料的搜集、整理、分析和推断的一门学科。

医学研究的对象是功能复杂的有机生命体。不同的个体在相同的条件下,对外界环境因素可以发生不同的反应,这种同质基础上个体特征值之间的差异,称为变异(variation)。而存在变异的现象正是统计学研究的对象。医学及其相关学科实践性极强,不可能完全脱离实验而仅仅依靠逻辑推理去获取新的知识,而单个实验所得到的结果几乎都带有或多或少的不确定性。统计学的介入可以帮助解决如何从这样一些不确定性中得出科学可靠、相对确切结论的问题。而且在科研工作中,常常必须根据有限的、不完全的信息作出评价或决策。例如,评价某年某地区儿童发育情况、某种新药对某病疗效如何等。限于人力、物力、时间等条件,研究人员不太可能调查到该地区所有儿童的发育情况,也不可能接收患该病的所有患者来研究该新药的疗效,仅仅能抽取有代表性的个体组成的集合来深入研究,这样获得的信息显然是有限、不完全的,这类问题需要用抽样研究(sampling research)来加以解决。统计学提供了理论和方法支持,使我们不仅能做出合理的判断与决策,而且知道判断与决策所承担风险的大小。

医学统计学的主要内容有统计研究设计、统计描述、统计推断、研究联系、分类和检测等。本章将就医学统计学作一概要性介绍,后续章节将会陆续介绍医学科研实践中常用的统计学方法,届时可以回顾本章以加深对统计学基本思想及相关概念的理解和领会。

统计工作一般经历以下几个主要步骤:



(1) 研究设计:对于研究全过程,如资料搜集、整理和分析等步骤作出总的设想和安排,是开展研究工作应遵循的依据和获得科学研究结论的前提。

(2) 搜集资料:按照统计研究设计的要求搜集资料,取得准确可靠的原始数据。需注意选择合适的指标,资料应尽可能保持完整,对于缺失值(missing)须有合理的说明等。

(3) 整理资料:根据研究设计的规定对原始资料进行检查整理、分组列表等。

(4) 分析资料:对资料进行统计分析,包括统计描述和统计推断两方面内容。

一、统计研究设计

医学研究开始阶段要制定研究计划。良好的研究计划除了要从所研究问题的专业特点考虑之外，还要从统计学角度进行考虑。即先理清专业问题，形成研究假说，再经过合理选择量化指标等过程将研究假说转化为统计假设，围绕假设，以较少的人力、物力和时间取得较多的、可靠的信息，使得搜集的资料和统计学检验能够回答所研究的问题。统计研究设计应当遵循3个基本原则：对照原则、重复原则和随机化原则。统计研究设计具体可分为两大类：调查研究设计和实验研究设计。调查研究又称观察性研究，只能对随机抽取的研究对象作被动观察，而不能对观察对象施加干预。例如，调查某地高血压的患病率及其影响因素。实验研究则人为设置处理因素或水平，受试对象接受何种处理因素或水平是由随机分配而定的。例如，比较两种药物治疗某疾病的疗效和安全性的临床试验，将研究对象随机分配到不同药物治疗组，观察比较各处理组的结果。

二、资料类型与搜集整理

（一）资料分类

资料一般可分成三大类，即计量资料、计数资料和等级资料。

1. 计量资料(measurement data)

计量资料又称为定量资料(quantitative data)，它是用度、量、衡等计量工具直接测定获得的每个观察单位某项指标值的大小，它有计量单位。根据各个观测值之间的变异是否连续性，分为连续型资料(continuous data)或离散型资料(discrete data)两类。连续型资料包括身高、体重、体温等；离散型资料包括正常人每分钟的心脏跳动次数、每个家庭现有的人口数、一年内的死亡人数等。

2. 计数资料(enumeration data)

计数资料又称为定性资料(qualitative data)，将观察单位按某种属性或类别用计数方式得到的资料，这些观察值只能以整数来表示。如调查1 483例居民，发现钩虫感染者144例、未感染者1 339例，这就是一个计数资料。

3. 等级资料(ranked data)

