



全国高等教育医药经管类规划教材

**Pharmaceutical  
Managerial Statistics**

**医药管理统计学**

主编 邱家学 席晓宇

中国医药科技出版社

全国高等教育医药经管类规划教材

---

# 医药管理统计学

---

主 编 邱家学 席晓宇  
编 者 (按姓氏笔画排序)  
田 磊 朱 玄 李伟霞  
李树祥 邱家学 武兴连  
茅宁莹 席晓宇 褚淑贞

中国医药科技出版社

## 内 容 提 要

本书是编者在多年医药管理学、统计学教学实践和研究成果的基础上,所编写的与医药产业紧密结合的管理统计学基础教程。内容包括医药数据处理、医药统计方法、数理统计原理等知识及常用统计方法等,并在每章最后部分对 Excel 软件对应统计功能的操作应用,辅之适当的课后习题。同时每章开始部分精选综合性典型案例进行分析,并配有学习目标、要求以及内容总结,使本教材能兼顾本科基础课教学和更高的学习要求(如考研)。全书针对医药经管类本科学生的基础和培养要求,适当选取教材的深度和广度,内容系统全面,例题案例结合医药产业实际,典型实用,编写力求简明易懂,深入浅出,富有启发性,适用性强。

本书主要用作医药类各专业概率统计、数理统计、应用统计等基础课教材或参考书,也可作为各类非理工科专业同类课程的参考书,还可以作为其他教学,如考研复习辅导及数学建模等的参考书。同时也是医药卫生工作者颇为实用的统计应用参考书。

### 图书在版编目(CIP)数据

医药管理统计学/邱家学,席晓宇主编. —北京:中国医药科技出版社,2015.2

全国高等教育医药经管类规划教材

ISBN 978 - 7 - 5067 - 7272 - 3

I. ①医… II. ①邱… ②席… III. ①医学统计 - 统计学 - 高等学校 - 教材  
IV. ①R195.1

中国版本图书馆 CIP 数据核字(2015)第 014103 号

美术编辑 陈君杞

版式设计 郭小平

出版 中国医药科技出版社

地址 北京市海淀区文慧园北路甲 22 号

邮编 100082

电话 发行:010-62227427 邮购:010-62236938

网址 www.cmstp.com

规格 787×1092mm<sup>1</sup>/<sub>16</sub>

印张 26<sup>1</sup>/<sub>4</sub>

字数 473 千字

版次 2015 年 2 月第 1 版

印次 2015 年 2 月第 1 次印刷

印刷 三河市国英印务有限公司

经销 全国各地新华书店

书号 ISBN 978 - 7 - 5067 - 7272 - 3

定价 55.00 元

本社图书如存在印装质量问题请与本社联系调换

# 前 言

近几年来,我国医药高等教育不断深入改革和发展,高校对医药人才的培养模式和要求有了极大的变化,同时也对我国医药经管类相关专业的统计及应用课程的教学提出了更高更新的要求。为此,我们通过对国内外管理统计及应用课程的教材和内容等进行深入的分析研究,同时结合学科发展动态,医药领域管理统计应用的特点要求和多年教改和实践经验,对课程结构、内容、教学模式和教材等各方面进行了有益的改革探索,取得的成果显著。在此基础上,我们编写了具有医药行业特色的《医药管理统计学》。

统计学是研究随机现象统计规律性的学科。本教材的编写既考虑到统计学知识结构的科学性和系统性,又结合医药管理领域对统计应用的具体要求和特点,同时针对医药本科生的基础和培养要求,对教材的内容的深度和广度选取适当,并结合时代发展的特点,内容系统而全面,案例典型实用,编写简明易懂,深入浅出。其主要特点是:

作为医药类相关专业基础课教材,在尽量保持统计学的科学性和系统性的基础上,不片面追求理论的推导和证明,而强调理论与实际的结合,体现了学以致用目的,并充分考虑更高层次(硕士研究生教育)的学习要求。

所选内容涵盖统计基础、医药应用领域数据处理和统计分析的基本原理、基本知识和常用统计方法,在系统、简明介绍统计学知识的基础上,以统计数据的处理和分析为核心,注重统计方法思想和实际医药应用的说明,结合数据和医药专业应用实例说明统计方法的特点,形成以概率论基础、统计原理、统计方法及统计软件应用为主体并面向医药领域实际应用的内容体系。

