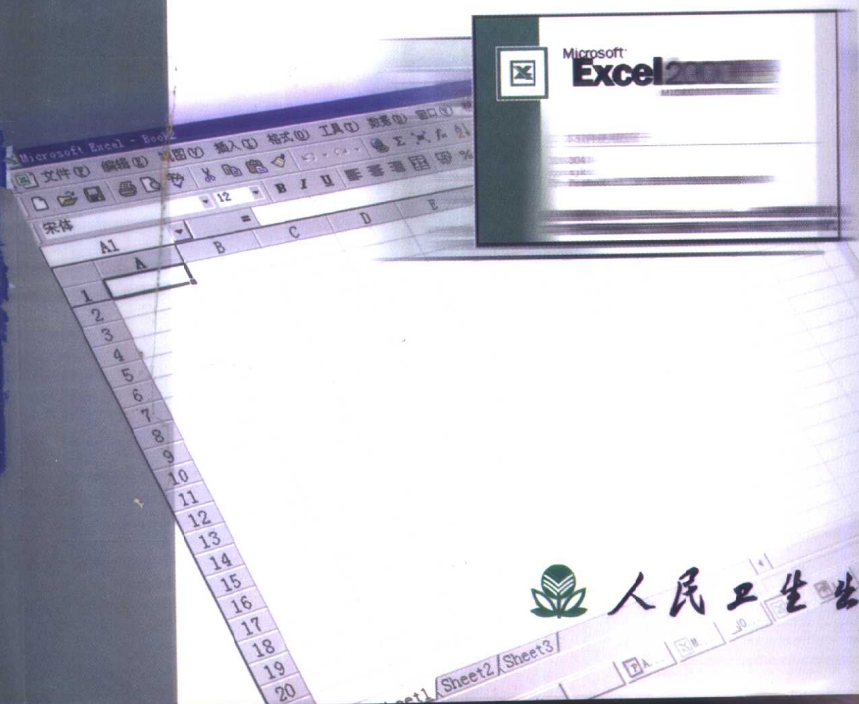




Excel

在统计分析中的 应用

主编 刘钢



 人民卫生出版社

Excel 在统计分析中的应用

主 编 刘 钢

编 者 徐一平 陶育纯 盖学良

杨湘山 李治国

人 民 卫 生 出 版 社

图书在版编目(CIP)数据

Excel 在统计分析中的应用/刘钢主编.

北京:人民卫生出版社,2002

ISBN 7-117-05163-9

I .E... II .刘... III .电子表格系统,

Excel-应用-医学统计-统计分析 IV .R195.1-39

中国版本图书馆 CIP 数据核字(2002)第 077276 号

Excel 在统计分析中的应用

主 编:刘 钢

出版发行:人民卫生出版社(中继线 67616688)

地 址:(100078)北京市丰台区方庄芳群园 3 区 3 号楼

网 址:<http://www.pmph.com>

E-mail: pmph@pmph.com

印 刷:北京市增富印刷有限责任公司(天运)

经 销:新华书店

开 本:850×1168 1/32 印张:11.5

字 数:270 千字

版 次:2002 年 11 月第 1 版 2002 年 11 月第 1 版第 1 次印刷

标准书号:ISBN 7-117-05163-9/R·5164

定 价:23.00 元

著作权所有,请勿擅自用本书制作各类出版物,违者必究

(凡属质量问题请与本社发行部联系退换)

前 言

医学、药学工作者在工作和研究中经常需要对数据作统计分析。有人报导，在医学期刊论文中应用的统计方法大多数属于 t 检验、方差分析、 χ^2 检验等基本的或经典的统计方法。在计算机相当普及的今天，利用计算机完成统计分析是人们很自然的选择。尽管目前有多种优秀的统计专业软件（例如 SAS、SPSS 等）可供选用，但是这些专业软件往往系统庞大（其中大多数内容只适合于专业统计人员）、结构复杂。即使经过相当的努力，大多数非统计专业人员也难以运用自如。而且正版统计学软件价格不菲，一般的应用者难以获得。

电子表格软件 Excel 提供了大多数基本的和经典的统计分析方法。通过简单的模板编制，还可以实现一批 Excel 没有直接提供的统计方法（如秩和检验、寿命表以及 logrank 检验等）。

用 Excel 作统计分析具有以下优点：

1. Excel 提供的统计描述和统计推断方法形式直观。
2. 学习与使用极为方便。
3. Excel 软件容易获得。几乎所有安装 Windows 操作系统的微型计算机都选择安装 Office。安装了 Office 就拥有了 Excel。用户不必要另外投资。
4. Excel 不但提供了大多数基本的统计分析工具，而且提供了丰富的统计图和统计表的编制功能。
5. Excel 与 Windows 平台上的文字处理工具 Word、多媒体制作软件交流十分方便。

