

教育部经济管理类核心课程教材

ECONOMICS AND MANAGEMENT

统计学案例与分析

贾俊平 郝 静 等编著



中国人民大学出版社

教育部经济管理类核心课程教材

ECONOMICS AND MANAGEMENT

统计学案例与分析

贾俊平 郝 静 等编著

中国人民大学出版社

• 北京 •

图书在版编目(CIP)数据

统计学案例与分析/贾俊平,郝静等编著.
北京:中国人民大学出版社,2010
教育部经济管理类核心课程教材
ISBN 978-7-300-11653-2

- I. ①统…
II. ①贾…②郝…
III. ①统计学-高等学校-教学参考资料
IV. ①C8

中国版本图书馆 CIP 数据核字 (2009) 第 019004 号

教育部经济管理类核心课程教材
统计学案例与分析
贾俊平 郝 静 等编著
Tongjixue Anli yu Fenxi

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	北京东方圣雅印刷有限公司		
规 格	185 mm×235 mm 16 开本	版 次	2010 年 2 月第 1 版
印 张	13 插页 2	印 次	2010 年 2 月第 1 次印刷
字 数	207 000	定 价	22.00 元

版权所有 侵权必究

印装差错 负责调换

前言

► PREFACE

《统计学案例与分析》是为学习统计学课程的读者而编写的。本书共编写了30个案例，内容涉及宏观经济、体育、医学、大学生活等多个研究领域。每个案例均包括问题背景和数据、需要解决的问题、学习目标、案例分析等几个部分。

目前，统计学案例方面的书籍并不多见，尤其是可直接用于课堂教学和帮助学生理解统计方法类的案例书籍就更少。在国外的统计学教材中，通常在各章最后给出少量案例，其规模一般也不大，可直接用于教学和学习。而国内教材限于篇幅，很少在教材中给出案例，虽有为数不多的案例集方面的书籍，但案例的写法也很难用于课堂教学和学习。原因是其中的案例大多是编写者自己感兴趣或曾专门研究过的一些问题，案例规模偏大，结构复杂而烦琐，更像是一篇论文。这样的案例要读懂首先就很困难，更谈不上用于实际的教学和学习。

我们认为，案例的目的是帮助学生理解课堂上所学的知识，并综合应用这些知识，而不是告诉学生如何进行科学研究。这是本书与目前已有统计学案例书籍的主要区别。基于这样的考虑，本书中的案例，在给出基本背景信息后，直接给出案例中所涉及的数据，然后提出针对这一问题的数据所要进行分析的内容。在案例分析部分，给出了针对这一数据进行分析的思路和可能用到的分析方法。这样写的原因是，我们认为案例的目的是告诉学生：如果你面临这些数据该如何思考和分析，重点是如何综合运用已学过的统计方法来分析这些数据，而不是告诉学生如何研究案例背景所涉及的专业问题。

本书共包括五个部分，即数据收集和描述、参数估计和检验、线性回归分析、时间序列预测、多元统计分析、非参数检验。其中，有些案例是综合性的，涉及几种统计方法的应用。案例使用的分析软件有SPSS，Excel，SAS等。本书



案例的数据文件可在中国人民大学出版社网站《统计学》第三版教学资源库 (www.rdjg.com.cn/zyk/tj) 下载使用。本书可作为学习统计学课程的辅助教材, 也可作为实际工作者的参考书。

参加本案例编写的有贾俊平、郝静、田梦阳、曹伟、宋森、燕敬宇等。全书最后由贾俊平修改定稿。编写一本能用于实际教学的统计学案例, 对我们来说也是一种尝试。因此, 书中的错误或不当之处在所难免, 希望读者批评指正, 并多提宝贵意见, 以便我们不断修改和完善。

