

Excel

在统计学中的应用

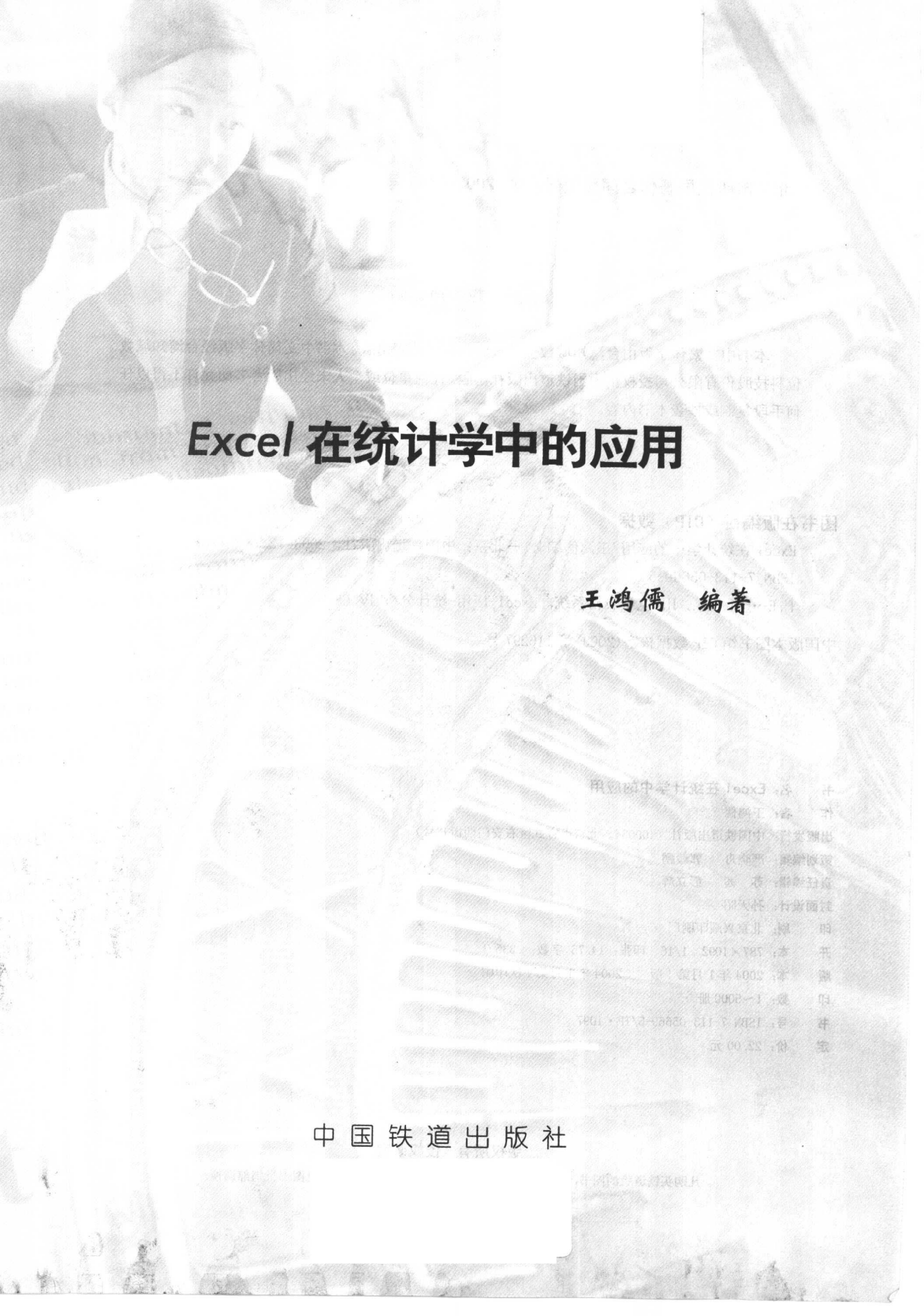
王鸿儒 编著



本书特色：

- 采用丰富的实例进行说明，清晰明了。
- 清楚的索引分类易于参考。
- 简化计算过程，易学易懂，降低统计学习门槛。
- 落实统计结果在商业中的概念，让学习Excel与统计相辅相成。

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE



Excel 在统计学中的应用

王鸿儒 编著

中国铁道出版社

北京市版权局著作合同登记号：01-2003-8748 号

版 权 声 明

本书中文繁体字版由台湾知城数位科技股份有限公司出版。本书中文简体字版经台湾知城数位科技股份有限公司授权由中国铁道出版社出版。任何单位或个人未经出版者书面允许不得以任何手段复制或抄袭本书内容。

