

SHANGWU TONGJIXUE

# 商务统计学

吴爱娟 / 主编



知识产权出版社

全国百佳图书出版单位

SHANGWU TONG.

# 商务统计学

北京市民办教育发展促进项目资金资助

吴爱娟 / 主编



知识产权出版社

全国百佳图书出版单位

## 图书在版编目(CIP)数据

商务统计学/吴爱娟主编. —北京:知识产权出版社,2015.2

ISBN 978 - 7 - 5130 - 3167 - 7

I. ①商… II. ①吴… III. ①商业统计学 IV. ①F712.3

中国版本图书馆 CIP 数据核字(2014)第 269219 号

### 内容提要:

商务统计学是用以收集数据、分析数据和由数据得出结论的一组概念、原理和方法。本书的内容设计比较精简,共分为 10 章节,分别为第 1 章,商务统计学概述;第 2 章,概率论基础;第 3 章,用图表掩饰数据;第 4 章,数值描述度量;第 5 章,时间序列分析;第 6 章,统计指数分析法;第 7 章,抽样分布;第 8 章,参数估计;第 9 章,假设检验;第 10 章,相关与回归分析。

本书针对财经类本科学学生及相关专业读者编写的,我们希望学生通过本书的学习,能运用所学的统计理论和统计方法对案例中待解决问题进行分析和研究,对计算过程和计算结果进行分析和评价。

责任编辑:唐学贵

执行编辑:牛 闯

## 商务统计学

吴爱娟 主编

出版发行:知识产权出版社有限责任公司

电 话:010-82004826

社 址:北京市海淀区马甸南村 1 号

责编电话:010-82000860-8571

发行电话:010-82000860 转 8101/8029

印 刷:保定市中国画美凯印刷有限公司

开 本:720 mm × 1000 mm 1/16

版 次:2015 年 2 月第 1 版

字 数:228 千字

ISBN 978 - 7 - 5130 - 3167 - 7

网 址:<http://www.ipph.cn>

<http://www.laichushu.com>

邮 编:100088

责编邮箱:21183407@qq.com

发行传真:010-82000893/82003279

经 销:各大网上书店、新华书店及相关专业书店

印 张:17

印 次:2015 年 2 月第 1 次印刷

定 价:38.00 元

出版权专有 侵权必究

如有印装质量问题,本社负责调换。

## 本书编委会

主 编：吴爱娟

副主编：徐 慧 张瑞亭 杜聪慧

编 者：贾鲁军 马新苗 赵 昕

# 前 言

本书是为了满足学生对于统计学日益增长而又尚未满足的需求,使他们能够熟练地掌握统计信息的特性而编写的.

统计学在一般人的心目中,是最不“文科”的学科了.要是有人说统计学的好话,多半是在说它的用处.医疗领域的专家必须有统计学知识,才能了解医学研究报告的内容;经理人需要统计,因为大量的数字需要经过有效浓缩,才能看得出所以然;老百姓需要了解统计,才能知道民意调查和消费者物价指数到底是怎么回事.因为数据和机遇无所不在,所以我们的广告词可以这样写:每个人都用得到统计,甚至因统计而获利多多.

尽管统计有这么多的应用,但是我们很难说大家对于统计信息不仅接受而且有了较多的了解.当人们看到一个研究结果时,他们如何判断结论是否正确?他们是否会问:这个研究中的变量是如何定义的?用了什么样的统计方法?什么是“统计显著”的结果?所报告的结果有什么样的不足?这些问题正是我们在本书中讨论的一部分内容.显然,理解了统计学的主要概念以后,大家才能够明白那些专门鼓捣数字的人们都干了些什么,并对他们的结果进行评价.

本书在编写过程中得到了北京工商大学嘉华学院基础部、科研处等单位领导的大力支持和帮助,在此表示衷心的感谢!同时本书的出版得到了北京市民办教育发展促进项目资金资助及知识产权出版社编辑的大力帮助,也一并致谢!