贾俊平

于中国人民大学统计学院

目 录

► CONTENTS

第 1 部分	数据收集和描述	1
	案例 1.1 金融危机形势下大学生就业状况调查方案 设计	3
	案例 1.2 北京奥运会奖牌的分布及构成分析	11
	案例 1.3 分析运动员发挥的稳定性	16
	案例 1.4 三个产业生产总值的描述性分析	22
	案例 1.5 基于行业的上市公司基本情况分析	28
第 2 部分	参数估计和检验	37
	案例 2.1 大学生月平均生活费的估计和检验	39
	案例 2.2 有重大科学突破时科学家年龄的估计	51
	案例 2.3 电视广告效果的比较分析	54
	案例 2.4 全国 31 个省区市城镇居民恩格尔系数差异 分析	59
第 3 部分	线性回归分析	63
	案例 3.1 用 PPI 预测 CPI	65
	案例 3.2 预测人均 GDP	72
	案例 3.3 城镇居民消费支出与可支配收入的关系 研究	79
	案例 3.4 促销对相对竞争力的影响研究	84
	案例 3.5 上市公司财务指标的线性回归分析	89



第 4 部分	时间序列预测	97
	案例 4.1 不同交通工具客运量的预测	99
	案例 4.2 预测社会消费品零售总额	108
	案例 4.3 农村居民现金收入的分解预测	116
	案例 4.4 我国人口总量的时间序列分析	122
第 5 部分	多元统计分析	129
	案例 5.1 全国 21 个城市社会经济发展指标的聚类 分析	131
	案例 5.2 我国各地区信息化发展阶段的聚类分析	137
	案例 5.3 我国各地区财政支出的因子分析	143
	案例 5.4 我国各地区农业发展状况研究	149
	案例 5.5 基于细胞的乳腺癌诊断分析	157
	案例 5.6 上市公司运营能力与盈利能力相关性分析	162
	案例 5.7 基于几个社会经济指标的城市判别分析	168
	案例 5.8 全国各地区农村居民消费指标的统计分析	174
第 6 部分	非参数检验	181
	案例 6.1 农村居民家庭年人均各项纯收入的比较	183
	案例 6.2 山东省农村和城镇居民收入指数比较分析	189
	案例 6.3 我国副省级和沿海开放城市实际利用外资额 的比较	194
	案例 6.4 不同选手的射箭成绩之间是否有显著 差异	198



第 1 部分

数据收集和描述

案例1.1 金融危机形势下大学生就业状况调查方案设计

案例1.2 北京奥运会奖牌的分布及构成分析

案例1.3 分析运动员发挥的稳定性

案例1.4 三个产业生产总值的描述性分析

案例1.5 基于行业的上市公司基本情况分析



案例 1.1

CASE 1.1

金融危机形势下大学生 就业状况调查方案设计

调查背景

2008 年 9 月开始,由美国次贷危机引发的全球性金融危机成为人们谈论的焦点。此次危机对全球经济产生了广泛影响,造成金融业萎靡,经济衰退。经济形势的剧烈变化冲击到我国的就业市场,很多企业为了降低成本采取了不同程度的裁员措施。金融危机影响到我国的出口,致使相关企业员工需求降低;另外,金融危机也使得外商投资减少,外资及合资企业提供的岗位数相应减少。在这些冲击下,大学生的就业压力也随之增加。

数据显示,2009 年我国约有 610 万大学应届毕业生准备就业,加上以往还没有就业的大学生,总人数至少达到 700 万。在这个特殊时期,大学生就业显得尤为引人注目,走在求职路上的大学毕业生们也显得比往常更加焦虑。新华网一项关于 2009 年大学生就业形势的调查问卷显示,高达 91.05% 的人认为现在就业形势严峻,就业难;大约一半的人表示会先就业再择业。

那么,在金融危机的影响下,大学生的就业现状究竟如何?他们的心态如何?又将作何选择?2009 年 3 月,我们对中国人民大学应届本科毕业生就金融危机形势下大学生的就业状况进行了调查。通过对他们就业的心态、选择及就业结果的调查,客观、真实地了解金融危机下大学毕业生就业的实际情况,并且利用调查的数据进一步分析和探讨了大学生的就业观念,对就业变化状况等做了分析。



抽样方法

本案例的抽样总体为中国人民大学全体应届本科毕业生,共2 525人。采用的抽样方法为两阶段抽样:第一阶段为分层抽样,以寝室为抽样单位,按专业和性别分层;第二阶段为随机抽样,在第一阶段抽到的寝室内随机抽取学生作为调查对象。本案例共抽取样本370人,采用问卷的方式展开调查,共发放问卷370份,回收有效问卷355份。

调查问卷

关于大学生就业状况的调查问卷

您好!我们是中国人民大学统计学院的学生,正在进行一项关于金融危机下大学生就业状况的调查,需要您协助完成此问卷,您的意见非常重要,谢谢您的配合!