强化以计算机应用为基础的统计技能的培养。现代医药领域数据处理和统计分析离不开计算机统计软件的应用,满足医药院校本科学生所用的软件普及、实用的要求,本教材选用了最为常用的 Microsoft Office 系统的 Excel 软件统计模块来进行教学,操作指导具体详实,便于自学,从而真正提高读者运用软件分析和解决实际问题的操作能力,达到学以致用目的。

本书内容系统全面,阐述深入浅出,用例经典实用,概括高度精炼,并兼顾本科基础课程教学和更高学习要求不同层次的需求,使之不仅适用于医药类各相关专业统计、

数理统计等基础课程教材或参考书，而且还可用于农林经管类非理工科专业同类课程的参考书，并可作为其他教学，如考研辅导及数学建模等的参考书。书中极具实用性的医药统计的 Excel 软件操作应用，是从事医药研究和工作的相关人员不可多得的统计应用参考书。

本教材共分 14 章。具体写作分工如下：第 1~6 章由邱家学教授编写；第 7~14 章由席晓宇编写；另外褚淑贞教授、茅宁莹副教授、李树祥老师、田磊老师以及研究生朱玄、李伟霞也参与了部分章节的编写，武兴连负责全书核对。

在本教材的编写过程中得到了有关专家学者的关心和帮助，并参考了大量的教材和文献，在此表示衷心的感谢。由于时间和水平有限，书中疏漏和不妥之处在所难免，恳请各位读者批评指正。

编者

2014 年 9 月

# 目 录

|                             |       |
|-----------------------------|-------|
| <b>第一章 总论</b> .....         | (1)   |
| 第一节 医药管理统计学及其应用 .....       | (2)   |
| 第二节 医药统计数据的类型 .....         | (5)   |
| 第三节 医药管理统计学中的基本概念 .....     | (7)   |
| 第四节 Excel 简介及基本操作 .....     | (13)  |
| <b>第二章 医药管理数据统计调查</b> ..... | (20)  |
| 第一节 原始医药管理数据的搜集 .....       | (21)  |
| 第二节 医药管理数据调查问卷 .....        | (28)  |
| 第三节 二手医药管理数据的搜集 .....       | (35)  |
| 第四节 医药产品市场调研方法学 .....       | (37)  |
| 第五节 怎样使用 Excel 的“帮助” .....  | (41)  |
| <b>第三章 医药管理数据统计整理</b> ..... | (46)  |
| 第一节 数据的预处理 .....            | (47)  |
| 第二节 品质数据的整理与图示 .....        | (49)  |
| 第三节 数值型数据的整理与图示 .....       | (54)  |
| 第四节 数据统计表 .....             | (62)  |
| 第五节 Excel 统计整理 .....        | (64)  |
| <b>第四章 医药管理统计综合指标</b> ..... | (76)  |
| 第一节 医药管理统计总量指标和相对指标 .....   | (77)  |
| 第二节 医药管理统计集中趋势的描述 .....     | (87)  |
| 第三节 医药数据离散程度的描述 .....       | (96)  |
| 第四节 医药管理数据分布偏态与峰度的测度 .....  | (102) |
| 第五节 Excel 进行综合指标的计算 .....   | (105) |
| <b>第五章 统计概率与概率分布</b> .....  | (113) |
| 第一节 统计事件及其概率 .....          | (114) |
| 第二节 离散型随机变量的概率分布 .....      | (122) |