为编好此书,我们尽了最大的努力,但我们深知自己的理论水平和实践经验非常有限,书中难免有不妥之处,恳请广大读者、学生和同行专家给予批评指正,我们将虚心接受,并且不断提高自身的水平.

编者

2014年10月

# 目 录

<b>第 1 章 商务统计学概述</b> .....	1
§ 1.1 引 言 .....	2
§ 1.2 什么是统计 .....	3
§ 1.3 描述统计学与推断统计学 .....	6
§ 1.4 标志、指标与指标体系 .....	10
§ 1.5 数据的类型 .....	14
习题 1 .....	17
<b>第 2 章 概率论基础</b> .....	21
§ 2.1 随机现象和随机事件 .....	22
§ 2.2 概 率 .....	26
§ 2.3 条件概率和事件独立性 .....	30
§ 2.4 随机变量和分布函数 .....	32
§ 2.5 离散型随机变量及分布律 .....	33
§ 2.6 连续型随机变量及概率密度 .....	35
§ 2.7 二维随机变量及分布 .....	40
§ 2.8 随机变量的独立性 .....	43
§ 2.9 随机变量的数字特征 .....	44
习题 2 .....	50
<b>第 3 章 用图表演示数据</b> .....	52
§ 3.1 引 言 .....	54
§ 3.2 频数分布表 .....	55
§ 3.3 茎叶图 .....	61
§ 3.4 直方图、频数折线图和频数累积图 .....	64
§ 3.5 其他的一些统计图 .....	67
习题 3 .....	75

<b>第 4 章 数值描述度量</b> .....	81
§ 4.1 引 言 .....	83
§ 4.2 集中趋势的度量 .....	83
§ 4.3 离散程度的度量 .....	96
§ 4.4 分布形态的度量 .....	104
§ 4.5 相对位置的度量 .....	107
§ 4.6 异常值检测 .....	112
习题 4 .....	113
<b>第 5 章 时间序列分析</b> .....	122
§ 5.1 时间序列及分析方法 .....	123
§ 5.2 时间序列的水平指标分析 .....	126
§ 5.3 时间序列的速度指标分析 .....	132
§ 5.4 时间序列的构成分析方法 .....	137
习题 5 .....	153
<b>第 6 章 统计指数分析法</b> .....	157
§ 6.1 指数的概念和计算 .....	158
§ 6.2 指数体系和因素分析 .....	162
习题 6 .....	167
<b>第 7 章 抽样分布</b> .....	170
§ 7.1 总体分布和样本分布 .....	171
§ 7.2 抽样分布 .....	174
习题 7 .....	182
<b>第 8 章 参数估计</b> .....	183
§ 8.1 点估计 .....	184
§ 8.2 区间估计 .....	186
习题 8 .....	191
<b>第 9 章 假设检验</b> .....	194
§ 9.1 假设检验中的基本概念 .....	195

---

§ 9.2 重要参数检验 .....	197
§ 9.3 假设检验与构造置信区间 .....	207
习题 9 .....	208
<b>第 10 章 相关与回归分析 .....</b>	<b>211</b>
§ 10.1 相关分析的对象和内容 .....	212
§ 10.2 简单相关分析 .....	215
§ 10.3 简单回归分析 .....	219
§ 10.4 相关系数 .....	225
习题 10 .....	228
<b>附 表 .....</b>	<b>231</b>
附表 1 泊松分布表 .....	231
附表 2 标准正态分布表 .....	237
附表 3 卡方分布表 .....	239
附表 4 $F$ 分布表 .....	247
附表 5 $t$ 分布临界值表 .....	259
<b>参考文献 .....</b>	<b>262</b>



## 第 1 章 商务统计学概述

### 学习目标和要求:

1. 了解什么是统计工作、统计资料和统计学;
2. 掌握总体、个体、样本等统计学中的一些基本概念;
3. 了解描述统计学和推断统计学的区别;
4. 掌握标志和标志表现的概念;
5. 掌握标志的分类;
6. 掌握总量指标的概念;
7. 了解总量指标的特点;
8. 掌握总量指标的种类;
9. 了解统计标志和统计指标的区别和联系;
10. 了解总量指标体系;
11. 能够确定数据的测量水平.