图书在版编目 (CIP) 数据

Excel 在统计学中的应用/王鸿儒编著. —北京: 中国铁道出版社, 2003. 12

ISBN 7-113-05669-5

I. E… II. 王… III. 电子表格系统, Excel-应用-统计分析 IV. C819

中国版本图书馆 CIP 数据核字 (2003) 第 116297 号

书 名: Excel 在统计学中的应用

作 者: 王鸿儒

出版发行: 中国铁道出版社 (100054, 北京市宣武区右安门西街 8 号)

策划编辑: 严晓舟 郭毅鹏

责任编辑: 苏 茜 彭立辉

封面设计: 孙天昭

印 刷: 北京兴顺印刷厂

开 本: 787×1092 1/16 印张: 14.75 字数: 336 千

版 本: 2004 年 1 月第 1 版 2004 年 1 月第 1 次印刷

印 数: 1~5000 册

书 号: ISBN 7-113-05669-5/TP·1097

定 价: 22.00 元

版权所有 侵权必究

凡购买铁道版的图书, 如有缺页、倒页、脱页者, 请与本社计算机图书批销部调换。

出版说明

微软的 Office 软件深入国内信息市场已近 10 年之久，随着计算机及软件的迅速发展，计算机已逐渐地走入校园、办公室及政府机构当中。

而现在，学习单一软件，是成为信息人才的基础目标。不论是在校园中学习或者是担任职场上的尖兵，只会应用软件的基本操作，将不足以应付需求。将信息融入教学及扩展商机当中，才是日后必须加紧努力的目标。

Excel 为微软公司特别针对运算工作所开发的一套软件。发展至今，软件本身除了蕴含着强大的计算功能外，高级的 VBA 程序自定义功能，更是让此软件较其他类似软件更具成长性的关键特征。而《Excel 在统计学中的应用》一书，则是融合上述二者，再辅以商业活动中经常会遇到的问题，将之通过统计方法，通过 Excel 的强大计算能力，整理成切合商业决策适用的推论准则。如利用 Excel 中简单的加载宏功能，计算出统计中最重要的推论及假设工作，为制造商及产品管理人员推论出品质监控过程中是否出现瑕疵；或者，利用回归分析工具，推论出公司未来的预测收益，让您除了应用 Excel 本身的公式、函数及数据分析工具外，更能以熟悉的操作程序，学到更多益于工作的方法。

另外，本书还配有书中的范例文件，读者如果需要，可到 <http://www.tqbooks.net> 下载。

本书在统计理论实务上，主要参考的著作如下：

❀ 应用统计学. 林惠玲、陈正仓，双叶出版社

❀ 统计学. 方世荣，五南出版社

❀ 商用统计学. 叶大述译，颜廷芳校订，桂冠出版社

❀ 应用统计学. 赖铁军、郭祝武著

❀ Basic Business Statistics concepts and Applications ,Mark L. Berenson, David M. Levine, PRENTICE HALL

读者在学习本书中如何应用 Excel 计算统计的同时，在加强统计概念上，也可参考以上书籍。通过对本书中实用范例的学习，将使您达到更深入的学习效果。

本书由台湾知城数码科技股份有限公司特提供版权，张涛、崔海宇、朱霞、李志文、艾学华、刘东进、李宏、张忠宝、郑华、左和群等完成了本书的整稿与编排工作。

目 录

第一篇 基础操作篇

第 1 章	Excel 数据表及基础统计数据整理原则.....	3
1-1	统计学概念介绍	4
1-2	统计学中的总体与样本概念解析.....	4
1-3	Excel 与统计学的关系	5
1-4	应用 Excel 所需注意的事项	6
	正确定义输入的统计数据.....	6
	正确定义单元格数据范围.....	7
	绝对引用与相对引用的应用.....	8
第 2 章	使用 Excel 整理及显示统计数据	11
2-1	取得统计数据	12
2-2	整理统计数据与编组	12
	求出全距	13
	决定组数	13
	决定组距	13
	组距上、下限的选择.....	13
	计算组中点	13
	计算各组的次数.....	13
	操作实例：应用 Excel 工作表整理数据	14
	累加次数分布.....	17
	操作实例：做以下累加.....	17
2-3	统计数据的显示	18
	折线图	19
	饼图	20
	柱形图	20
	操作实例：使用图表向导，制作次数分布柱形图.....	21