等级资料又称为半定量资料(semi-quantitative data)，它是将观察单位按某种属性的不同程度分组计数的资料。这类资料既有计数资料的特点，又有程度或量的不同。例如，用某药治疗慢性肾炎102例，其中无效49例，好转30例，显效23例，这就是一个等级资料。

不同的资料类型有其相应的统计学处理方法。有时可根据研究目的和统计处理的需要可以进行资料类型的转换。例如，年龄是计量资料，有时需将年龄划分成几个年龄段，这时就成为等级资料，而当需要划分为两个组别如老年组与非老年组时又成了计数资料。

（二）资料搜集

根据研究目的，按照研究设计开展相关的信息搜集，其中所确定的结局指标与研究目的应有本质的联系。例如，该指标能够确切反映处理因素的作用。资料搜集一般借助于调查表、报告卡、统计报表等原始记录用表格，原始数据应尽可能获取细致的信息。项目具体开展时要争取在较短的时间内、用尽可能少的投入获取高质量的研究资料，同时要开展质量控制以确保资料准确、完整，保证所收集的数据能充分反映研究对象的真实情况。

（三）资料整理

对于所获取的原始数据要进行审核，进行数据清理、检查、核对与纠错，通过归纳汇总使之系统化、条理化。数据量小时可以手工处理，当记录多、数据量大时需要借助计算机工具将原始数据数量化编码录入数据库。数据量大且要求严格时则一般要进行独立双遍录入，核对两遍录入的数据，找出不一致者，根据原始资料进行数据库修改确认。经核对后进行逻辑检查，以保证数据的准确可靠。

三、统计描述

统计描述(statistical description)是指将研究数据加工提取，用统计指标、统计表、统计图等方法，对资料的数量特征及其分布规律进行测定和描述。一个统计问题所涉及的对象的全体称为总体(population)，总体中每一个研究对象即观察单位(observed unit)称为个体(individual)。开展研究的最终目的是要了解总体的数量特征及其规律性，如果在研究中可以得到总体中的每个个体资料，那么只进行统计描述就够了。

四、统计推断

实际研究工作中，受条件所限，在研究中很难得到整个总体，往往只能得到总体中的一个子集。即实际工作中往往按随机的方式从总体中抽取若干有代表性的同质个体所构成的一个样本(sample)进行研究[①]，这就需要通过样本有限的、不确定的信息来推论有关总体的特征，这就是统计推断(statistical inference)。简言之，统计推断是指由样本所提供的信息对总体数量规律性做出推断。

为了描述总体和样本的数量特征，需要计算出几个特征量。由总体计算所得的特征量称为参数(parameter)；由样本资料计算所得的特征量叫统计量(statistic)。总体参数一般是未知的，参数常用希腊字母表示，例如后续章节将要学习的总体均数 μ、总体标准差 σ、总体率 π、总体回归系数 β、总体相关系数 ρ 等。统计量常用拉丁字母表示。例如样本均数 $\bar{x}$、样本标准差 s、样本率 p、样本回归系数 b、样本相关系数 r 等。在总体确定的情况下，总体参数是固定的常数，而样本统计量是样本观测值的函数，在总体参数附近波动。

统计推断主要是通过统计量来实现的。统计推断分为两个部分：参数估计和假设检验。

1. 参数估计(estimation of parameter)

根据研究目的从相应总体中随机抽取样本进行研究，由样本统计量估计总体分布中的未知参数。参数估计可分为点估计和区间估计。

选择一个适当的样本统计量作为总体参数的估计值称为点估计(point estimation)。点估计方法是用一个确定的值去估计未知的参数。由于估计量是来自一个随机抽取的样本，不同的样本就会有不同的估计量，一个样本估计量恰好等于总体参数(某未知的常数)的可能性极小。由于个体间存在变异性，在抽样研究中样本统计量与总体参数的差别称为抽样误差(sampling error)。