请填写您的基本信息

院系_____专业_____年级_____性别_____

以下各题请在您要选的选项前的“□”上划“√”

1. 大学本科生活即将结束,请问您的毕业去向。

☐在国内读研究生 ☐出国留学 ☐就业

(如果您选择了读研或出国,您只需再回答第2题和第14、15、16、17题,问卷就结束了;如果您选择了就业,请跳过第2题,从第3题继续回答)

2. 您毕业后不打算就业,是否受到了金融危机下就业压力的影响。

☐受到很大影响 ☐有一定的影响 ☐没有影响,本来就是这样打算的

3. 您预期的工作属于以下哪种类型(可多选)。

☐IT/互联网 ☐制造行业 ☐房产建筑 ☐金融行业 ☐广告传媒
☐公务员 ☐医药行业 ☐其他(请注明)_____

4. 您预期的就业单位属于哪种类型(可多选)。

☐国有企业 ☐民营企业 ☐外资企业 ☐合资企业 ☐政府部门
☐其他(请注明)_____

5. 您愿意到中小城市或西部去发展吗?

☐是 ☐否

6. 您是否已经找到工作?



☐是 ☐还没有

(如果您选择的“是”，请继续回答；如果您选择“还没有”，请跳过第7、8、9题，从第10题继续回答)

7. 您找到的工作属于以下哪种类型。

☐ IT/互联网 ☐ 制造行业 ☐ 房产建筑 ☐ 金融行业 ☐ 广告传媒
☐ 公务员 ☐ 医药行业 ☐ 其他(请注明) _____

8. 您所就业的单位属于以下哪种类型。

☐ 国有企业 ☐ 民营企业 ☐ 外资企业 ☐ 合资企业 ☐ 政府部门
☐ 其他(请注明) _____

9. 您对找到的工作满意程度为：

☐ 非常满意 ☐ 比较满意 ☐ 一般 ☐ 不太满意 ☐ 非常不满意

10. 您找工作的原则是：

☐ 不管有多难，一定要坚持找到自己特别满意的工作
☐ 有自己的标准，但如果条件差不多，可以作出让步
☐ 不管什么样的工作都可以，先就业再说

11. 在以下影响择业的因素中，请按顺序列出您最在意的前三项的序号（可以少于三项）。

①薪酬/福利 ②工作地点 ③工作稳定性 ④工作的潜力 ⑤工作压力 ⑥个人的兴趣 ⑦是否与专业对口 ⑧家人、朋友等的影响

12. 您找工作的主要途径有（可多选）：

☐ 直接去公司的相关部门咨询 ☐ 参加各种现场招聘会
☐ 上网投简历 ☐ 通过同学、朋友、家庭的关系
☐ 其他(请注明) _____

13. 在找工作的过程当中，您遇到的主要困难是（请注明）：

14. 在您看来，目前找工作的难易程度如何。

☐ 非常容易 ☐ 比较容易 ☐ 一般 ☐ 不容易 ☐ 非常不容易

15. 您认为金融危机对您的就业有哪些具体影响？（请注明）



16. 您是否知道国家最近出台了一系列促进大学生就业的新政策。

☐不清楚 ☐听说过,但不太了解 ☐知道大概内容 ☐非常了解

17. 请谈谈您对这些政策的看法(请注明)。

问卷到此结束,再次感谢您的配合!