|             |                                 |       |
|-------------|---------------------------------|-------|
|             | 第三节 连续型随机变量概率分布 .....           | (129) |
|             | 第四节 Excel 进行常用分布的概率计算 .....     | (134) |
| <b>第六章</b>  | <b>抽样调查与统计推断</b> .....          | (144) |
|             | 第一节 抽样调查概述 .....                | (145) |
|             | 第二节 抽样推断 .....                  | (145) |
|             | 第三节 大数定律与中心极限定理 .....           | (146) |
|             | 第四节 三种不同类型的分布 .....             | (149) |
|             | 第五节 一个总体参数推断时样本统计量的抽样分布 .....   | (151) |
|             | 第六节 两个总体参数推断时样本统计量的抽样分布 .....   | (157) |
|             | 第七节 Excel 进行抽样调查 .....          | (160) |
| <b>第七章</b>  | <b>参数估计</b> .....               | (165) |
|             | 第一节 参数估计的一般问题 .....             | (166) |
|             | 第二节 一个总体的参数区间估计 .....           | (172) |
|             | 第三节 两个总体的参数区间估计 .....           | (179) |
|             | 第四节 样本容量的确定 .....               | (183) |
|             | 第五节 Excel 进行参数估计 .....          | (187) |
| <b>第八章</b>  | <b>假设检验</b> .....               | (197) |
|             | 第一节 假设检验的一般问题 .....             | (198) |
|             | 第二节 总体均值、比例的假设检验 .....          | (203) |
|             | 第三节 Excel 进行假设检验 .....          | (206) |
| <b>第九章</b>  | <b>时间序列分析</b> .....             | (215) |
|             | 第一节 时间序列的有关概念 .....             | (216) |
|             | 第二节 时间序列的因素分析 .....             | (218) |
|             | 第三节 Excel 进行时间序列分析 .....        | (240) |
| <b>第十章</b>  | <b>统计指数</b> .....               | (252) |
|             | 第一节 指数基本概念 .....                | (253) |
|             | 第二节 指数的构造方法 .....               | (254) |
|             | 第三节 指数的应用与调整 .....              | (262) |
|             | 第四节 几种常用的经济指数 .....             | (265) |
|             | 第五节 Excel 统计指数分析 .....          | (271) |
| <b>第十一章</b> | <b>相关与回归</b> .....              | (279) |
|             | 第一节 相关分析 .....                  | (281) |
|             | 第二节 一元线性回归分析 .....              | (288) |
|             | 第三节 残差分析 .....                  | (302) |
|             | 第四节 Excel 进行相关分析和一元线性回归分析 ..... | (306) |
| <b>第十二章</b> | <b>多元线性回归分析</b> .....           | (314) |
|             | 第一节 多元线性回归模型 .....              | (316) |

|             |                                     |       |
|-------------|-------------------------------------|-------|
| 第二节         | 多元线性回归方程的评价 .....                   | (318) |
| 第三节         | 多重共线性 .....                         | (322) |
| 第四节         | 利用回归方程进行估计与预测 .....                 | (327) |
| 第五节         | Excel 进行多元线性回归分析 .....              | (334) |
| <b>第十三章</b> | <b>综合评价分析</b> .....                 | (342) |
| 第一节         | 综合评价分析的概念与构成要素 .....                | (343) |
| 第二节         | 综合评价分析的特点与一般步骤 .....                | (345) |
| 第三节         | 评价指标选择与数据预处理 .....                  | (347) |
| 第四节         | 评价结果的综合 .....                       | (353) |
| 第五节         | 几种常用的综合评价方法 .....                   | (359) |
| 第六节         | SPSS 软件简介 .....                     | (367) |
| <b>第十四章</b> | <b>医药管理统计学在日常生活和学习中的实践与应用</b> ..... | (372) |
| 第一节         | 医药管理统计学课程设计教学大纲 .....               | (372) |
| 第二节         | 医药管理统计学课程设计调查分析报告及评析 .....          | (374) |
| 第三节         | 课程设计题目参考 .....                      | (389) |
| <b>附录</b>   | <b>常用统计表</b> .....                  | (391) |
| 附表 1        | 随机数表 .....                          | (391) |
| 附表 2        | 二项分布表 .....                         | (394) |
| 附表 3        | 泊松分布表 .....                         | (398) |
| 附表 4        | 标准正态分布函数表 .....                     | (401) |
| 附表 5        | $t$ 分布临界值表 .....                    | (403) |
| 附表 6        | 卡方分布表 .....                         | (406) |
| 附表 7        | $F$ 分布表 .....                       | (409) |



## 第一章

## 总 论

统计的思维方法，就像读和写的能力一样，有一天会成为效率公民的必备能力。

——英·威尔斯



### 医药管理统计学应用：6 $\sigma$ 在上海卡乐康的实施

医药管理统计学在医药众多领域都有应用，生产过程中医药产品的质量控制是这些领域之一。在统计学中， $\sigma$  代表标准差，反映了数据之间的差异程度，在符合正态分布的数据中，6 $\sigma$  是指偏离正态分布的中心 6 个标准差。在药品生产线上，6 $\sigma$  则意味着生产 100 万件药品，出现不合格药品的数量平均不超过 3.4 件。当今，6 $\sigma$  是企业管理的一种重要手段，其本质是指通过设计、监管每一道生产工序和业务流程，以最少的投入和损耗赢得最高的客户满意度，从而提高企业的利润。