由于抽样误差不可避免，或者说估计值(统计量)不可能正好等于真值(参数)。估计值与

① 样本中所包含的同质个体的数目称为样本含量(sample size)，简称样本量，一般在研究设计阶段根据相关设定条件进行估计。

真值近似程度到底是多少，点估计中没有提供任何信息，因此，在点估计之外最好能给出估计精度，即将抽样误差考虑在内。在一定把握程度下估计出总体参数处于某一个小区间内，则更能说明问题。因而，根据一定的正确度和精确度要求，确定一个概率水平，由样本统计量计算出一个适当的区间作为未知总体参数真值所在的范围，称为区间估计(interval estimation)，称此概率水平为置信度，简称信度，也可称为置信水平(confidence level)。所估计的区间称为置信区间 (confidence interval)。也有人将置信区间称为可信区间。区间的端点称为可信限(confidence limit)，有上限、下限之分。

区间估计给出的信息显然多于点估计。例如，从患某病的患者总体中随机抽得 n 例患者进行治疗，治愈 x 例，则可得样本治愈率，对于总体治愈率的点估计、区间估计结果如表 1.1 所示(具体的计算，详见第八章)。

表 1.1 从样本率对其总体率的估计

	样本含量 n	治愈例数 x	样本治愈率/%	总体治愈率/%		
				点估计	95%置信区间	99%置信区间
样本 1	10	5	50	50	19～81	13～87
样本 2	100	50	50	50	40～60	37～63
样本 3	1000	500	50	50	47～53	46～54

由表 1.1 可见，置信区间的大小与样本含量及置信度的大小有关，随着置信度的加大，置信区间也加大，随着样本含量的加大，置信区间缩小。

2. 假设检验(hypothesis testing)

假设检验又称显著性检验(significance testing)，是统计推断的另一种基本形式，是统计分析中的主要内容。假设检验先对总体的参数或分布作出某种假设，然后用适当的方法，根据样本对总体提供的信息推断是否拒绝该假设。其结果将有助于研究者作出具体判断和选择。例如，一项临床实验要比较两种药物治疗某疾病的疗效，研究者获得的样本资料显示两种药物疗效不同。产生差异的原因是什么？①可能是由于进行比较的处理间事实上就有实质性的差异；②可能是由于无法控制的偶然因素所引起。假设检验的目的就在于承认并尽量排除这些无法控制的偶然因素的干扰，将处理间是否存在本质的差异揭示出来，那么研究者的目标就是要区分事实和偶然性，只有证实实验表现出来的效应显然不是偶然性波动所致，才能合乎逻辑地作出正确的结论。

假设检验的方法很多，常用的有 t 检验、方差分析、卡方检验等，后续章节将会逐一介绍。假设检验的一般过程如下：

1) 先对总体的参数或分布作出某种假设。例如，两个总体均数相等、两总体治疗有效率相同、两总体分布相同等。假设检验中将假设分为两种：①检验假设(null hypothesis)，也称为无效假设，用 H_0 表示；②对立假设或备择假设(alternative hypothesis)，用 H_1 表示。H_0 与 H_1 是相互联系、相互对立的假设[①]。

① 一般情况下，无效假设 H_0 是研究者期待证伪的假设，而备择假设 H_1 是研究者期望证实的假设。逻辑上证实一项假设不可行，那么就通过反证法来拒绝 H_0，从而接受 H_1。

2）然后选择适当的样本统计量，在 H_0 成立的情况下计算所得概率 P 值的大小（P 值可以理解为 H_0 成立时得到目前研究结果甚至更极端情况的可能性），以此决定究竟是拒绝 H_0，还是不拒绝 H_0，完成统计推断。

假设检验的基本步骤为：

（1）建立 H_0，H_1。

（2）选择合适的统计检验方法，计算统计量。

（3）根据检验统计量的分布，直接计算概率 P 值，或者将检验统计量与检验水准 α 相应的临界值进行比较，根据 P 值与 α 的大小关系进行判断：

如果 $P>0.05$，则在 $\alpha=0.05$ 水平上，不拒绝 H_0；

如果 $0.01<P\leqslant 0.05$，则在 $\alpha=0.05$ 水平上，拒绝 H_0；

如果 $P\leqslant 0.01$，则在 $\alpha=0.01$ 水平上，拒绝 H_0。

由上述内容可知，统计学的目的是探索总体的数量规律性。统计方法的精髓是通过随机样本信息对总体特征作出科学的推断。统计方法中统计描述与统计推断的关系、统计学探索客观现象数量规律性的过程可以总结为如图 1.1 所示。