第二篇 数值与概率篇

第 3 章	计算数值分散趋势	27
3-1	衡量未分组数据中央集中趋势.....	28
	平均数	28

操作实例：计算数列的平均数.....	28
中位数	30
操作实例：求取中位数.....	31
众数	32
操作实例：求取众数值.....	32
偏态观察	32
操作实例：使用数据分析工具制作分布图.....	33
数值选择	36
3-2 计算数值离散程度	37
方差	37
操作实例：计算甲班及乙班成绩的方差.....	37
标准差	38
操作实例：计算甲班及乙班成绩的标准差.....	38
3-3 总结	38
第4章 概率分布	39
4-1 概率的基本概念	40
随机实验	40
事件及样本空间.....	40
4-2 计数原理	40
独立事件出现的概率.....	41
操作实例：投掷骰子三次都得到最大点数的获胜概率.....	41
排列	42
操作实例： $n!$ 自定义函数的使用	42
操作实例：使用 $\frac{n!}{(n-x)!}$ 自定义函数	44
组合	45
操作实例：使用 $\frac{n!}{x!(n-x)!}$ 自定义函数	46
4-3 概率分布	47
二项式分布	47
操作实例：计算灯泡工厂产品不合格率.....	48
二项式分布的期望值与方差.....	49
操作实例：计算灯泡工厂不合格产品生产的期望值与方差.....	50
操作实例：评估保险公司收取的费用是否合理.....	50
4-4 超几何检验	51
操作实例：用超几何检验计算灯泡工厂产品不合格率.....	51
超几何分布的期望值与方差.....	52

第三篇 估计与检验篇

第 5 章 正态分布	57
5-1 正态分布	58
正态分布的意义	58
正态分布的特性	59
操作实例：制作正态分布图	59
5-2 标准正态分布	62
操作实例：使用 NORMSDIST 函数实例	62
利用 Excel 制作标准正态分布图	64
操作实例：制作标准正态分布图	64
5-3 正态分布计算实例	65
操作实例：究竟考多少分才能上大学	66
第 6 章 抽样与一个总体的估计检验	69
6-1 抽样与抽样分布	70
6-2 抽样的方式	70
抽签式	70
使用随机数表	70
利用 Excel 产生随机数	71
操作实例：使用 Excel 抽出随机数	71
6-3 抽样误差	73
6-4 样本平均数与方差	73
操作实例：样本平均数的抽样分布	74
中央极限定理	76
6-5 样本方差及标准差计算实例	76
无限总体的方差及标准差的计算	77
操作实例：计算无限总体的方差及标准差	77
有限总体的方差及标准差的计算	78
操作实例：计算有限总体的方差及标准差	78
第 7 章 估 计	81
7-1 统计估计	81
点估计	82
操作实例：估计每家分店的销售收入	82
区间估计	83
7-2 大样本的总体平均数区间估计	83
利用 Confidence 函数计算置信区间	84
操作实例：计算布公主咖啡店 95%置信度下的平均销售数额	85

7-3	小样本的总体平均数区间估计.....	86
	t 分布	86
	TINV 函数	86
	操作实例：使用 TINV 函数取代查表	86
	操作实例：t 检验实例操作	87
7-4	总体方差区间估计——卡方分布	89
	操作实例：使用 CHIINV 函数计算 95%置信度下的卡方分布	90
第 8 章	检 验	93
8-1	假设检验的意义	94
	单测检验与双测检验	94
	检验的步骤	95
	P 值法	95
8-2	总体平均数的假设检验——Z 检验	96
	操作实例：使用 Excel 进行品质检验	96
8-3	小样本平均数假设检验——t 检验	98
	操作实例：使用 t 检验评估选手实力	98
8-4	方差及标准差的检验—— χ^2 检验	100
	操作实例：使用 χ^2 检验评估产品	100
第 9 章	两总体的检验	103
9-1	两总体间假设检验的方法	104
	检验的步骤	104
9-2	两总体平均差的假设检验	104
9-3	两独立总体平均差的假设检验——Z 检验	106
	操作实例：灯泡工厂的生产线产品检测	106
9-4	小样本平均差的假设检验——t 检验	109
	操作实例：市场分析师的准确度检验	109
9-5	两独立样本总体方差比或标准差比的假设检验—— χ^2 分布	111
	操作实例：生产线的预测差异程度检验	111
第 10 章	方差分析——ANOVA	115
10-1	两个以上正态总体的平均数分析——方差分析	116
10-2	单因素方差分析	116
	案例解析：工厂生产流程检验	116
10-3	双因素方差分析——交叉分析	118
	案例解析：生产差异因素控制检验	118
	操作实例：用 ANOVA 工具进行双因素方差分析	119
10-4	双因素方差分析——非交叉分析	120
	操作实例：用 ANOVA 工具进行双因素方差非交叉分析	121

