

21 世纪高职高专精品教材·经贸类通用系列

应用统计和 Excel 运用

主 编 方向阳
副主编 凌彩金
主 审 黄莉萍

中国人民大学出版社

· 北京 ·

图书在版编目 (CIP) 数据

应用统计和 Excel 运用 / 方向阳主编

北京: 中国人民大学出版社, 2009

21 世纪高职高专精品教材·经贸类通用系列

ISBN 978-7-300-11476-7

I. 应…

II. 方…

III. ①应用统计学-高等学校: 技术学校-教材

②电子表格系统, Excel-应用-统计分析-高等学校: 技术学校-教材

IV. C8

中国版本图书馆 CIP 数据核字 (2009) 第 213357 号

21 世纪高职高专精品教材·经贸类通用系列

应用统计和 Excel 运用

主 编 方向阳

副主编 凌彩金

主 审 黄莉萍

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

邮政编码 100080

电 话 010-62511242 (总编室)

010-62511398 (质管部)

010-82501766 (邮购部)

010-62514148 (门市部)

010-62515195 (发行公司)

010-62515275 (盗版举报)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

经 销 新华书店

印 刷 北京东君印刷有限公司

规 格 170 mm×228 mm 16 开本

版 次 2010 年 1 月第 1 版

印 张 12

印 次 2010 年 1 月第 1 次印刷

字 数 213 000

定 价 20.00 元

版权所有 侵权必究

印装差错 负责调换

前言

随着知识经济和网络时代的到来,无论是自然科学领域、社会科学领域的研究,还是宏观经济管理和企业生产经营管理,甚至人们的日常生活,信息需求量日益增多,信息处理技术愈加复杂,作为信息技术支柱的统计方法的应用也越来越广泛。统计学作为研究事物数量规律的一门方法论科学,是对各类社会经济现象进行定量分析的重要工具,市场经济越发展、越完善,统计工作的作用就越广泛、越重要。计算机的普及使人类从繁重的体力和脑力劳动中解放出来,如今人们可以借助计算机及其统计分析系统或统计软件包完成繁冗的数据处理,在海量的数据中获得有价值的信息,从而分析、解决问题。在西方发达国家,SAS和SPSS是大学学生的必修课和毕业后谋生的“摇钱树”。在我国,SAS和SPSS也逐渐成为统计从业人员和大学学生的“看家本领”。本书强调统计知识的应用,将统计知识与计算机相结合,着力提高学生运用计算机解决统计问题的能力。SAS和SPSS等专业统计分析系统或统计软件包由于其费用昂贵、推广不易,本教材所有涉及的统计计算,都由目前使用最广的Excel软件的统计功能完成。

本书是笔者在多年的教学实践基础上,将授课讲义进一步充实凝练而成。每章都由案例导入,然后介绍有关统计思想与基本概念、方法,用Excel完成统计计算,配备章节上机实验、小结和习题,并附有习题答案与教学课件。编者力图使本书成为一本融Excel的高级应用与初、中级统计学于一体的教材,成为学过统计者的计算参考书和未学过统计者的“傻瓜式”统计教程。本书由方向阳主编,凌彩金担任副主编,刘大建、周礼静等参加了有关章节的撰写,全书由方向阳修改、统稿、定稿,黄莉萍老师审阅了本书。凌彩金、许秀芝、任峰编制了本书的课件。本书可供经管专业在校本、专科学生使用,特别适合自学统计分析的读者和统计学爱好者使用。