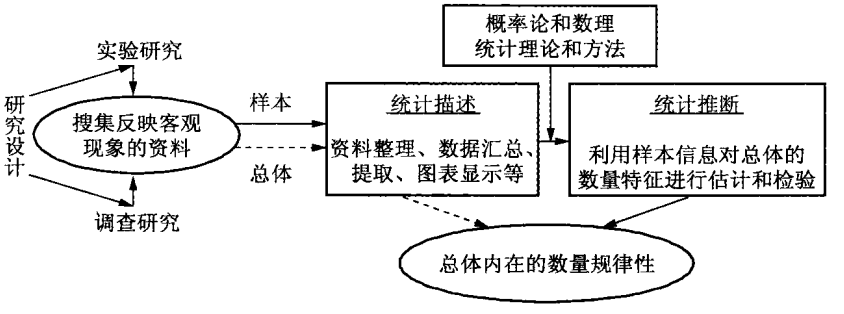


图 1.1　统计学探索现象数量规律性的过程

第二节　概念与术语

一、随机现象、随机事件与随机变量

在物质世界、社会生活中发生的现象是多种多样的，归结起来大致可分为两大类：确定性现象（又称必然现象）和不确定性现象（又称偶然现象，亦称为随机现象）。确定性现象包含必然事件和不可能事件。这类现象是在一定条件下，必定会导致某种确定的结果。例如：在标准大气压下，100℃的纯水必然沸腾。确定性现象其结果可以事先预言，这种没有变异的现象不是统计学研究的对象。

实际上，另一类客观现象即随机现象在现实生活中更为普遍。所谓随机现象，就是在基本条件不变的情况下，各次实验或观察可能会得到不同的结果，而且无法准确地预测下一次所得结果的现象。例如，用同一种药物治疗患者，由于个体差异等原因，有的患者治疗有效，而有的患者治疗无效。这种结果的不确定性，是生物个体变异性及其他一些偶然的因素影响所造成的。

对于某个现象，如果能让其条件实现一次，就是进行了一次实验。而实验的每一种可能的结果，都是一个事件，将随机现象的每种结果称为随机事件。随机事件的数值性描述称为随机变量(random variable)①，简称变量。例如，抛掷一枚硬币，其结果可用一个随机变量 X 来描述，若用数值 1 表示正面朝上，0 表示反面朝上，由于实验的观察结果不能事先确定，则掷硬币之前可以说实验结果变量 X 可能取 0，也可能取 1，即随机变量的数量化取值与事件相对应。随机变量分为两类(如图 1.2 所示)：①离散型随机变量(discrete random variable)，即仅取数轴上有限个点或可列个点；②连续型随机变量(continuous random variable)，即所有可能取值充满数轴上一个区间$[a,b]$，$a,b\in(-\infty,\infty)$。前者如某药治疗患某病的 n 个患者，其治疗有效例数 X：随机变量 X 可能的取值为 $0,1,2,3,\cdots,n$；后者如正常成年男子的身高 Y：随机变量 Y 可能的取值处于区间(100cm，300cm)。

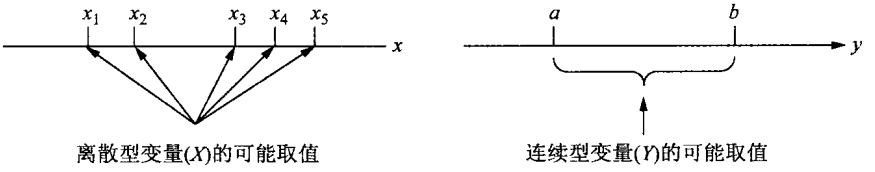


图 1.2　离散型随机变量及连续型随机变量示意图

二、概率与频率

在一定条件下，随机事件可能发生也可能不发生，需要知道的不仅仅是可能会发生哪些结果，我们更感兴趣的是各结果即随机事件发生可能性的大小，即事件发生的概率。

概率(probability)表示一个事件在一次试验或观测中发生的可能性大小。直观上，将某事件记为 A，我们用一个数 $P(A)$ 来表示随机事件 A 发生可能性的大小，$P(A)$ 就称为 A 的概率。概率是在 0～1 之间的一个数，概率为 0 时表示事件不会发生，概率为 1 时表示事件必定发生。