第四篇 回归与时间序列篇

第 11 章 回归分析	125
11-1 回归分析的意义及用途	126
回归直线	126
相关系数	127
11-2 广告播放次数回归分析案例研究	127
操作实例：使用回归分析分析广告播放次数与销售量之间的关系	128
第 12 章 时间序列	133
12-1 时间序列的意义与性质	134
12-2 测定时间趋势的方法	135
线性趋势法	135
操作实例：使用线性趋势法预测未来营业方针	135
平滑法	137
操作实例：使用移动平均法求得未来数据	137
指数平滑法	139
操作实例：使用指数平滑法求得未来数据	139

第五篇 高级应用篇

第 13 章 数据透视表应用实例——咖啡店消费者分析实例	143
13-1 问卷数据整理——客户满意调查实务	144
整理问卷数据	144
13-2 制作数据透视表	145
激活数据透视表向导	145
操作实例：激活数据透视表向导	146
咖啡总销售量的计算	148
操作实例：利用数据透视表进行求和	148
客户消费行为统计	149
操作实例：运用数据透视表计算消费行为	149
性别与口味是否有所差异	151
操作实例：数据透视交叉分析功能运用	151
第 14 章 自定义函数的应用	155
14-1 自定义函数	156
14-2 Function 函数的制作与使用	156
操作实例：制作 M! 函数	156
14-3 Function 程序解析	158
自定义函数的结构	158

自定义函数中可用的数据类型.....	158
算术运算符	159
流程控制	160
14-4 动手制作简单的函数	163
操作实例：制作温度度量衡转换公式.....	163
第 15 章 数据筛选工具	167
15-1 导入数据的途径	168
从数据库和文件导入数据.....	168
用 Microsoft Query 导入数据	169
从 Web 导入数据.....	170
15-2 使用自动筛选工具	171
操作实例：自全数问卷数据中筛选条件为大专的数据.....	172
操作实例：自全数问卷数据中筛选出自定义条件的数据.....	174
第 16 章 简单制作统计模板	177
16-1 成绩表模板概念解说	178
16-2 制作计算成绩工作表	178
定义字段	179
插入公式	179
制作成绩统计信息.....	181
制作次数分布表.....	182
绘制次数分布图形.....	184
保护公式单元格.....	187
保护工作表	188
局部保护工作表.....	189
取消工作表保护.....	190
16-3 制作成绩表模板	191
模板的制作及使用.....	191
自动激活特定模板.....	192
清除自动加载工作簿的功能.....	193
16-4 与 Office 文件的结合应用	194
制作并编辑演示文稿.....	194

第六篇 附录篇

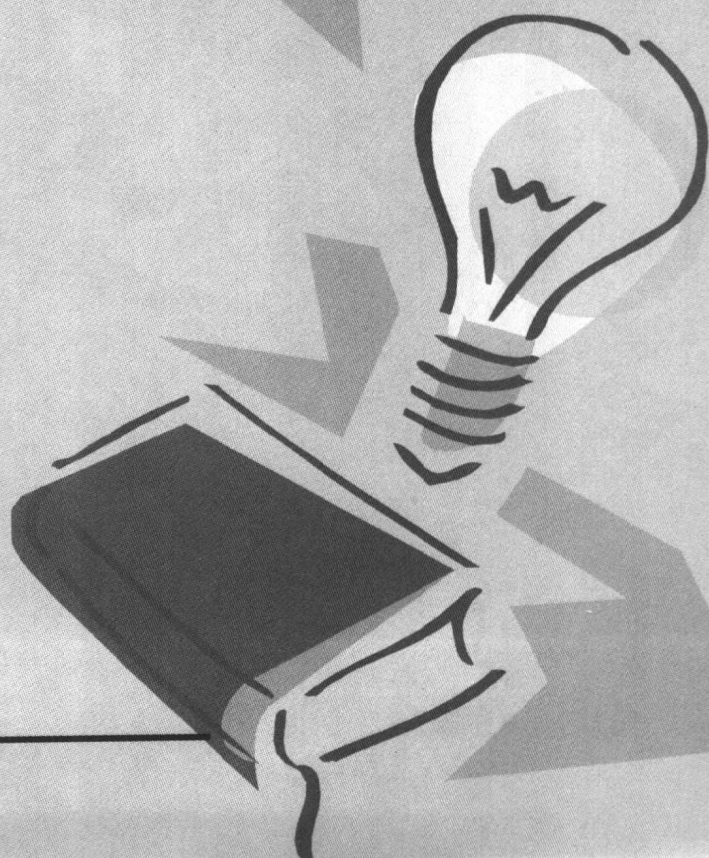
附录 A Excel 实用统计函数整理	199
AVERAGE	200
BINOMDIST	201
CHIDIST.....	201