上海卡乐康包衣技术有限公司是美国著名的卡乐康公司在中国投资的一家专业生产药片薄膜包衣的公司。在进入中国市场后，它给中国制药企业带来了美国的高科技产品和服务。2003 年 4 月，上海卡乐康公司聘用了美国著名的 6 $\sigma$  管理专家 Ben，并将 6 $\sigma$  管理理念引入到了公司的战略层面。随后，将公司内所有的部门经理和骨干员工分为 6 个小组，在 Ben 的指导下，运用统计学知识并结合日常工作实践，对公司的流程进行分析，并从各项活动中挑选出 7 个较有价值的项目进行研究。2003 年 9 月，他们向公司管理层所做的项目汇报的结果显示这 7 个项目都取得了良好的成果，起到了以项目带动学习的作用，并获得了一定的经济效益，同时为公司自行开展 6 $\sigma$  管理活动奠定了良好的基础。2004 年，在王学东总经理的大力推行下，卡乐康公司继续开展 6 $\sigma$  活动。并成立了 6 $\sigma$  小组，由小组负责组织和推动公司的 6 $\sigma$  管理活动。同时，公司还进行了平衡积分卡管理，在确定战略目标后，他们采用 6 $\sigma$

的方法来解决平衡积分卡中的 KPI。2005 年,上海卡乐康在 6 $\sigma$  小组的基础上,扩大了 6 $\sigma$  活动的规模,将公司质量、生产、技术、客服等大多数部门的员工都组织起来参加 6 $\sigma$  知识培训,并要求各部门经理、主管必须通过 6 $\sigma$  知识培训,从而在公司内部掀起了一股“学 6 $\sigma$ 、做 6 $\sigma$ ”的热潮。并且在每月的管理会议上,总经理要亲自主持 6 $\sigma$  的汇报。此外,公司还定期组织全体学员进行交流,分享学习 6 $\sigma$  的技巧,启发员工的新观念及创意。通过引进 6 $\sigma$  管理理念并将其融入到企业文化中,上海卡乐康公司明显提高了员工的工作积极性、提高了生产效率,并极大地提高了卡乐康在客户心中的满意度。

**【学习目标】** 本章主要介绍医药管理统计学的含义、研究对象、应用领域及数据分类等内容。通过本章的学习使读者从总体上对统计学有基本的认识,了解医药管理统计学的学科性质、研究对象和国家统计的职能、统计研究的基本方法,重点掌握统计学中的几个基本概念。

#### **【学习要求】**

1. 重点掌握:统计总体、个体、统计样本、统计标志、统计指标、统计指标体系、参数、变量的概念;
2. 掌握:统计学中的数据分类;
3. 熟悉:医药管理统计学的应用领域及 Excel 的基本操作;
4. 了解:医药管理统计学的含义、研究对象。

## 第一节 医药管理统计学及其应用

在医药管理领域内,管理者和医药研究者常常会接触到“统计”这一术语,在电视、网页等有关媒体中也常常会看见一些新闻报道会使用各种统计数据、图表等,而在各类学术期刊杂志和会议报道中,各种统计学方法的运用和各类统计图表的使用更是不胜枚举。本节将介绍医药管理统计学的含义、研究对象以及运用领域等。

### 一、医药管理统计学的含义

医药管理统计学是统计学在医药管理领域的应用而衍生出的一门应用型学科,即运用统计学的原理和方法研究医药管理领域内的各种问题,以揭示医药行业发展的内在数量规律。其主要任务在于收集、分析、表达和解释医药管理相关领域内关于医药企业、医药产品、医药市场等医药管理数据,从而为该领域的研究拓展研究方法,提供科学的定量佐证。

与统计学的分类相似,医药管理统计学按其不同用途也可以分为描述统计学(descriptive statistics)和推断统计学(inferential statistics)。对给定的一组数据,统计学可以找出其关键信息并且描述这组数据,这种方法称为描述统计学;而研究者

以数据的形态建立一个用以解释其随机性和不确定性的数学模型，从而推断出研究的母体和制定研究步骤，这种方法则称为推断统计学。这两种方法在医药管理领域都具有重要的应用价值，前者适用于实证调研后的数理统计和分析，而当研究者提出新的理论和假说时，则往往需要依赖推断统计学来进行模型的构建和验证。

从以上的含义中，我们可以看出医药管理统计学的应用与医药管理领域的研究遵循着这样一个规律：第一，提出研究问题；第二，收集相关数据；第三，运用描述统计学对这些数据进行数理统计和分析；第四，提出理论假说和构建模型；最后，验证假设，并说明结果（图 1-1）。