在相同的条件下，独立重复做 n 次试验，随机事件 A 发生了 m 次，则比值 m/n 称为随机事件 A 在 n 次试验中出现的频率(freqency)，计为 $f(A)=m/n$。一般情况下，当实验次数 n 越来越大，直至 $n\to\infty$时随机事件 A 发生的频率 $f(A)=m/n$ 趋向一个常数 π，我们将这个常数 π 称为随机事件 A 发生的概率 $P(A)$，即我们利用实际频率数据(m/n)来估计概率 $P(A)$，这就是概率的统计定义。

例如：掷一枚制作均匀的硬币，抛出去之前预先并不知道结果会是什么，每实验一次有两个可能的结果："正面"、"反面"两个不同的事件。我们掷币 10 次后可以总结"出现正面"的次数是多少，当重新再来 20 次或 100 次可以总结"出现正面"的次数是多少，历史上很多人做过掷币试验(表 1.2)。结果表明随着重复掷币次数的增加，出现正面这一随机事件(A)的频率在 0.5 附近波动，当实验次数越多，一般波动会越小，出现正面的可能性越来越接近于一个常数：50%。因而，我们可以说在一次掷硬币实验中"出现正面"这一事件的概率为 50%。

---

① 医学应用中习惯上将随机变量称为指标。随机变量常用大写字母 X、Y、Z 等表示，随机变量的取值常用小写字母 x $(x_1,x_2\cdots)$、y、z 等表示。

表 1.2　掷币实验

实验者	掷币次数 n	正面次数 m	频率 $f(A)=m/n$
蒲　丰	4040	2048	0.5069
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005
维　尼	30000	14994	0.4998

掷币实验结果很好地反映了多次重复的随机实验中的频率趋于稳定性的特点，表明了随机事件发生的可能性大小是随机事件本身固有的一种客观属性，说明随机现象有其偶然性的一面，更有其必然性的一面。这种必然性表现为大量观察或试验中随机事件发生频率的稳定性，这种规律性称为随机现象的统计规律性，即我们主要依靠频率稳定性来数量化刻画随机现象的内在规律。例如，前述例子中用同一种药物治疗患者，尽管对于不同患者疗效具有不确定性，而当观察一定数量的治疗病例后，通过有效例数(m)与总治疗例数(n)之比，我们可以估计该药治疗的有效率。

注意：概率是一个确定的数值，而频率是大量试验的结果。频率具有随机性，是一个随着试验次数变化而变化的数值，它随着试验次数的无限增加，以一种趋势无限接近概率。

三、小概率原理

随机事件的概率表示了随机事件在一次试验中出现的可能性大小。若随机事件发生的概率很小，例如<0.05、0.01、0.001，则称之为小概率事件。人们积累的大量实践经验表明：当事件发生的概率接近 100%时，在一次试验中几乎一定会发生。同时，当事件发生的概率很小，那么可以认为小概率事件在一次试验中该事件实际上不可能发生，在统计学上称为小概率事件实际不可能性原理，亦称为小概率原理。如果小概率事件在一次试验中居然发生了，我们就有理由怀疑该事件是小概率事件的正确性，或者说有较充分的理由怀疑导致这一小概率事件发生的前提条件的正确性。小概率事件实际不可能性原理是统计学上进行假设检验的基本依据。

那么这个概率小到什么程度人们才能将一小概率事件接受为一次试验中实际不可能事件？即小概率的设定标准应该是多少？著名的英国统计学家 R. A. Fisher 把 1/20 作为标准，也就是 0.05，这种惯例沿袭了下来。于是，在生物医学研究中常常称 $P\leqslant 0.05$（或者 $P\leqslant 0.01$）的事件为小概率事件。

四、Ⅰ类错误与Ⅱ类错误

由前述假设检验内容可知，统计学假设检验是先对总体的参数或分布作出某种假设即 H_0，再根据检验统计量的分布，由样本信息计算概率 P 值。若 $P\leqslant 0.05$，由实际原理推断：有理由怀疑导致这一小概率事件发生的前提条件有问题，即认为原假设 H_0 是错误的。于是，就