## 生活中的统计：目前什么专业就业率高



2013年6月,麦可思研究院发布《2013年中国大学生就业报告》.报告是基于麦可思对2012届大学生毕业半年后的调查研究及对2009届大学生毕业三年后的跟踪调查研究得出的.此报告中宣称,本科院校2012届大学生毕业半年后的就业率为91.5%,比2011届(90.8%)略有上升;2012届本科生毕业半年后就业率最高的学科门类是管理学(92.9%),最低的是法学(87.2%).2013年本科就业红牌警告专业包括动画、法学、生物技术、生物科学与工程、数学与应用数学、体育教育、生物工程、英语等;2013年本科就业绿牌发展专业包括地质工程、港口航道与海岸工程、船舶与海洋工程、石油工程、采矿工程、油气储运工程、矿物加工工程、过程装备与控制工程、水文与水资源工程、审计学.在就业的2012届本科生中,58%对就业现状满意,其中“211”院校毕业后的就业满意度为62%,非“211”本科院校为57%.在2012届本科学科门类中,毕业生毕业半年后就业满意度最高的为经济学,为61%;就业满意度最低的为农学,为53%;2012届本科生毕业半年后就业满意度最高的职业是“税收监察者、征收人和税收代理人”,为82%.最低的职业是“零售售货员”,为32%;2012届本科生毕业半年后就业满意度最高的行业是国有金融管理业,为76%;最低的行业为电子产品和电器用品零售业,为40%.

(资料来源:麦可思研究院发布的《2013年中国大学生就业报告》)

## § 1.1 引言

相信你对“统计”一词并不陌生,因为统计数据与统计术语已经融入了我们现在的生活、学习和工作.我们每个人都会自觉或不自觉地关心统计数据.比如:

- 2013年12月份新建商品住宅(不含保障性住房)环比价格变动中,最高涨

幅为1.1%，最低为下降1.7%；同比价格变动中，最高涨幅为21.9%，最低为下降2.8%。<sup>①</sup>

• 2014年1月上旬，全国猪肉平均价格为15.24元/斤，价格差为6.56元/斤。<sup>②</sup>

• 新发地农产品批发市场数据显示，2014年1月6日北京市蔬菜的加权平均价为每公斤2.63元，1月12日，北京市蔬菜加权平均价为每公斤2.62元。<sup>③</sup>

• 在2013年第15届青歌赛美声唱法总决赛中，王传越的自选歌曲《你是我心至爱的至爱》的评委打分分别为99.9, 99.8, 99.9, 99.9, 99.9, 99.9, 99.9, 99.9, 99.9, 99.9, 99.7, 99.4, 99.8, 99.8和99.8，去掉一个最高分99.9分，去掉一个最低分99.4分，王传越在第一轮的成绩是99.85分。

• 基差均值达到131元，基差幅度3.04%，基差幅度方差为1.75%，沥青期货市场存在期现套保机会。<sup>④</sup>

从上面这些例子可见，统计数据在现实生活中已经随处可见。而作为处理数据的方法之一——统计学，也被人们所青睐。特别是近几年，随着大数据的全面推广，统计学已经成为最炙手可热的学问，学习统计学将有助于你今后的工作中整理、分析、解读数据。比如，作为教师，你可能需要计算学生的平均成绩、及格率等；作为研究员，你可能需要知道新方法是否优于旧方法等；作为教练，你可能需要统计自己队员的各种得分和对手的各种得分等；作为企业管理人，你可能需要掌握生产销售情况、本企业产品的市场占有率等；作为政府领导，你可能关心居民消费价格指数、失业率等。这些只是统计学在不同行业中应用的几个简单的例子，而事实上，统计学的应用已经遍及各行各业。值得一提的是统计数据不是天然就有的，需要利用统计方法进行收集和整理才能得到。