本书在编写过程中,参考了国内外相关文献资料,均在书后一一列出。由于水平有限,书中难免有疏漏或错误之处,恳请读者多提宝贵意见,以便今后进一步修改与完善。

作者

2009年8月

目 录

第一章 市场调查与问卷设计 \ 1

第一节 调查研究的主要方法 \ 2

第二节 问卷调查 \ 3

第三节 问卷调查表的创建 \ 10

第四节 上机实验一 用 Excel 制作电子调查问卷 \ 22

第二章 数据与数据整理 \ 26

第一节 数据和数据的类型 \ 27

第二节 用 Excel 收集、整理数据 \ 32

第三节 上机实验二 用 Excel 进行数据整理 \ 46

第三章 描述统计 \ 51

第一节 数据描述 \ 51

第二节 用 Excel 进行数据初步分析 \ 60

第三节 上机实验三 用 Excel 进行数据统计描述 \ 72

第四章 参数估计 \ 75

第一节 几种常见的概率分布 \ 75

第二节 参数估计 \ 85

第三节 用 Excel 进行参数估计 \ 92

第四节 上机实验四 用 Excel 进行参数估计 \ 100



第五章 假设检验 \ 103

- 第一节 假设检验的基本概念 \ 103
- 第二节 一个正态总体的统计假设检验 \ 107
- 第三节 两个正态总体的统计假设检验 \ 112
- 第四节 用 Excel 做假设检验 \ 117
- 第五节 上机实验五 用 Excel 进行假设检验 \ 124

第六章 方差分析 \ 127

- 第一节 方差分析基础 \ 127
- 第二节 用 Excel 做方差分析 \ 132
- 第三节 上机实验六 用 Excel 进行方差分析 \ 141

第七章 相关分析 \ 145

- 第一节 相关分析基础 \ 145
- 第二节 用 Excel 对普通相关分析的处理 \ 151
- 第三节 上机实验七 用 Excel 进行相关分析 \ 154

第八章 回归分析 \ 156

- 第一节 回归分析基础 \ 156
- 第二节 用 Excel 进行回归分析处理 \ 164
- 第三节 上机实验八 用 Excel 进行回归分析 \ 172

附录 思考与练习题参考答案 \ 175

参考文献 \ 181

第一章

市场调查与问卷设计

“没有调查就没有发言权”，要想了解社会、了解市场，应先学会调查。市场调查是市场运作中必不可少的一个环节，同时也是一个项目策划前期准备工作的重要组成部分，只有在深入、彻底的市场调查的前提下，营销策划和市场操作才有规可循，也只有这样，才有可能获得成功。学会调查研究，可掌握第一手的数据信息，为统计分析提供数据，从而获得科学的指导，把握机遇，以便在激烈的市场竞争中稳操胜券。统计就是从数据中获取信息的一种方法，这就是统计的概念。

本章首先向读者介绍市场调查的实例，然后介绍市场调查的主要方法，最后介绍问卷调查表的创建。

【实例描述】

某超市与一个大学的校企合作超市

“工学结合、校企合作”是最近几年职业教育的热点，为了提高相关专业生源的实践能力，推进高校与企业间的紧密联系，促进高校教学与社会实际工作进一步紧密结合，某超市与一个注册学生6 000人的大学欲签订一份校企合作协议。

校企双方合作在学校指定地点开设超市，由超市所属公司负责装修、货架、商品、促销等方面的投入，学校负责提供一个面积不小于100平方米的店面并负责水电。本着公益与实践的原则，超市不以盈利为目的，超市的销售毛利在支付学校勤工俭学人员的工资后用于校企双方设立的助学金/奖学金。

超市公司前期所做校园超市可行性分析如下：

目前的校园超市大多为非专业经营，商品品质、品类结构、营销促销等方面都处于低端运作阶段，主要依靠为学生提供的便利性服务来获得高毛利的经营回报。由于不规范的经营，高毛利、低品质的商品满足不了学生的需求，大多数学生顾客处于被动消费的地位。非专业的经营无法衍生超市延伸服务项目。

店铺预计投入 27.5 万元，其中包括货架及固定资产投入 5 万元、商品投入 15 万元、信息设备 1.5 万元、装修投入 5 万元、营销投入 1 万元，以上投入完全由超市公司负责。

学校拥有学生逾 6 000 人，以每人每天超市消费 1 元计算，一个校园超市零售的日市场容量在 6 000 元以上。

以日均有效客流 800 人计，客单 10 元，每日营业额 8 000 元，扣除销售费用后，按 3% 的纯利计算，每年盈利： $8\,000 \times 3\% \times 30 \times 9 = 64\,800$ （元）。

因此，专业的零售进入学校存在着较大生存空间。

这个分析唯一的问题是，超市不清楚该大学每周的销售额到底是多少，大学生的消费习惯与普通市民的差异是什么。该超市委托学校市场营销专业的学生对该校学生进行市场调查以获得所需要的信息。而开展调查工作的首要任务是编制调查问卷。