6. Excel 可以直接从常用数据库读取数据。

7. 只要计算机资源允许, Excel 几乎可以处理任意样本量的资料。

本书编者均为多年从事统计学、计算机应用基础课教学和科研的大学教师。经多年的实践和比较我们认为, 与 SAS、SPSS、STATISTICA、MINITAB 等统计学专用软件相比, Excel 对广大非统计专业人员的确是一个易学、易用、廉价、高效的好助手。

“少投入(不仅指物化的, 还包括时间、精力等), 多产出”是全人类科学技术不断进步的基本动力之一, 也是人们在处理各种事物时普遍追求的原则之一。本书可以为那些需要进行统计分析的读者提供一种体现这个原则的选择或途径。为此, 我们积多年教学和使用各种软件进行统计分析的感受, 集思广益编写此书, 奉献给需要基本统计分析方法的广大医学、药学和其他领域的读者。本书完全可以作为医学或药学专业本科生与硕士研究生统计软件课程教材的一部分。

本书假定读者对统计学基本概念和基本方法已经了解。因此不拟对统计学的基本内容做说明。虽然本书采用的数据例子多数属于医学范畴, 但是我们相信这里使用 Excel 的方法也会使从事其他领域工作的读者受益。

本书第一章是关于电子表格软件 Excel 基本概念和功能的一个自给自足的介绍。考虑到本书的主题, 本章对那些与统计分析有关的内容给予了更多的注意。对 Excel 了解较多的读者可以跳过这一章而直接阅读后面的章节, 或者必要时再作参考。

第二章对所有读者都是必要的。这部分篇幅介绍了利用 Excel 进行统计分析必须借助的工具以及本书编者提出的“统计分析模板”的概念。熟练掌握 Excel 提供的统计分析工具(统计函数与分析工具)和模板的编制技巧(按照西方人的说法,

这也许是一种挑战)意味着读者驾驭 Excel 的能力的提高,同时也会使读者感受到成功的乐趣。

第三章到第十四章分别介绍了如何在 Excel 中编制统计表与统计图,如何实现描述统计、正态分布应用、区间估计、t 检验、方差分析、二项分布与 Poisson 分布的应用、Pearson χ^2 检验、秩和检验、直线相关与回归分析、正态性检验、寿命表分析和 logrank 检验等统计学方法。各章的内容互不依赖,基本上是平行的。一个例外是第三章(统计表)与第四章(统计图)介绍的方法在后面的个别章节中有所引用。

第十五章将 Excel 作出的统计分析结果与专业软件 SAS 或 SPSS 的进行了比较。这对本书的编写可能具有某种支撑意义。根据这一章的内容我们认为,用 Excel 作出的统计分析结果的科学性毋庸置疑,可以在各种报告中引用。

本书主要根据 Office97 编写。但是 Excel2000 中与统计分析有关的内容没有明显变化。

尽管各位编者竭尽了自己的所能,但是限于我们的学识与专业水平,书稿中还是可能存在错误与不尽人意之处。我们诚恳地希望读者不吝赐教,以期有机会再版时修正。

最后,对热情关心、支持本书编写的吉林大学公共卫生学院的有关领导者与同事,对为了本书出版付出心血的编辑们,致以衷心的感谢。

编者 刘 钢

2002 年 4 月

目 录

第一章 Excel 操作简介	1
第一节 Excel 基础知识	1
第二节 工作表的建立	4
第三节 工作表的编辑和格式设置	18
第四节 数据的管理与分析	28
第五节 工作页面设置与打印	35
第二章 Excel 的函数、分析工具库与统计分析模板	40
第一节 函数	40
第二节 数据分析工具库	44
第三节 统计分析模板	48
第三章 统计表	52
第一节 统计表的结构和制表的基本要求	52
第二节 统计表的制作	54
第四章 统计图	77
第一节 制图的基本要求	77
第二节 常用的统计图	78
第三节 Excel 提供的其他统计图	104
第四节 Excel 图表应用——质量控制图	109
第五章 描述统计	120
第一节 频数表的编制	120
第二节 利用函数进行描述统计	126
第三节 利用分析工具进行描述统计	130