## § 1.2 什么是统计

当人们谈及“统计”时，一般包括统计工作、统计资料和统计学。

### 1.2.1 什么是统计工作

人们对客观事物数量方面进行调查研究的认识活动，包括数据资料的收集、整

① 来源：<http://www.stats.gov.cn>

② 来源：<http://www.21food.cn/>

③ 来源：<http://sqjt.beijing.cn/>

④ 来源：<http://finance.sina.com.cn/>

理和分析工作,称为统计工作。

过去,只有各级政府部门设有专门的统计机构从事统计数据的收集、整理工作,然而随着信息社会的到来,人们发现身边存在着大量可使用的数据,而且这些数据往往具有重要的商业价值、经济价值和社会价值。典型的案例是一个关于尿不湿与啤酒的故事。超级商业零售连锁巨无霸沃尔玛公司拥有世界上最大的数据仓库系统之一。沃尔玛利用数据挖掘工具对这些数据进行分析 and 挖掘。一个令人惊奇和意外的结果出现了:“跟尿不湿一起购买最多的商品竟是啤酒!”产生这一现象的原因是:美国的太太们常叮嘱她们的丈夫不要忘了下班后为小孩买尿不湿,而丈夫们在买尿不湿后又随手带回了他们喜欢的啤酒。另一种情况是丈夫们在买啤酒时突然记起他们的责任,又去买了尿不湿。既然尿不湿与啤酒一起被购买的机会很大,那么沃尔玛就在他们所有的门店里将尿不湿与啤酒并排摆放在一起,结果是得到了尿不湿与啤酒的销售量双双增长。

因此,统计工作不再只是政府部门的事情。越来越多的人将会投入到统计工作中来。其职能不仅仅限于填报统计报表,更重要的职能在于对统计数据进行分析,发现事物发展变化的规律。

### 1.2.2 什么是统计资料

统计工作的成果,包括各种统计数据、图表和文字资料等,统称为统计资料。例如国内生产总值数据,说明整个国家的生产规模;居民消费支出,说明居民的生活水平。而这些数据经常会在报纸、杂志、电视、广播、网站、书籍等出现。

这些间接得到的数据都称为二手数据。二手数据往往比较容易得到,相对来说也比较便宜。但使用时应当谨慎,因为二手数据有它的局限性,比如时效性差、可靠性低等。比如一个最重要的二手数据的来源《人口普查资料》,这项普查每10年一次,但因为数据如此之多,普通读者要看到有关数据还必须等待大约4年的时间,而此时,有些数据已经过时了。

二手数据一般为初期研究提供了很大的帮助,然而要深入研究还是应尽可能地获得一手数据,即原始数据。原始数据更及时、更准确、更有效,但一手数据并不像得到二手数据那么轻松。因为获得最原始的数据往往成本是比较高的,比较费时的,代价是比较大的。比如,想要了解非洲狮的生活习性,调研人员首先需要大量的时间和金钱,同时还要面临着各种生命危险。

### 1.2.3 什么是统计学

统计学是处理数据的一门科学。一些人认为统计学是数学的一个分支,因为统



计学与概率论联系紧密;而另一些人认为统计学不是数学的一个分支,因为很多统计学的内容都不是数学,比如数据的收集、试验的设计等.从人们的争论中可以看出,现在的统计学应该归属为一门交叉科学,其技术方法是建立在数学、计算机学、经济学等多个学科基础之上的.需要注意的是要想学好统计学,一定的数学基础是不可或缺的.那么什么是统计学?从古至今,统计学家们给统计学下了诸多的定义,但目前广泛使用的定义如下:

统计学(statistics)是收集、处理、分析、解释数据并从数据中得出结论的科学.

统计学的定义告诉我们,统计学是用来研究数据的,数据可以来自各行各业.统计学是提供一套系统的、完整的收集数据、处理数据、分据数据、解释数据并得出结论的方法.

#### 1.2.4 为什么要学习统计学?

首先,看看一些统计学家是怎么评价统计学的.