## 二、医药管理统计学的研究对象

医药管理统计学的研究对象，狭义地说，是指研究者所搜集的各类有关医药的数据，例如某年度医药工业主要经济指标数据、一家上市医药公司年度报告中的所有财务数据、某家药品生产车间生产数据等。本书认为医药管理统计学的研究对象更为广泛，即广义地说，不仅包括所采集的医药管理领域内的所有数据，还包括研究者所构建或验证的理论模型，甚至包括研究者进行系统综述时所涉及到的所有文献。

## 三、医药管理统计学的应用领域

医药管理统计学的理论方法已广泛应用到医药管理研究与实践的各个方面，如医药企业的技术创新业绩的评估，企业管理与发展，调查问卷的评估与分析，医药管理领域内相关研究理论的实证检验以及对医药产业的发展预测等。此外，对于医药企业管理者而言，医药管理统计更是他们制定企业发展战略、构建产品营销策略以及建立员工业绩考核评估制度的必备工具。下面将对医药管理统计学在医药企业管理中的一些应用进行讲述。

### （一）新药研发风险评估

医药管理统计学在新药研发的风险评估中具有非常重要的应用价值。医药产品的研发具有高风险、高投入、高回报、长周期和低效率的特点，而新药研发是制药企业取得竞争优势的核心力量。因此，对新药研发的前期、研发过程以及研发后期的风险评估是制药企业降低投资风险、提高研发效率以及增加新产品收益的重要保障工作。在新药研发的前期，制药企业需要做新药研发项目决策的风险评估，即对该药物的研发技术难度、市场环境、新药审评机制等新药研发的内部条件和外部环

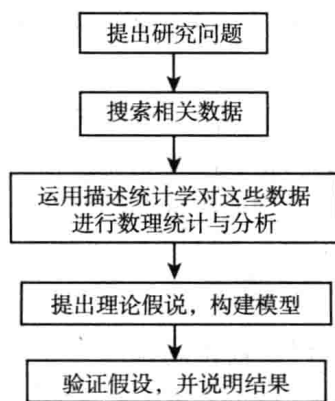


图 1-1 医药管理统计学运用规律

境进行分析,在充分考虑各方面因素之后,制定出与企业发展战略相一致且可行的新药研发项目决策,并启动该项目。在新药研发过程中,需要对研发的预算、资源配置、融资渠道等方面进行风险评估,以降低研发资金来源单一或不足的风险。最后,新药研发的后期,即新药上市期间,同样需要进行风险评估。一方面,医药市场竞争激烈,制药企业需要对新药上市的时间、新药上市后的仿制药市场等方面进行风险评估,并采取应对措施减轻可能的负面影响;另一方面,新药上市后,制药企业还需要进一步收集新药的安全性数据,以便制药企业对新药的安全性进行监控,从而为制定新药的营销战略提供依据。

## (二) 药品质量管理

药品作为人们用以防病治病、康复保健的特殊商品,其质量好坏直接关系到人们的生命健康。药品安全问题是重大的民生和公共安全问题,事关全社会的和谐稳定。因此,对于药品生产企业而言,保证药品的有效性、安全性和质量可控是其立足于医药市场的根本。一般来说,生产符合国家标准的药品仅仅是药品企业生存的底线,提高药品质量、增强企业的技术创新能力才是药品生产企业持续发展的关键,亦是企业树立品牌效应的关键所在。为了提高药品质量、加强药品质量管理,越来越多的药品生产企业采用 $6\sigma$ 准则实行药品质量的监控,从而保证药品质量的均一性和稳定性。

## (三) 医药市场预测

任何一家医药企业,无论是药品生产企业还是药品流通企业,要想在激烈的竞争中获取一席之地,首先必须了解医药市场。因此需要进行广泛的市场调查,搜集相关数据,并对这些信息进行科学的分析,以便作为药品研发、生产和经营的依据。无论是数据的搜集,还是对数据的分析与预测都需要医药管理统计学的支持。举例来说,仿制药厂商想要仿制即将过期的专利药,首先需要搜集这些即将过期的专利药的信息,了解市场需求以及企业自身与竞争对手的研发能力等基本情况,然后利用医药管理统计学中的相关方法对信息进行处理和分析,从而预测仿制药的预期收益,从而做出最终决策。

## (四) 医药企业人力资源管理

人力资源作为企业的一种无形资源,不仅是保证企业运转的源动力,也是推动企业成长的助动力。因此,人力资源战略越来越受到医药企业的重视,而在实际操作中关于如何制定人力资源战略,并非是一个简单的问题。这就需要利用统计学方法对职工的年龄、性别、受教育程度、工资、业绩等进行分析,从而为医药企业制定员工绩效考核标准、奖惩制度提供依据,以不断优化资源配置,加强人力资源的管理。