CHIINV	202
CHITEST	202
CONFIDENCE	202
COUNT	203
FINV	203
FDIST	203
FTEST	204
MAX	204
MEDIAN	205
MIN	205
MODE	206
NORMDIST	206
NORMINV	207
NORMSDIST	207
NORMSINV	208
STDEV	208
TDIST	208
TINV	209
TTEST	209
VAR	210
ZTEST	210
附录 B 常用数据函数整理	211
CONVERT	212
IF	213
HLOOKUP	214
VLOOKUP	214
ABS	215
ROUND	215
SQRT	216
SUM	216
附录 C 公式跟踪修订功能介绍	217
跟踪公式错误的方法	218
检查单元格范围引用	218
公式审核	219

第一篇

基础操作篇

- 第 1 章 Excel 数据表及基础统计数据整理原则
- 第 2 章 使用 Excel 整理及显示统计数据



Chapter 1

Excel 数据表及基础统计数据整理原则

本章知识要点

- 统计学概念介绍
- 统计学中的总体与样本概念解析
- Excel 与统计学的关系
- 应用 Excel 所需注意的事项



1-1 统计学概念介绍

2002年12月7日的台湾地区台北和高雄市长选举,台北市现任市长马英九,得票率约为64%,连任下一届台北市市长。高雄市现任市长谢长廷,得票数与对手差距约为4个百分点……。

自2002年7月1日起,12月31日止,台湾地区平均失业率约为5%,创历年来新高……。

打开电视,民意调查数字、失业、就业人口等这些数字,几乎每天都可以在新闻中听到,而这些数字,多多少少反应出一些社会现象。如失业人口数反应出目前台湾地区经济的景气状况,台北、高雄市长的得票结果,也反应出了市民对于市长的支持度。而这些数字,则是通过统计学的计算方法计算出来的。

通过统计学所计算出来的数字,在我们的日常生活中,几乎随处可见,但一般人可能觉得,统计学似乎必须是专业的人士才用得到,我们只是普通百姓,统计离我们似乎还很遥远。

那么,让我们来看看这样的一个例子。

卖小吃,或者是提供饮食服务,在台湾地区是最普遍的情形,而餐旅学校的学生,在毕业前,也多半必须在餐饮店里服务,以培养经营商店的能力。

就读于餐旅学校的小明,在暑假打工时,就到忠孝东路边的咖啡馆去实习。在实习一段时间之后,他发现,当地的客户在点餐时似乎比较习惯于守旧的口味,对于新产品,似乎总是缺乏兴趣。因此,店里想要举办咖啡促销活动,让消费者能因为价格降低的因素而愿意品尝新的产品。身为店长的小明,要怎么从所有的商品中选择最合适促销的商品进行促销呢?

此时,小明会不计成本,以随机的方式随便选出一种产品,作为主要促销的商品?还是选择先计算出店里热卖的商品种类,以便于推出组合式套餐,让每项商品都可以因为促销,而有机会让消费者品尝到?

其实,以上所述的任何方法几乎都与统计学息息相关。所谓统计学,简单的解释,指的就是搜集、整理、显示及分析数据的方法。也就是说,通过统计学,我们可以将手边的数据,归纳整理为可提供商业决策,或者是政府施政方向参考的信息。

1-2 统计学中的总体与样本概念解析

在对统计学有了简单的了解之后,接下来,就让我们来介绍在统计学中,最重要的概念——总体及样本。

所谓“总体”,指的就是由具有某些相同特质的元素或个体所组成的群体,也就是作为观察对象的全体集合。让我们先回到前一节的咖啡店的例子,当身为咖啡店店长的小明,想要了解店里哪项商品最为热卖时,可行的方法将是,先计算出店中一周内所卖出的咖啡总杯数,而这里的总杯数,即统计概念中所谓的“总体”。