第六章 正态分布及其应用	133
第一节 正态分布曲线的绘制	133
第二节 正态分布分界点的计算及应用	139
第三节 正态分布密度曲线下面积的计算与应用	143
第七章 t 检验	146
第一节 t 界值的获得	147
第二节 总体均数的估计	148
第三节 t 检验	154
第八章 方差分析	162
第一节 单因素方差分析	162
第二节 无重复观察的两因素方差分析	164
第三节 有重复观察的两因素方差分析	167
第四节 多个样本均数间的两两比较	170
第五节 多个方差的齐性检验	174
第九章 二项分布与 Poisson 分布	178
第一节 二项分布的概率和累积概率的求法	178
第二节 Poisson 分布的概率和累积概率的求法	184
第十章 χ^2 检验	190
第一节 两样本率比较的四格表 χ^2 检验	190
第二节 行 $\times$ 列表的 χ^2 检验	196
第三节 配对四格表的 χ^2 检验	204
第四节 Fisher 确切概率法	206
第十一章 非参数统计分析方法	213
第一节 符号秩和检验(Wilcoxon 法)	213
第二节 完全随机设计两样本比较的秩和检验	220
第三节 完全随机设计多样本比较的秩和检验 (Kruskal-Wallis 法)	230
第四节 多个样本间两两比较的秩和检验	240

第十二章 直线回归与相关分析	243
第一节 直线回归分析	243
第二节 直线相关分析	247
第三节 秩次相关系数	251
第四节 曲线拟合	254
第五节 多元线性回归分析	256
第十三章 与其他软件交换数据	260
第十四章 其他统计方法	269
第一节 正态性检验	269
第二节 寿命表	272
第三节 对数秩检验	275
第四节 随机数的生成与应用	277
第五节 根据概括数据的 t 检验	286
第六节 大样本问题	290
第十五章 与其它统计软件的比较	291
第一节 概率分布计算的比较	291
第二节 Excel 的统计分析结果与 SPSS、SAS 的比较	296
附录 1 Excel 提供的统计函数	305
附录 2 Excel 与 SAS、SPSS 分析结果比较	341
附录 3 T 界值表与 r_s 界值表	354

第一章 Excel 操作简介

第一节 Excel 基础知识

一、Excel 概述

Excel 是 Microsoft Office 软件的主要组成部分，是目前应用最普遍的电子表格软件。它具有较强的数据处理与分析能力，丰富的图表功能，在统计分析中有着广泛的应用。

1. 主要功能

(1) 电子表格制作：提供了丰富的电子表格模板，可做出各类美观实用的统计表格。

(2) 图表制作：具有丰富的图表功能，依据表格中数据的结构生成不同类型直观、实用的统计图表。

(3) 数据的组织、运算与分析：提供了丰富的函数和数据分析工具，用以进行复杂的表内及表间数据运算和统计分析。所包含的大量统计函数及具备专业水准的统计分析工具，可对电子表格中数据进行组织、管理和分析。

(4) 网络功能：具有强大的 Web 功能。用户可直接从浏览器中拖动 Excel 的表格数据，在网页上对图形和对象进行操作。

2. 主要特点

(1) 三维电子表格：采用了工作簿——二维工作表的三维结构，便于建立表格的多层关系。

(2) 智能感知技术：可根据当前操作，预测下一步可能进行的操作，如列表、自动填充、自动求和等，自动探测可能需要进行操作的数据区域，提高工作效率。

(3) 各种向导：利用图表向导、数据透视表向导及模板向导等，规范了相关操作。

二、启动与退出

1. 启动

Excel 的启动方法与 Office 中其它应用程序相同，常用的方法有：

(1) 通过“开始”菜单启动 Excel：单击任务栏左端[开始]按钮→“程序”菜单→“Microsoft Excel”项；

(2) 在 Windows 桌面上双击快捷方式图标启动 Excel；

(3) 利用 Excel 文档启动 Excel。

2. 退出

必须首先保存当前编辑的工作簿文件，利用以下方法：

(1) 单击“文件”菜单→“退出”命令；

(2) 单击 Excel 窗口右上角[关闭]按钮；

(3) 双击 Excel 左上角控制菜单图标。

三、窗口组成

启动 Excel 后就会打开 Excel 的用户界面窗口。该窗口由以下几部分组成(图 1-1)：

1. 标题栏 位于窗口最上端，显示“Microsoft Excel”应用程序标题及当前工作簿名称。

2. 菜单栏 将 Excel 丰富的命令集合，按类别分为文件、编辑等九项菜单。

3. 工具栏 将 Excel 常用功能设成按钮形式，组成“常用”