第一节 调查研究的主要方法

调查研究的方法很多，但主要有访谈调查法、问卷调查法和文献追踪法等。

一、访谈调查法

访谈调查法又称访谈法、谈话法或访问法，是指通过与研究对象交谈搜集所需资料的调查方法。访谈是一种研究性交谈，也就是两个人（或更多人）之间一种有目的的谈话，其中由访谈员一方通过询问来引导被访者回答，以此了解调查对象的行为或态度，最终达到调查目的。

访谈一般以面对面的个别访谈为主，也可采用小型座谈会、调查会的形式进行团体访谈，还有通过电话进行的电话访谈。访谈既可以作为一种独立的研究方法，也可以作为其他研究方法中搜集资料的辅助方法。

访谈调查法有灵活、准确、深入的优点，但访谈调查法也有成本较高、缺乏隐秘性、受访谈员影响大、记录困难和处理结果难等局限。

二、问卷调查法

问卷调查法又称书面调查法或填表法，是指用书面形式间接搜集研究材料的一种调查手段。它是通过向调查者发出简明扼要的征询单（表），请其填写对有关问题的意见和建议来间接获得材料和信息的一种方法。其优点是清晰、

直观,不易遗漏,可作为调查研究的关键手段。问卷一般有三种形式:

(1) 报刊问卷。即在报纸或刊物上公布调查表、号召读者做出书面问答,并指定地址寄回答案。

(2) 邮寄问卷。即把已印好的调查表寄给一定类型的对象,并请他们填写答案后寄回调查表。

(3) 发送问卷。即由研究人员把调查表发给集中在一处的一群调查对象,要求他们当场填写后直接收回。

三、文献追踪法

文献是宝贵的信息资源,它记载着各个企业、部门的生产史和经营史。文献包括各类经济报刊和广告,特别是政府部门的统计数据及新闻报刊。文献追踪法是对各种有价值的文史档案进行追踪调查,从中分门别类地进行分析对比、去粗存精,以预测近期市场的发展趋势,进而制定出相应的竞争对策。文献追踪法对社会调查和市场营销具有普遍的实用价值。

下面将着重介绍最常用的问卷调查法。

第二节 问卷调查

问卷调查法是把被调查的问题有条不紊地罗列在调查表中,让被访者如实回答。问卷调查的施行步骤则分为研究问题的设定、编制问卷、问卷的修改与试测、正式施测、数据统计分析、形成调查研究报告。本节将重点介绍问卷的类型、结构和设计制作方法。

一、问卷调查的类型

问卷(questionnaire)一词源于法文,意为“一种为了统计或调查用的问题表格”。它研究有机体变量(性别、年龄、文化程度等)和反应变量(意见、期望、动机、兴趣、态度、信念、行为等)。

问卷调查,按照问卷填答者的不同,可分为自填式问卷调查和代填式问卷调查。其中:自填式问卷调查,按照问卷传递方式的不同,可分为报刊问卷调查、邮政问卷调查和送发问卷调查;代填式问卷调查,按照与被调查者交谈方式的不同,可分为访问问卷调查和电话问卷调查。

上述几种问卷调查方法的利弊,可简略概括如表1-1所示:

表 1—1 各种问卷调查方式的利弊

项目	自填式问卷调查			代填式问卷调查	
	报刊问卷调查	邮政问卷调查	送发问卷调查	访问问卷调查	电话问卷调查
调查范围	很广	较广	窄	较窄	可广可窄
调查对象	难控制和选择, 代表性差	有一定控制和选择, 但回复问卷的代表性难以估计	可控制和选择, 但过于集中	可控制和选择, 代表性较强	可控制和选择, 代表性较强
影响回答的因素	无法了解、控制和判断	难以了解、控制和判断	有一定了解、控制和判断	便于了解、控制和判断	不太好了解、控制和判断
回复率	很低	较低	高	高	较高
回答质量	较高	较高	较低	不稳定	很不稳定
投入人力	较少	较少	较少	多	较多
调查费用	较低	较高	较低	高	较高
调查时间	较长	较长	短	较短	较短

二、问卷的编制

(一) 研究问题的确定

即确定要研究问题的具体内容, 这是确定问卷结构与项目的重要依据。如对中学生“追星”的调查研究中:

现状描述:

究竟有多少学生加入了“追星”行列?