(1)著名统计学家 C. R. Rao 有一句名言:在终极的分析中,一切知识都是历史;在抽象的意义下,一切科学都是数学;在理性的世界里,所有的判断都是统计学.

(2)著名的英国护士和统计学家弗洛伦斯·南丁格尔(Florence Nightingale, 1820~1910)有一句名言:如果要想了解上帝在想什么,我们就必须学统计,因为统计学就是在测量上帝的旨意.这里所说的上帝是我们的客观世界.

(3)著名的统计学家约翰·图基(John Tukey, 1915~2000)有一句名言:作为统计学家最美妙的事就是你能有跨界的乐趣.

(4)哥伦比亚大学统计学家兼政治科学家安德鲁·格尔曼称:“文化已经发生了改变.现在人们的想法是,数字和统计学是有趣的,是一种很酷的东西.”

其次,看看其他的关于统计学的一些名人名言、评论与报道.

(1)原北大校长马寅初曾说过:“学者不能离开统计而研究,政治家不能离开统计而施政,事业家不能离开统计而执业.”

(2)美国著名科幻作家威尔斯(H. G. Wells)曾说过:“对于追求效率的公民而言,统计思维总有一天会和读写能力一样必要.”

(3)斯大林曾经说:“任何建设工作,任何国家工作,任何计划工作,没有正确的计算是不可想象的.而没有统计,计算是不可想象的.”

(4)国务院《关于加强统计工作的决定》(1984年1月6日)中提到:“统计是社会主义建设的一项重要基础工作.我国要实现工业、农业、科学技术和国防现代化,必须实现统计工作的现代化.”

(5)在间谍中有一句名言:世界上没有破译不了的密码,只是时间问题,因为有了统计学.统计学、代数学为我们破译密码提供了良好的工具.

虽然读者今后从事的工作可能不是统计工作,但现在学习统计学仍然具有重要的意义,可以归纳为如下3个主要的原因:

(1)你今后的学习可能需要用到统计学.统计学是处理数据的一门基础学科.如果你对数据感兴趣,喜欢研究数据,那么今后你要学习的计量经济学、数据挖掘等课程都离不开统计学,因此学好统计学将有助于你后续一些课程的学习.

(2)你今后的工作可能需要用到统计学.虽然你今后可能不是统计工作者,但你的工作仍然有可能存在大量的数据需要处理,而统计学将提供你处理数据的一种方法.即使你不用统计学的方法,但你的对手或是合作伙伴可能会使用统计学的方法,解读数据也是你在统计学课程中将要学到的.

(3)学习一点统计学可以帮助你生活中更好地做决策.比如据统计,因肺癌死亡的患者中,87%是由吸烟包括被动吸烟引起的,男性吸烟者肺癌死亡率是不吸烟者的8到20倍.看到这样的统计信息,你还会考虑吸烟吗?再比如统计研究表明青少年的鞋号与阅读能力之间相关性很高,那么你是否会认为青少年阅读能力越强,鞋号就应该越大呢?最后的一个例子,电视上某去痘广告声称在大街上随便抽取脸上有痘的四个人,实验表明四个人去痘效果都非常明显,现在又有很多的优惠,如果你的脸上也有痘,你会冲动地去买这款产品吗?如果你学点统计学,就可以利用一些信息帮助你和别人更好地生活,同时也不会被某些信息所迷惑而得出错误的结论.

### § 1.3 描述统计学与推断统计学

统计学经过300多年的发展,形成了自己的学科体系.描述统计学(descriptive statistics)和推断统计学(inferential statistics)是初等统计学重要的组成部分.学好描述统计与推断统计对掌握诸多统计技术至关重要.在给出描述统计学和推断统计学的概念之前,需要先了解统计学里几个常用的统计术语.

总体(population)是我们所要研究的所有基本单位(通常是人或其他)的总和.

也就是说,总体是一同质群体.例如,总体可以是:①中国的全部人口;②一个班级的所有学生;③10个公司的所有员工;④大海里的鱼类;⑤某厂某日生产的所有汽车;⑥去年北京市肯德基所有餐馆的销售量;⑦北京某一地点每一天(包括过去、现在和将来)的最高气温,等等.