需要分析的问题:

对调查方案进行详细的剖析,提出调查方案设计的基本思路。

一、学习目标

1. 了解数据收集的主要过程。
2. 了解抽样方法的确定思路。
3. 了解问卷设计的过程及应注意的主要问题。
4. 通过本案例的学习,熟悉数据收集的整体过程,了解数据收集过程中要考虑的主要问题,重点体会抽样方法的确定和问卷设计的方法,达到能自主思考、自主设计调查方案、自主收集数据的目的。

二、案例分析

一次成功的调查涉及多方面的问题。就本次研究而言,涉及样本的选取、问卷的设计、调查的具体实施、调查数据的处理和分析、调查报告的撰写等。下面结合本次调查就样本选取、问卷设计等几个主要问题做详细的剖析,以供广大读者设计调查方案时参考。

(一) 调查方法的确定

本案例要了解大学生就业的现状,是面向大量同学的调查,所以焦点小组访



谈话、头脑风暴法、德尔菲法等小组访谈形式的调查方法显然不适合。在这种情况下,面访调查、电访调查和邮寄调查显得较为合适。

本案例的调查对象是同一学校的同学,很容易接触到,考虑到调查的经费问题,没有必要采用电话调查。接下来要对面访调查和邮寄调查的优劣作出比较,邮寄调查分为几种,其中一种为留置问卷调查,这种方式实际上介于邮寄调查和面访调查之间,是指调查员按照面访的方式找到被调查者,说明调查目的和填写要求之后,将问卷留在被调查者手中,由被调查者自主完成,之后由调查员取回填好的问卷。可以看出,这种方式综合了邮寄调查保密性强的优点,同时还结合了面访调查回收率高的优势,而且可以避免由于调查员自身素质的原因带来的偏差。综合上述分析,本案例采用自填问卷的形式展开调查。

(二) 样本量的确定

样本量的大小取决于总体元素的多少,同时也与估计所要求的精度以及调查所需的人力、物力和财力有关。

本案例抽样总体为2 525名应届本科毕业生。根据调查的时间安排以及当时的人力、物力、财力状况,我们认为样本量确定为总体的10%~15%比较合适,在本案例中,我们选取14%这一比例,约350人,考虑到抽取的样本由于种种原因可能成为无效样本,实际中我们抽取了370个样本。

(三) 抽样方法的确定

采用何种抽样方法取决于总体元素的分布、抽样框等多种因素。本案例要从2 525名本科毕业生中抽取样本量为370的样本,可供选择的抽样方法有随机抽样、系统抽样、分层抽样。在确定抽样方法时,主要考虑以下两个因素:样本的有效性和抽样的可操作性。

根据本次调查对象的特点,本案例最终采取两阶段抽样:第一阶段为分层抽样,以寝室为抽样单位,按专业和性别分层,按相应比例抽出各层应该抽取的寝室,得到一系列的群体;第二阶段为随机抽样,在第一阶段抽到的寝室内随机抽取学生作为调查对象。

采取该抽样方法主要出于以下几点考虑:

第一,对于大学生就业这一研究话题,我们认为大学生的就业状况会受到专业、性别等因素的影响,即总体中的个体之间存在显著的差异,所以第一阶段我们采取分层抽样。抽取的样本按相应的概率覆盖所有的专业和不同的性别,抽取



的样本与随机抽样相比更具有代表性。采取分层抽样的另一个优势是在进行数据分析时可以对各层进行单独分析,提高了数据的利用率。

第二,我们的目的是要抽取毕业生,但是考虑到抽样的可实施性,如果直接选择对毕业生进行分层抽样,调查实施时就会涉及毕业生和在校本科生的甄别问题,同时还增加了寻找抽到指定毕业生的困难。因此,在本案例中,我们将调查的地点定在毕业生的寝室,以寝室为抽样单位采取分层抽样抽得相应群体。

第三,在第一阶段抽得寝室后,每个寝室有6名同学,总人数不多,并且他们之间也不存在较大差异,所以第二阶段采取简单随机抽样,由调查员根据实际情况随机从寝室的6名同学中抽取1名作为调查对象。

(四) 问卷设计

问卷设计是调查中十分重要的一个环节。问卷设计的过程中主要考虑问卷的内容和问卷的格式两大部分,其中问卷的内容包括要明确调查的目的、要获取的信息、问卷中各个问题的内容等;问卷的格式主要考虑问卷的版面格式、问卷中问题的形式与措辞、问题的前后顺序等。