## (五) 医药企业发展战略制定

发展战略关乎一个企业的长远发展方向,是企业体现其发展愿景和文化理念的重要途径。制定医药企业的发展战略除了需要了解和把握整个宏观经济的发展状况、

医药行业的发展趋势，还需要分析企业自身的优势、劣势，以及面临的发展机遇和潜在威胁，从而明确企业在市场中的定位。所有这些资料的分析都离不开医药管理统计，需要在获得可靠数据的基础上，运用统计学方法对所有与企业运营相关的数据进行分析与预测。

当然，医药管理统计学并不是万能的，不能解决医药管理领域内的所有问题。医药管理统计学不可能为企业的管理者提供所有的管理技巧和方法，而只能为管理者做出决策提供依据，同时，管理者也并非完全凭借统计结果来处理所有日常事务，最重要的是从这些数据和统计结果中获取影响其决策的有用信息。对于社会科学的工作研究者而言，统计只是搜集、处理和分析数据的工具，而对统计结论的进一步解释则需要研究者的专业知识。总之，医药管理统计学是医药管理领域最为基础、应用最为广泛的工具，为人们探寻医药经济与管理现象的本质和经济发展的规律提供一定的帮助。

## 第二节 医药统计数据类型

在医药管理领域中，研究者常常需要对观察单位（个体）的某项或某些特征（指标）进行观测，这些特征（指标）在统计学中被称为变量（variable），其值则被称为变量值（variable value）或观察值（observation）。统计学中所指的变量通常是指随机变量（random variable），它是指取值不能事先确定的观察结果。在医药管理统计学中，随机变量一般用大写的拉丁字母表示。例如，2010年某省的所有药品生产企业的经营业绩情况，可记为 $Y$ ，其中某一家药品生产企业2010年的营业总收入记为 $Y_i$ 。通常一家企业的经营业绩受到该企业的规模、管理水平、创新能力以及外部政策等多方面因素的影响，从而使得企业之间存在个体差异。在医药管理实践中，有些变量能够直接获取具体数值，而有些变量却只能进行定性判断。对于这些不可直接获取数值的变量，在统计学中一般以虚拟变量替代。

总的来说，统计数据是对现象进行测量的结果，由于使用的测量尺度或收集方式等的不同，统计数据可分为不同的类型。常见的分类方法有以下三种：

### 一、按计量尺度分类

根据所采用的计量尺度，可以将统计数据分为分类数据、顺序数据和数值型数据。

#### （一）分类数据

分类数据（categorical data）是指那些只能归于某一类别的非数值型数据，一般以文字表述。分类数据依据不同的分类尺度会形成不同的组别，从而产生不同的结果。例如，对我国市场上所售的所有药品进行分类，以是否为处方药可以分为处方

药和非处方药两种；而按照药品的结构属性可分为化学药品、中成药和生物制品；按照药品的来源不同又可分为国产药品和进口药品等。虽然分类结果不同，但我们得到的这些数据均为分类数据。为便于统计处理，一般用数字代码来表示分类数据，比如，用1表示“处方药”，0表示“非处方药”；用1代表“化学药品”，2表示“中成药”，3表示“生物制品”等。

## （二）顺序数据

顺序数据（rank data）是指只能归于某一有序类别的非数值型数据，即这些类别是有顺序的。依照顺序尺度计量形成的数据在一定程度上具有逻辑性，有利于数据的收录和整理，在对事物分级以及问卷调查中应用非常广泛。比如，我国将化学新药分为5类，即Ⅰ类化学新药、Ⅱ类化学新药、Ⅲ类化学新药、Ⅳ类化学新药和Ⅴ类化学新药；一家医药企业内所有员工的受教育程度可以分为大专以下、大专、本科、硕士和博士；李克特量表<sup>①</sup>中常会出现这样的选项，“非常不同意”、“不同意”、“不一定”、“同意”、“非常同意”，这是依照人们的态度层次来计量的。类似地，对顺序数据也可以用数字代码来表示，比如，1表示“非常不同意”，2表示“不同意”，3表示“不一定”，4表示“同意”以及5表示“非常同意”。值得注意的一点是分类数据中的所有的数字代码仅是代号，不能说明这些非数值型数据之间存在偏序关系；而在顺序数据中，数字代码所代表的类别之间往往存在一定的逻辑关系。