①至③这三个总体是人的总和;④这一总体是动物的总体;⑤这一总体是物体

的总体;⑥这一总体是交易的总和;⑦这一总体是事件的总和。

组成总体的每一个事物,称为总体单位,简称单位或个体。

例如,生活在中国的每一个人都是①的个体;班里的每一个学生都是②的个体;10个公司中的每一个员工都是③的个体;海洋里的每一种鱼类都是④的个体;某厂某日生产的每一辆汽车都是⑤的个体;去年北京市肯德基每一个餐馆的销售量都是⑥的个体;北京该地点每天的最高气温都是⑦的个体。

总体中所包含的个体的个数称为总体容量。

总体可以分为有限总体和无限总体。容量有限的称为有限总体,容量无限的称为无限总体。①,②,③,⑤,⑥都是有限总体,④,⑦是无限总体。

总体具备三个特性:同质性、变异性和大量性。同质性是指构成总体的全部个体,至少在某一点上或某一方面必须是相同的。如妇女总体是性别上的同质;职工总体是就业身份上的同质,等等。同质性是总体的根本特征,只有总体单位是同质的,才能通过对个体特征的观察研究,统计归纳和揭示出总体的综合特征和规律性。然而同质性又并不意味着所有总体单位在一切方面都是同质,它们在其他方面往往是不同的,具有差异的。如妇女们在年龄、职业等方面是有差异的;职工们在工龄、工资等方面是不同的。这就是总体的变异性。总体的变异性是统计研究的前提。大量性是指总体所包括的总体单位要有足够多的数量,个别或很少几个单位不能构成总体。如果总体单位数过少,则将丧失统计观察和研究的意义,而总体的大量性,可使总体单位某些偶然因素的影响相互抵消,从而显示出总体的本质和规律性。

总体和个体是相对而言的,二者的确定均决定于统计研究的目的。同一事物在不同目的的研究中,既可以作为总体,也可以作为个体。例如,当我们要了解全国工业企业职工的工资收入情况时,那么全国工业企业为总体,各个工业企业为个体;但当我们了解某个工业企业职工的工资收入情况时,则该企业就成了总体,而每位职工的工资就是个体。

考虑到获得总体的成本、时间、局限性等各种因素,大多数时候想要得到总体是不太可能的。因此,在现实的生活中,只能抽取一部分个体来进行研究,将其定义为样本。

样本(sample)是总体中所抽取的一部分个体。

**例 1.3.1** 某学院 2013 年秋季学期统计学期末考试的试卷共有 1430 份,为了检验试卷核分的准确性,随机地从 1430 份中抽取 100 份进行检查,那么 1430 份试卷就构成了一个总体,抽出的 100 份试卷就构成了一个样本。

**例 1.3.2** 想要监测北京市内空气污染的情况,我们不可能得到每一块土地

上方的空气质量指标,只能在北京市内建立一些监测站,获得监测站上方的空气质量数据,而北京市内每一块土地上方的空气质量指标就构成了一个总体,而监测站所获得的空气质量指标就构成了一个样本。

如果样本选择的合适,在大多数的时候它们将拥有与总体相同或相似的特征。正因为两者之间存在着这种亲密的关系,利用样本来推断总体才成为可能。

收集到的数据,不管它是样本还是总体,往往是杂乱无章的。想要找到其中的规律,需要一定的方法和技术。初等统计学把处理数据的方法和技术归属为两类,即描述统计学与推断统计学。

描述统计学就是以某种信息化的方式收集、组织、概括和展示数据的方法。

具体来说,描述统计就是研究如何取得客观现象的数据,并通过图表形式对所收集的数据进行加工处理和显示,进而通过综合、概括与分析得出反映客观现象规律性的数量特征。内容包括统计数据的收集方法、数据的加工处理方法、数据的显示方法、数据分布特征的概括与分析方法等。