“追星”现象究竟到了何种程度?

它对学生的身心健康、学业等是否造成了不良影响?

原因分析:

“追星”受到哪些因素的影响呢?

究竟是明星的哪些特质吸引学生去追逐?

对策研究:

在原因分析的基础上, 找到一些具体的措施与方法能将学生的不当行为引导到正常轨道上来, 建设良好的校园文化。

(二) 问卷的内容

在一般情况下, 问卷的基本结构由以下几部分组成: 标题、封面信、指导语、问题、结语。

1. 标题

标题要直接、明白地表明自己的研究内容，应该与研究目的相符，简单明了。如：“中学生‘追星’现状调查”。

2. 封面信

封面信 (cover letter) 的 WWW 三原则：要说明调查者的身份 (who)；要说明调查的大致内容和进行这项调查的目的 (why)；要说明调查对象的选取方法和对调查结果保密的措施 (how)。

3. 指导语

指导语是指导被调查者填写问卷的说明，作用与仪器的使用说明书相似。有些指导语集中在封面信之后，并标有“填表说明”的标题，其作用是对填表的方法、要求、注意事项等作一个总体说明。如：

同学，你好：

为了解你对“追星”这一普遍的社会现象的态度和看法，加深我们对你的了解，创建一个良好的校园文化氛围，请你协助进行这次调查。本次调查不记名，所以你不必顾忌，只需按照自己的真实情况，如实回答就可以了。谢谢！

填表说明：

- a. 请在每一个问题后适合自己情况的答案号码上画圈或者在画线处填上适当的内容。
- b. 问卷每页右边的数码及短横线是计算机使用的，你不必填写。
- c. 若无特殊说明，一个问题只能选择一个答案。
- d. 问题没有对错之分，填写问卷时，请不要与他人商量。

4. 问题

问题一般包括两部分资料。

一是个人的背景资料：可放在问卷前面，也可放在问卷的后面。

如：性别：男 _____ 女 _____

二是具体调查的内容。

例如：你对自己喜欢的明星的熟悉程度如何？

- a. 只是认识，知道名字，看过他们的表演或节目
- b. 知道他们的基本情况，例如身高、体重、爱好、家庭等
- c. 努力地搜集有关他们的一切消息和新闻



5. 结语

结语用一段短语内容表示。包括对被调查者的合作再次表示感谢，以及关于不要漏填与复核的请求。

如：题目已完，谢谢你真实、有效的回答，万分感谢！

结语还可以提出关于本次调查形式与内容、感受等方面的问题，以征询被调查者的意见。

如：题目已完，如果您对本次调查有什么意见或想法，请写在下面。谢谢！

三、问题答案的设计

(一) 填空式

即在问题后面画一短横线，让回答者填写。

例如：a. 您的年龄是_____岁？
 b. 您家有几口人？_____人
 c. 您有几个孩子？_____个
 d. 您每天上班需要_____（分钟）时间？

(二) 是否式

答案只有“是”和“不是”（或者其他肯定和否定形式）两种，回答者根据自己的情况选择其一。一般在民意测验中使用最多，回答简单明了，可以严格地把回答者分成两类不同的群体，缺点是得到的信息量太小，不能了解和分析回答者中客观存在的不同层次。

例如：a. 您是共青团员吗？ 是 不是
 b. 您住在本市吗？ 是 不是
 c. 您家有空调吗？ 有 没有
 d. 您赞成学分制吗？ 赞成 不赞成

(三) 多项单选式

即给出的答案至少在两个以上，回答者根据自己的情况选择其一，这也是问卷中采用最多的一种问题形式。

例如：(1) 您的文化程度是（ ）。

a. 小学及以下 b. 初中
 c. 高中或中专 d. 大专以上

(四) 多项限选式

即在所列举的答案中,回答者根据自己的情况选择若干个答案。

例如:您喜欢看哪一类电视节目?(请最多选择三项打√)

- | | |
|---------|------------------|
| a. 新闻节目 | b. 电视剧 |
| c. 体育节目 | d. 广告节目 |
| e. 教育节目 | f. 歌舞节目 |
| g. 少儿节目 | h. 其他(请注明) _____ |