## （三）数值型数据

数值型数据（metric data）是指那些能够依照数值尺度测量的观测值，并以具体的数值呈现。数值型数据是按照特定的度量衡单位对事物进行测量的结果，因此，这些数值往往具有单位量纲。比如，某家药店的年销售额，其单位一般为元或万元；此外，药品的一些理化性质如片剂的溶出度、重量、有效成分含量等都是数值型数据。实际上，数值型数据是我们生活中使用最为广泛的数据类型。

综合比较上述3种数据类型，可以看出分类数据和顺序数据适用于描述事物的品质特征，有助于统计者进行归类和整理，因而又可称为定性数据或品质数据；数值型数据描述的是现象的数量特征，有助于统计者进行计算和数学处理，因而也可称为定量数据或数量数据。

## 二、按收集方法分类

依照统计数据的收集方法，数据又可以分为观测数据和试验数据。

### （一）观测数据

观测数据（observational data）是指通过调查或观测而收集到的数据。观测数据

<sup>①</sup> 该量表由一组陈述组成，每一陈述有“非常同意”、“同意”、“不一定”、“不同意”、“非常不同意”五种回答，分别记为1, 2, 3, 4, 5，每个被调查者的态度总分就是他对各题目的回答所得分数的加总，这一总分可说明他的态度强弱或她在这一量表上的不同状态。



往往是在没有人为控制或干预的情况下获取的数据，如《中国卫生统计年鉴》、《中国统计年鉴》等统计年鉴中收录的宏观统计数据都是这种类型的数据，是客观存在的数据。

## （二）试验数据

试验数据（experimental data）是指在试验中控制试验对象而收集到的数据。这类数据的获取往往与科研工作者所设计的试验方案有关。因此，试验数据在一定程度上会受到个体的主观影响。如在新药研发的过程中，依据药品不同的治疗目的，研究者所设计的研究方案就会有所差异，最终将会获得差异较大的试验数据和截然不同的试验结果，美国辉瑞公司对药品“万艾可”的发现就是最为典型的例子。

## 三、按时间特征分类

按照被描述对象与时间的关系，可以将统计数据分为截面数据、时间序列数据和面板数据。

### （一）截面数据

截面数据（cross-sectional data）是指在相同或近似相同的时间点上收集的数据，因此获得数据往往来自同一时间点的不同空间上的数据。如2010年我国各省市医药制造业的总产值就是截面数据。在对我国各地区医药产业的综合竞争力所进行的比较中，研究者通常采用的数据就是截面数据。

### （二）时间序列数据

时间序列数据（time series data）是指在不同时间点或者时间段上收集的数据，主要用于描述某一事物随时间而变化的情况。如2005~2010年5年内我国每年新药注册申请的数量就是一组时间序列数据。

### （三）面板数据

面板数据（panel data）是截面数据与时间序列数据综合起来的一种数据类型。这类数据是依照时间序列和截面两个维度排列的，因而最终排在一个平面上，使得整个数据结构像是一个面板。如2005~2010年5年间我国各省市医药产业产值就是一组面板数据。面板数据能够更好地发现和测量纯粹横截面数据或纯粹时间序列数据不能观测到的效应。此外，面板数据还增加了样本容量，有利于减少计量偏差。

## 第三节 医药管理统计学中的基本概念

医药管理统计学的概念众多，其中有些概念在日常管理工作中是经常用到的，比如：样本、总体、个体、指标等。还有一些概念如参数、变量等则经常出现在各类研究文章与报告中。这些概念在本书中将频繁出现，本节将着重介绍医药管理统计中的一些基本概念。

## 一、统计总体和个体

统计总体 (population) 是指包含一项研究中全部数据的集合。如由多个药品生产企业构成的集合、多种药品构成的集合、多个实验构成的集合等等。其中, 构成总体的每一个元素称为个体 (individual), 在由多个药品生产企业构成的总体中, 每一家企业就是一个个体; 在多种药品构成的集合中, 每一种规格的药品就是一个个体; 在多个实验构成的集合中, 每一次实验就是一个个体。

一般来说, 确定一个统计总体的范围较为容易。例如, 要检验一批药品的有效成分含量, 这批药品构成的集合就是统计总体, 而这批药品中的每一片药就是一个个体。但也有一些统计总体的范围却难以确定。例如经济学家在描述和分析影响医药企业长期发展的因素时, 就必须找出所有的影响因素, 然而社会环境、组织结构的复杂性使得经济学家很难确定这个统计总体的范围。因此, 经济学家们根据所研究的视角和目的不同, 定义的统计范围也就有所不同。