例如,国务院定于2010年进行第六次人口普查,普查员对每家每户进行调查,内容包括性别、年龄、民族、受教育程度、婚姻生育等。收集回来的数据需要进行录入、编辑、审核和汇总。最后,汇总出来的结果会用各种形式进行发布,比如:

### 1. 数字描述

这次人口普查登记的全国总人口为1 339 724 852人,与2000年第五次全国人口普查相比,十年增加7390万人,增长5.84%,年平均增长0.57%,比1990年到2000年的年平均增长率1.07%下降0.5个百分点。数据表明,十年来我国人口增长处于低生育水平阶段。

### 2. 表格描述

如表1-1所示。

表 1-1

指标	1953年 (一普)	1964年 (二普)	1982年 (三普)	1990年 (四普)	2000年 (五普)	2010年 (六普)
普查时点	7月1日零时	7月1日零时	7月1日零时	7月1日零时	11月1日零时	11月1日零时
总人口(人)	594 346 737	694 581 759	1 008 175 288	1 133 682 501	1 265 825 048	1 339 724 852
新增人口(人)		100 235 022	313 593 529	125 507 213	132 142 547	73 899 804
增长(%)		16.86	45.15	12.45	11.66	5.84
年平均增长率(%)		1.43	2.09	1.48	1.07	0.57



## 3. 茎叶图(图 1-1)描述和条形图(图 1-2)描述

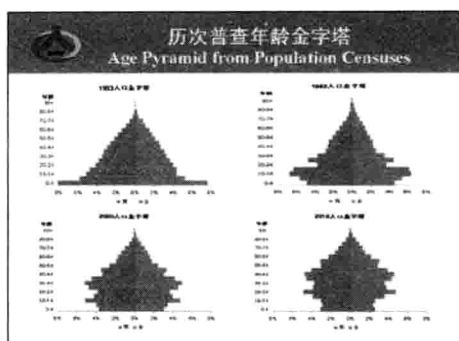


图1-1 茎叶图

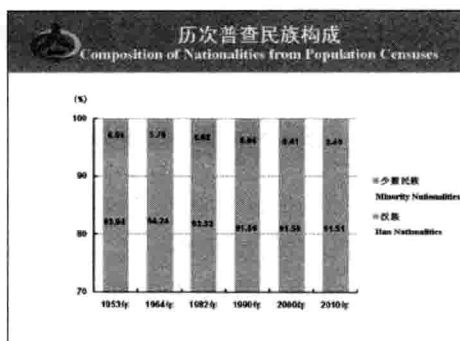


图1-2 条形图

## 4. 饼图(图 1-3)描述和时间序列图(图 1-4)

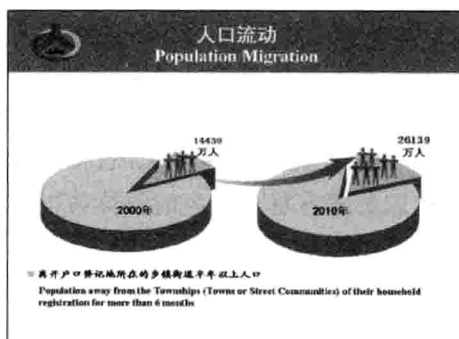


图1-3 饼图



图1-4 时间序列图

从上述的例子可以看出,描述统计学主要是在已有的数据上进行分析、总结和展示. 如果我们需要预测未来,需要推断具有相似性质个体的特性,那么需要另一种方法,我们一般称为推断统计学.

推断统计学是在样本的基础上确定总体的某些特征的方法.

推断统计其实就是利用样本数据的信息,对总体的情况作出估计、推断、预测或其他归纳. 包括从样本推广到总体,进行估计和假设检验,确定变量之间的关系,进行预测等. 推断统计要用到大量概率论里的内容与方法.

例如,美国盖洛普(Gallup)调查公司在美国总统大选前通常会从全美国的选民中随机抽取 1500 人左右,对大选结果进行调查和预测,并会给出 2% 左右的预测误差.