在本案例中,我们要完成金融危机下大学生就业现状的调查分析,调查过程中我们想要知道大学生的毕业去向,本科学习生活结束后大家是选择在国内读研、出国留学,还是就业。如果选择就业,他们是不是已经找到了工作,就业的难易程度又如何。对于选择就业的毕业生,我们又想知道他们心仪的工作是什么样子的,工作属于哪种类型,工作单位又是什么类型。谈到理想的工作,难免会想到现实,在金融危机影响下他们的毕业去向又是什么样的,工作职位和就业单位又分别属于什么类型。了解了大学生就业的理想与现实状况之后,我们难免要挖掘背后的原因,大学生就业有什么样的原则,择业过程中最在意什么,找工作都有哪些途径。为了缓解大学生的就业压力,国家颁布了一系列政策,大学生对此又知道多少,这些政策在现实生活中对大学生就业是否起到了积极作用。这些都是我们针对这一调查题目想到的内容,梳理一下思路,去掉头脑风暴过程中重复的部分,补充不足,可以将本案例的调查内容划分为以下五个部分:

- 大学生毕业去向及就业难易度的调查。
- 毕业生就业取向的调查。
- 毕业生就业去向的调查。
- 影响大学生就业的因素调查。
- 大学生对国家就业政策的认知状况调查。



由此,调查问卷要包括的主要内容就非常清晰了,各个部分要调查的细节也显而易见。在问卷设计过程中只要按照条理,把这些内容反映到问卷中就可以了。在这一过程中,要反复检查问卷的内容是否有效地反映了要获取的信息,考虑是否有遗漏。

问卷设计具有很强的艺术性,即使问卷的调查内容是相同的,问卷格式的不同也会影响到调查的质量。良好的问卷格式,可以使被调查者回答问题时心情舒畅,答题过程顺畅无阻,起到事半功倍的效果。在问卷格式这方面,要注意以下几点:

1. 要有卷头语和卷尾语。在卷头语中,向被调查者问好,说明此次调查的目的,劝说其参与调查。在问卷的末尾要向被调查者完成问卷再次表示感谢。
 2. 关于背景信息的调查。一般的问卷都会涉及这一部分,主要用于身份的甄别或数据分析的需要,被调查者背景的调查部分有时会比较敏感,因此通常将这一部分放在问卷的最后。本案例中调查对象是应届本科毕业生,因此在问卷的背景部分,我们要了解被调查者的年级来做身份的甄别,案例中我们认为专业和性别会对大学生就业产生影响。因此在这一部分,我们还需要了解被调查者的院系、专业和性别以满足数据分析的需要。本案例的背景部分问题比较简单且不敏感,不会对被调查者的情绪造成影响,所以放在了问卷的开头。
 3. 关于问卷答题的说明。本案例调查实施时采取问卷自填式,因此要向被调查者介绍问卷填写的方式。涉及整个问卷的说明,要放在问卷的开头,这样可以使被调查者开始就有清楚的了解。涉及个别问题的说明在题后单独说明,比如说跳答题的情况,是否为多选,选择题中其他内容要注明,等等。在对答题要求进行说明时,要用统一的字体,并且采用非问卷主体使用的字体,以达到醒目的目的。
- 问卷题目设置要有条理,避免给被调查者带来思维上的跳跃感,问题的设置要遵循由易到难的顺序,这样可以使被调查者答题时感到更加轻松。问卷的题目不宜过多,否则会给被调查者带来烦躁感,从而放弃调查。
4. 主观题放在问卷的最后。通常调查问卷都包括客观题和主观题两大部分,通过主观题来反映被调查者的主观看法和意愿,一般来说主观题的难度要比客观题高,为了提高被调查者回答问题的积极性,通常将主观题放在问卷的最后,这样做还可以使被调查者回答完整个问卷,对调查内容有全面了解后更准确地表达自己的看法。
 5. 问卷题目主干用词要准确,避免歧义;客观题选项设置要周全,不能让



被调查者无选项可选，若不能列出所有选项，要设置“其他”选项。

基于上述考虑，本次调查最终形成了案例中给出的问卷，并根据这一问卷成功地完成了调查任务，取得了所需的数据。在对调查数据进行分析后，得出了较为切合实际的结论。

——田梦阳