统计总体根据其所包含的单位数目是否可数, 可分为有限总体和无限总体。

### (一) 有限总体

顾名思义, 有限总体 (finite population) 是指总体的范围能够明确确定, 且个体的数目是有限可数的。举例来说, 我国所有的医药企业所构成的总体就是一个有限总体; 一批待检药品所构成的总体也是一个有限总体。

### (二) 无限总体

与有限总体相对地, 无限总体 (infinite population) 就是指总体中所包括的个体是无限的、不可数的。举例来说, 在新药的Ⅳ期临床试验中, 一次试验就是总体中的一个个体, 而药品生产企业在新药上市后, 为了监测新药的不良反应和控制新药的安全性, 会在较长的时间内持续进行Ⅳ期临床试验, 因此由这些Ⅳ期临床试验构成的统计总体就是一个无限总体。

人们对有限总体和无限总体进行抽样时, 产生的结果是不一样的。对于有限总体, 每一次从中抽取一个个体, 都会减少有限总体的个体数目, 因此, 前一次的抽样结果往往会影响到下一次的抽样结果。也就是说, 有限总体的每次抽样是不独立的。而对于无限总体来说, 每一次抽样则可以看作是独立的, 因为前一次的抽样不会影响到下一次的抽样结果。

## 二、统计样本

统计样本 (statistical sample) 是指从总体中抽取的一部分个体的集合。其中, 构成样本的个体数目, 称为样本容量 (sample size)。样本存在的目的在于帮助研究者根据样本提供的有关信息来推断研究总体的特征。比如, 从一批感冒胶囊中随机抽取 100 粒, 这 100 粒胶囊就构成了一个样本, 然后再根据这 100 粒胶囊的有效成分含量来推断这一批胶囊的平均有效成分含量。



### 三、统计标志

统计标志 (statistical characteristics) 是指统计总体中所有元素都具有的共同特征的名称。根据研究者的需要, 可以从不同的角度考察这些个体的特征。例如, 每家医院的日住院人次、床位使用率、年收入等。这些特征的具体表现, 如一家综合医院平均每日住院人次为 9 人, 床位使用率为 89.2%, 年收入为 1.5 亿元, 称为标志表现。

#### (一) 不变标志和变异标志

当一项标志在统计总体内所有个体的具体表现都相同时, 这个标志就是不变标志; 反之, 如果一项标志在所有个体的具体表现都不同, 则这个标志就称为变异标志。我国《药品管理法》对新药的定义为: “新药是未曾在中国境内上市销售的药品。”从这个定义上看, 未在国内上市就是一个不变标志, 而药品的种类、规格、来源等则是变异标志。不变标志是构成统计总体的基础, 因为必须至少存在一个不变标志才能将离散的个体统一联系起来, 从而依赖其“同质性”而构成一个统计总体; 变异标志则是统计研究的主要内容, 显然如果所有个体的所有标志表现都一致的话, 就失去了进行统计分析的必要性。

#### (二) 品质标志和数量标志

品质标志用于指示事物的特性, 是不能用数值表示的, 例如药品的种类、来源等; 而数量标志则显示事物量的特性, 可以用具体数值来表示, 如药品的剂量、规格、价格等。一般来说, 品质标志能够将不同性质的个体分开, 适用于分组, 便于计算各组的个体数目, 计算结构和比例指标; 数量标志则既可以用于分组, 也可以用于计算各种数量指标。

### 四、统计指标

#### (一) 统计指标及其构成要素

在医药管理统计学中, 对统计指标有两种解释。

1. 统计理论和统计设计所使用的解释认为, 统计指标是指反映总体现象数量特征的概念, 包括指标名称、计量单位和计算方法三个要素。

2. 实际的统计工作中, 更常用的另一种解释为统计指标是反映总体现象数量特征的概念和具体数值。因此, 统计指标应当具有指标名称、计量单位、计算方法、时间限制、空间限制和指标数值六个要素。

这两种解释在统计研究的全过程都有所体现。举例来说, 在未调查之前, 研究者只能设计统计指标的名称、内容、计量单位和计算方法, 而没有具体的数值; 而在经过数据的搜集、整理、计算和分析之后便能得到统计指标的具体数值, 从而反映总体现象的实际数量状况及其发展变化的情况。