工具栏和“格式”工具栏。

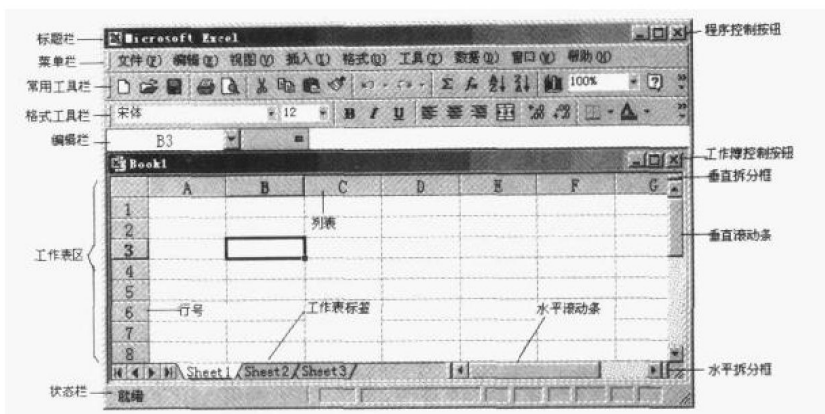


图 1-1 Excel 窗口组成

4. 编辑栏 左边是名称框，显示当前活动单元格或区域相对应的地址及名称。右边为编辑区，用于输入或编辑活动单元格的数据及公式。输入数据时编辑栏中出现三个按钮，按钮功能如下：

× 按钮：撤消按钮，用于撤消该单元格的本次输入或修改，退出编辑状态。

√ 按钮：确认按钮，用于确认和锁定该单元格此次输入内容。

= 按钮：编辑公式按钮，单击该按钮时，左边名称框变为函数列表框，供用户选择所需函数，并自动录入到编辑区。

输入到单元格的数据会同时显示在单元格和编辑区。对于长数据，单元格默认宽度(8个字符)通常不能满足数据内容全显示，而编辑区的显示将是完全的。

5. 工作表区 数据窗口主体，由单元格组成，用以记录数据。

6. 工作表标签 位于工作簿文档窗口左下底部，显示工作表名称。单击工作表标签将激活相应工作表为当前工作表，以实现在同一工作簿中不同工作表之间切换。一个工作簿中可包含多个工作表，默认为 sheet 1、sheet 2、sheet 3 三个工作表。

7. 滚动条及拆分框 滚动条用于水平、垂直方向改变工作表的可见区域。

拆分框分别位于垂直滚动条顶端和水平滚动条右端。拖动拆分框对工作表进行横向或纵向分割，有助于同时查看同一工作表的不同部分。双击拆分框可取消工作表的拆分。

8. 状态栏 位于 Excel 窗口底部，左边显示当前选定命令的相应信息，右边是几个特殊键状态区，如<Caps lock>、<Scroll lock>及<Num lock>等键。

第二节 工作表的建立

一、工作簿、工作表和单元格

1. 工作簿

一个工作簿就是一个 Excel 文档，是用来储存并处理数据的文件。工作簿名称显示在标题栏上(如 Book 1)，其扩展名为“.xls”。一个工作簿最多可由 255 个不同的工作表组成。用户可根据实际情况增减工作表或选择工作表。

2. 工作表

工作表是 Excel 实施数据管理的基本单位。每张工作表有一个标签与之对应(如“sheet1”)，工作表名称显示在窗口底部左侧标签处。

3. 单元格

单元格是组成工作表的最小单位。工作表中每一行、列交

叉处即为一个单元格。每个单元格由列标和行号标识,用来指明单元格在工作表中所处位置。工作区的第一行为列标,用 A—Z、AA—AZ……表示,共 256 列;左边第一列为行号,用 1~65536 数字表示,即 Excel 的一个工作表有 256×65536 个单元格。如单元格“A2”,表示位于表中第 A 列、第 2 行。单元格内可输入并保存由字符串、数字、公式等组成或产生的数据。

4. 单元格区域

由相邻单元格组成的矩形区域,如图 1-2 中由 A1 至 E5 相邻单元格组成单元格区域。一般用主对角线两端的单元格将该区域记为“A1:E5”。

单击某单元格即被加黑框标注,表示为“当前单元格”(或“活动单元格”),当前输入的数据或公式保存在该单元格内。当前工作表中只能有一个单元格是活动的。

工作簿与工作表的关系类似账簿与账页之间的关系。打开了工作簿就打开了它的所有工作表,被选定标签的工作表是可进行编辑的当前工作表。

例 1-1,有 A、B 两种药物用于治疗某疾病,效果对比数据见表 1-1

表 1-1 A、B 两种药物治疗某病效果比较

分 组	治 愈	未 愈	合 计	治愈率
A	25	4	29	86.21
B	18	14	32	56.25
合计	43	18	61	70.49

将表 1-1 中数据输入 Excel 工作表的 A1:E5 区域,如图 1-2 所示。

	A	B	C	D	E	F
1	A、B两种药物治疗某病效果比较					
2	分组	治愈	未愈	合计	治愈率	
3	A	25	4	29	=B3+C3	
4	B	18	14	32		
5	合计	43	18	61	=SUM(B3:C4)	
6						
7						