(五) 多项排序式

即对一个问题提供若干个答案项,要求被回答者根据自己认为的重要性程度对其排序。

例如:学习动机的调查

你每天在学校或家里读书是怎样一种兴趣?在下列答案中按照你自己的情况把它们排上顺序。对你来说最主要的就在()中写入1,其次比较主要的就在()中写入2,余者类推。

- () 因为取得了好成绩,心里高兴,而读书能取得好成绩。
- () 因为对书有兴趣。
- () 没有什么想法,只是默默地读。
- () 因为获得新知识,懂得新事物,觉得很快乐。

(六) 矩阵式

这是一种将同一类型的若干问题集中在一起,构成一个问题的表达方式。

例如:您对苏州电信提供的下列服务看法如何?(请在所选方框内打√)

	很满意	满意	基本满意	不满意	很不满意
a. 装机移机服务	<input type="checkbox"/>				
b. 话费查询服务	<input type="checkbox"/>				
c. 电话障碍修复	<input type="checkbox"/>				
d. 公用电话服务	<input type="checkbox"/>				

矩阵式的优点是节省问卷的篇幅,同时由于同类问题集中在一起,回答方式也相同,因此也节省了回答者阅读和填写的时间。

(七) 表格式

这是矩阵式的一种变体，其形式与矩阵式十分相似。

比如，上述矩阵式问题对应的表格式问题就是：

例如：您对苏州电信的下列服务看法如何？（请在所选方框内打√/）

	很满意	满意	基本满意	不满意	很不满意
(1) 装机移机服务					
(2) 话费查询服务					
(3) 电话障碍修复					
(4) 公用电话服务					

表格形式的问题除了具有矩阵形式的特点外，还显得更为整齐、醒目。应当注意的是，这两种形式虽然具有简单、集中的优点，但是也同时使人产生呆板、单调的感觉，在一份问卷中这两种形式的问题不易用得太多。另外，这两种形式只能减少问题在问卷中的篇幅，并不能减少其数量，比如上述例子中实际上就包含了四个问题。

(八) 条件式（追问式）

有些问题仅适合一部分回答者回答，因此，要使用有条件的题目（过滤性题目）。再者，有时对某些问题需要更深入的了解，也需要用有条件的题目。

例如：你有喜欢和崇拜的明星吗？ a. 有 b. 没有
如果有，请写下他们的名字 _____

四、利克特量表

1932年R. 利克特提出了一个简化的测量方法，被称为相加法。它不需要搜集对每个项目的预先判断，只是把每个项目的评定相加而得出一个总分数。利克特量表（Likert-type scale）由一系列陈述组成，调查者利用5点或7点量表让被试者作出反应。5点量表是从“强烈赞同（5）”、“赞同（4）”、“中性（3）”、“不赞同（2）”到“强烈不赞同（1）”。一般以5点量表最常用。

例如：
和其他同学比较起来，你每天花在听歌，观看电视、电影，阅读报纸杂志，游览网页上的时间如何？
a. 非常少 b. 很少 c. 差不多 d. 很多 e. 非常多

五、问卷试测与修改

问卷试测的目的在于发现问卷中潜在的问题：受访者对于问题的了解是否有困难；问题是否有模棱两可之处；问卷是否能够达到研究目标。试测对象不需要多，但应尽量考虑样本的多样性。调查者可根据试测情况对问卷做适当的修改，因而试测的过程不可以省略。

上例中，可以在目标学校的各个年级各找数名学生，其中男女各半，人数在50~100人，先对他们进行集中测试，或者也可以采用个别访问的方式进行测试。可要求学生问卷提出意见。试测之后及时对问卷的题目逐条进行检查，好的保留，不好的淘汰，需要修改的地方及时修改。

六、正式施测

正式施测要考虑抽样的代表性和问卷的回收率，否则问卷调查将失去统计意义。

上例中，如有条件最好使用“随机抽样”或“随机整群抽样”。对回收的问卷，在剔除废卷的同时要统计有效问卷的回收率。一般回收率在30%左右，取得的资料只能作参考；回收率在50%以上，可以采纳建议；回收率在70%以上时，方可作为研究结论的依据。