那什么又是“样本”?所谓“样本”,即为由总体中抽取出来组成的集合。就拿上面咖啡店的例子来分析,若店内的商品包含了特调咖啡、美式咖啡、单品咖啡、巴西咖啡、曼特宁咖啡、摩卡及低咖啡等6种商品,则每一种商品均可以被视为总体中的一个单一样本,如图1-1所示。

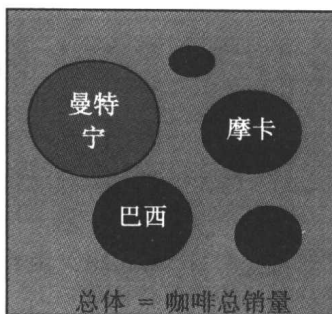


图 1-1

最后，我们可以做出这样的归类——在前面的问题中，身为咖啡店的店长，在提出促销方案之前，首先要做的工作是先分别统计出店中一周内销售量最好的产品及销售量最差的产品，再做出总销售数字的统计。接着，针对这两项商品，提出红配绿的促销方案，让消费者有机会以合理的价格，品尝新口味的咖啡。

在这个例子当中，一周内所销售出的总咖啡杯数，即为总体，而每项商品，则为这个统计事件中的样本。

1-3 Excel 与统计学的关系

在对统计中的意义及重要概念，有了最初步的了解后，接下来，就让我们来看看要利用什么工具，才能计算出我们所要的数字，顺利地在促销商品的选择上做出决策。

如果说咖啡店一周内所销售出的总咖啡杯数大概只有几十杯的话，要做出上述的决策，就非常的简单；如果一周内所销售的咖啡数量，约为千杯以上时，整个计数的工作，就会相当的痛苦和麻烦，此时，我们就可以利用 Excel 作为我们统计上的计数工具，计算出我们所要的结果。

如图 1-2 所示，即为以 Excel 工作表，作为我们统计上的计数工具，分别将统计期间为 2002/12/21~12/27 一周内销售商品的内容列在表格当中（即总体内容）。往后，若我们想要计算出在这个期间每项商品的销售总数时，利用 Excel 中的 countif()函数，就可以快速地帮我们计算出所要的结果。如在上面的例子当中，我们利用 countif()函数，求出每种咖啡在本周内的销售总杯数（即样本数量），如图 1-3 所示。

原始母
体数据

图 1-2



图 1-3

通过这个例子，不难看出，当我们想要利用统计方法求得想要的答案时，除了利用人工的计数方式，一笔一笔地求出次数汇总外，善用手边的应用程序，更可以简化得到结果的过程，得到应有的效率。

1-4 应用 Excel 所需注意的事项

通过上述的范例，我们可以发现 Excel 强大的运算能力，可以用来辅助我们得出所需的统计结果。本书的写作目的，也就是针对如何应用 Excel 中内建的功能及工具，为您计算出所要的统计分析结果。

“工欲善其事，必先利其器”，在进一步介绍如何利用 Excel 中的工具计算出所需的统计数据前，笔者要提醒读者在使用 Excel 时所必须注意的一些问题。

正确定义输入的统计数据

当我们要使用 Excel 工作表计算统计数据时，通常会将每一个统计的样本数值输入工作表的单元格当中，以便进一步地利用这些数据。但在输入这些数据时，必须注意到所输入的数据是否符合单元格中的数据类型规范。

如果所要输入的数据属于文本类型，请在数据前加注“'”单引号（见图 1-4），即可将数据表示为文字类型的数据。如在一般的问卷统计数据当中，常常会以编号的方式来代表某项答案，此时，这些编号即代表文本，而非数值数据。因此，在输入这些问卷的答案数据时，所要表现的数据内容虽然表面上看来为数值数据，但实际上却代表着每一个问题的答案。因此，在输入数据进行往后的数据计算时，我们必须用文本的方式来表现这些数据。

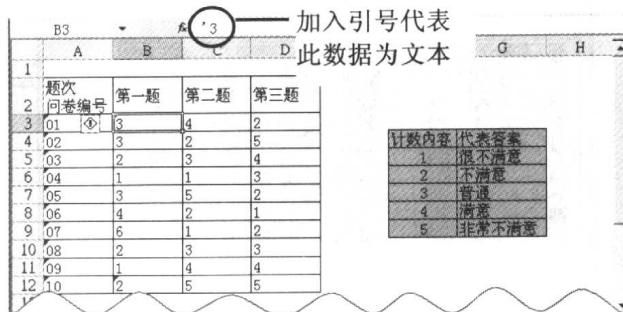


图 1-4