图 1-2 单元格区域

二、单元格或区域选中

只有选中单元格或区域后方可对其进行输入、编辑等操作。选中是电子表格常用操作，分为选中单个单元格、选中多个连续单元格及选中多个不连续单元格。各种选中方法见表 1-2 所示。

表 1-2 选中单元格及区域方法

区 域	方 法
单元格	单击该单元格
整行(列)	单击工作表相应的行(列)号
整张工作表	单击工作表左上角行列交叉按钮
相邻行或列	指针拖过相邻的行号或列标
不相邻行或列	选定第一行(列)后，按住<Ctrl>键，再选择其他行(列)
相邻单元格区域	单击区域左上角单元格，拖至区域右下角(或按住<shift>键再单击右下角单元格)
不相邻单元格区域	选定第一个区域后，按住<Ctrl>键，再选择其他区域

单击工作表内任意位置，即可取消选定，各单元格又成为独立的单位。

三、数据输入

数据不仅可由键盘直接输入，也可充分利用 Excel 提供的自动输入及函数、公式的自动计算方法，每个单元格输入结束后按回车键、<Tab>键或单击编辑栏的〈√〉按钮均可确认输入；按<Esc>键或单击编辑栏〈×〉按钮可取消输入。

1. 数据的输入

输入的数据类型分为字符型、数值型和日期时间型。

(1) 字符的输入：字符的输入在单元格中默认左对齐。一个单元格内，最多可输入 32000 个字符；每个单元格缺省宽度为 8 个字符宽，若字符数超出单元格宽度时，输入内容扩展到右侧列上，如果右侧列上有数据，输入内容的显示被截断，但编辑栏中显示该单元格的全部内容。

输入数字字符时，在数字字符串前加单引号，如'123，单引号仅在编辑栏中出现。

(2) 数字的输入：数字的输入在单元格中默认右对齐。当输入数据长度超出单元格宽时，Excel 自动以科学计数法表示。编辑栏中至多显示前 15 位数字原样。如单元格数字格式设置为带两位小数，当输入三位小数时，末位将进行四舍五入。Excel 计算时以输入数值而不是显示数值为准。

(3) 日期和时间输入：Excel 内置了一些日期、时间格式，如：yy-mm-dd 或 yy/mm/dd。两个日期型数据相减结果为相隔天数。时间型数据格式一般为：hh:mm(am/pm)，"am/pm"与分钟之间应留有空格。

(4) 区域的数据输入：采用区域输入方法可实现在选取区域内自动回行选择单元格。先选取一个区域，在区域中某一单

单元格输入结束后，使用<Tab>键即可按行或列的顺序跳转到下一个单元格输入数据。当区域内一行(列)满时，会自动进入该区域的下一行(或列)的第一个单元格。

2. 数据自动输入

利用自动输入功能，可以方便快捷地输入等差、等比及自定义的数据序列。

(1) 自动输入数据序列：由初始项决定后面的填充项，光标指向初始值所在单元格右下角的填充柄，拖动至要填充行(列)的最后一个单元格即完成自动填充操作。

初始值为纯字符或纯数字，填充时相当于数据复制；初始值为字符数字混合体，填充时字符不变，最右边的数字递增。如：初始值为 A1，填充为 A2, A3...；初始值为 Excel 已定义的自动填充序列中成员，则按该序列填充，如：甲、乙、丙、丁…。

(2) 自动完成输入数据：可利用“记忆式输入”功能，在工作表中输入一列具有许多相同或部分相同的数据。当在单元格中输入第一个字符或前几个字符(或汉字)时，Excel 会与该列已输入的数据项进行比较，找出相符的数据项，自动显示剩余字符(见图 1-3)，如欲输入数据项与给出数据项相同，按回车键即完成输入，否则继续输入后面的字符，此法只限字符形数据。

“选择列表”自动输入重复字符项的方法，右击欲输入数据单元格，在快捷菜单中单击“选择列表”命令，Excel 会显示该单元格所在列已键入的数据项所组成的列表(如图 1-4)，从中选择需要的数据项即可。

(3) 输入有效数据：Excel 具有预先设置对某一单元格允许输入的数据类型、范围，并可设置数据输入提示信息和输入错误提示信息。