在问卷调查中，由于各种各样的原因，被调查者可能会漏填各种数据，也就是所谓的数据缺失。缺失数据处理是统计界的一个大课题，数据缺失有无意缺失和故意缺失两种。缺失数据可用以下办法处理：（1）用一个样本统计量去代替；（2）用某些统计模型进行推算；（3）问卷整体剔除；（4）部分排除。

七、抽样

为了确保生产过程正常进行，企业运营经理需要了解所生产产品的质量情况。如果质检人员为了检测产品是否存在质量问题而必须破坏产品的话，那么除了抽样别无他选：因为对所有产品进行检测将会破坏生产的所有产品。做问卷调查也是一样，当我们不能对总体的所有单位都进行调查时，我们一般会抽取总体的部分单位作为样本进行调查（有关总体和样本的概念在后面的章节专门介绍），用这部分单位的信息来推断总体的相关情况。

抽样就是从所研究的对象中随机地取出其中的一部分来观察，由此获得有关总体的信息。在抽样时每个单位有同等被抽中的机会，这是抽样必须遵守

的随机原则，只有遵守随机原则，才能使抽取的部分单位具有充分的代表性。

随机抽样和判断抽样都是从总体中抽取样本的方法。随机抽样是按随机原则，即按概率规律抽取样本，在总体中所有单位被抽中的机会是均等的。被抽中的样本单位不带任何个人或集体的主观意见。在本书中，只限于有关随机抽样的阐述。

随机抽样一般有简单随机抽样、分层随机抽样和整群随机抽样三种不同的抽样方案，简单随机抽样是抽样中最基本的方式。简单随机抽样是对总体的所有容量不作任何的分类和排队，完全按随机原则逐个抽取样本容量。为了保证随机样本的随机性和代表性，在抽取样本时可采用抽签法和随机数字法。抽签法就是将总体容量全部加以编号，并编成相应的号签，然后将号签充分混合后逐个抽取，直到抽到预定需要的样本容量为止。然而，如果总体容量较大，这时编制号签的工作量很大，并且很难充分混合。因此，当总体容量很多时采用这种方法是不合适的。随机数字法是最简便易行，又符合随机原则的方法。这种抽样方式在某种程度上都依赖于随机数表（或用计算机产生的随机数），目的是为了消除抽取样本时的人为偏差。

第三节 问卷调查表的创建

一、随机抽样

如果要统计一所学校教职工的性别比例，由于人数众多，要得到最终结果会是一件十分麻烦的事。一般都是将学校所有教职工看成是一个总体，从中随机抽取一部分教职工作为样本，使用随机数函数“RAND”和索引函数“INDEX”来完成（另用抽样分析工具进行抽样在后面介绍）。

下例就是从一个拥有 148 位教职工的学校随机抽取 11 名教职工进行抽样调查，以推断整个学校的性别构成。

打开学校教职工性别构成工作簿，里面是某校 148 位教职工的姓名和性别记录，要通过随机抽样统计教职工的性别比例，如图 1—1 所示。

第一步：为了方便起见，先对教职工按顺序编号，在 A 列前插入列，在 A2 单元格中输入“序号”，在 A3、A4 单元格中分别输入数字 1 和 2，同时选中这两个单元格，按下鼠标并向下拖动到最后一个教职工的单元格时松开鼠标，给所有教职工编上号码，见图 1—2。

教职工档案	
教职工姓名	性别
李明	1
张诚	1
何佳	2
李明诚	1
李涛	2
张晓群	2
邓捷	1
舒小英	2
李军	1
何明天	1
林立	1

图 1—1 学校教职工性别构成

教职工档案		
序号	教职工姓名	性别
1	李明	1
2	张诚	1
3	何佳	2
4	李明诚	1
5	李涛	2
6	张晓群	2
7	邓捷	1
8	舒小英	2
9	李军	1
10	何明天	1
11	林立	1

图 1—2 插入一列并添加序号

第二步：进行随机抽样，本例要抽取 11 名教职工进行统计。在 E2 单元格中输入“抽样”，选中 E2:E11 单元格区域，单击菜单栏“插入→函数”命令，在系统弹出的“插入函数”对话框的“选择函数”列表框中选择“RAND”函数，如图 1—3 所示，单击“确定”按钮。