



全国高等农林院校“十一五”规划教材

统计学原理

TONGJIXUE YUANLI

江 华◎主编



 中国农业出版社

欢迎登录：全国农业教育教材网
<http://www.nongyejiaocai.com>

本书采用出版物数码防伪系统
刮开涂层将16位防伪密码发短信至95881280
免费查询 辨别真伪
详情请查询中国扫黄打非网
<http://www.shdf.gov.cn>
防伪、网络增值服务说明见书内“郑重声明”页

明码 4102 2416 8793 6971
密码 XXXXXXXXXX

封面设计 姜欣

ISBN 978-7-109-11961-1



9 787109 119611 >

定价：25.30 元

全国高等农林院校“十一五”规划教材

统计学原理

江 华 主编

中国农业出版社

图书在版编目(CIP)数据

统计学原理/江华主编. —北京: 中国农业出版社, 2007. 11

全国高等农林院校“十一五”规划教材

ISBN 978-7-109-11961-1

I. 统… II. 江… III. 统计学—高等学校—教材 IV. C8

中国版本图书馆 CIP 数据核字 (2007) 第 168414 号

中国农业出版社出版

(北京市朝阳区农展馆北路 2 号)

(邮政编码 100026)

责任编辑 夏之翠

北京智力达印刷有限公司印刷 新华书店北京发行所发行

2008 年 1 月第 1 版 2008 年 1 月北京第 1 次印刷

开本: 720mm×960mm 1/16 印张: 18.25

字数: 318 千字

定价: 25.30 元

(凡本版图书出现印刷、装订错误, 请向出版社发行部调换)

主 编 江 华 (华南农业大学)

副主编 徐学荣 (福建农林大学)

平 瑛 (上海水产大学)

范文正 (华南农业大学)

参 编 (按编写章节先后顺序排序)

陈利昌 (华南农业大学)

胡瑞卿 (惠州学院)

张士海 (广东海洋大学)

祝建民 (仲恺农业技术学院)

李明月 (华南农业大学)

欧阳涛 (湖南农业大学)

谢小芸 (湖南农业大学)

序 言

“统计学原理”是高等院校经济管理类专业必须开设的专业基础课程，它是从事经济分析与研究的重要工具。本学科通过研究社会经济现象的数量表现及其相互关系变动特征去揭示社会经济发展运行的规律，从而为各级经济预测与决策提供必要的基础依据。

全国高等农林院校“十一五”规划教材《统计学原理》的编写目的就是为适应在进入 21 世纪新形势下各高等农林院校经济管理类专业开设统计学原理课程所用。本教材的特点是：一方面站在学科前沿，力求反映目前统计学理论和方法论的新成果，增强内容的时效性，特别是突出了数理分析在统计学原理中的应用。另一方面，以全面而丰富的新的内容及知识点和严谨的逻辑结构确保全书的系统性与完整性。

本教材于 2006 年春经中国农业出版社审定立项后，主编和副主编共同商议形成写作提纲，经编写组成员近一年的努力写出本教材初稿。2007 年 2 月初，在本教材的主编单位华南农业大学召开了教材定稿会。会上各位参编者对每章内容均提出了详细的修改和充实意见。在此基础上，各位参编者根据审稿会意见进行进一步修改后成稿。

本教材的主要数据资料和案例一部分来自于《中国统计年鉴》和各省、市统计部门，另一部分来自本教材参编者的实际调查研究，这就更加突出了本教材的工具性和实用性。为了强化对学生的基本理论、基本知识和技能的训练，本教材十分注重理论联系实际。在

各章节中每讨论一种方法和公式后均给出了实际运用的案例，并用实证案例来说明有关的基本概念和方法。本教材在每章开始时均有本章重点提示，在每章结束时还专门进行了本章小结，并附有复习思考题和参考答案，以帮助学生巩固与掌握各章的主要知识点。另外，在本教材最后还附有近年来我国颁布的统计法规和趣味统计学事例以及 Excel 统计函数等，其目的是增强学生对统计学应用范围的了解和提高学生学习统计学原理的兴趣。

本教材由江华主编，徐学荣、平瑛、范文正副主编。全书共分 10 章，参加编写的人员有：范文正（第一章、附录），陈利昌（第二章），胡瑞卿（第三章），张士海（第四章），祝建民（第五章），徐学荣（第六章），平瑛（第七章），李明月（第八章），欧阳涛、谢小芸（第九章），江华（序言、第十章），最后由江华统一修改和定稿。

由于时间仓促和水平所限，书中难免有欠缺之处，敬请读者给予批评指正。

编 者

2007 年 5 月

目 录

序言

第一章 绪论	1
一、统计学的特点	2
二、统计学的作用	6
三、统计学的源流	7
本章小结	14
复习思考题	14
第二章 统计数据的收集	15
第一节 统计数据的类型	15
一、统计数据的计量尺度	15
二、观测数据和实验数据	16
三、变量的准确定义与变量选择	16
第二节 统计数据的收集方法	17
一、统计调查方法	17
二、统计调查方式	18
第三节 统计调查问卷设计	21
一、问卷的结构	21
二、问卷提出问题的形式和设计程序	23
三、问卷设计的基本要求	24
四、问卷范本举例	25
第四节 统计数据的质量	26
一、统计数据的误差	26
二、统计数据的质量要求	27
三、数据质量的相对性及影响因素	27
四、数据质量的检查和控制	28
本章小结	28

复习思考题	29
练习题	29
第三章 统计数据的整理与显示	31
第一节 数据的审核与编录	31
一、数据的审核	31
二、数据的编码	34
三、数据的录入	39
第二节 统计数据的整理	39
一、数据整理的概念与意义	39
二、数据整理的原则	40
三、数据汇总的组织和技术	40
四、数据分组	41
五、数据的分布数列	43
第三节 统计数据的显示	48
一、图示法	48
二、次数分布的主要类型	54
三、绘制统计图应注意的问题	55
四、表示法	55
本章小结	56
复习思考题	57
练习题	57
第四章 统计数据的描述	61
第一节 统计分布的集中趋势	61
一、算术平均数	61
二、交替标志平均数	66
三、几何平均数	67
四、众数	69
五、中位数与四分位数	71
第二节 统计分布的离散趋势	74
一、标准差和方差	74
二、标准差系数	75
三、交替标志的标准差	77

四、四分位差	77
五、方差的加法定理	78
第三节 统计分布的形态	81
一、偏度	81
二、峰度	83
本章小结	84
复习思考题	85
练习题	85
第五章 抽样与参数估计	87
第一节 抽样推断概述	87
一、抽样推断的含义	87
二、抽样推断的理论基础	88
三、概率抽样的主要方法	93
第二节 抽样分布	96
一、正态分布	96
二、 χ^2 分布	100
三、 t 分布	101
四、 F 分布	104
第三节 参数估计的基本方法	105
一、点估计与区间估计	105
二、总体均值的区间估计	107
三、总体比例的区间估计	111
四、两总体均值差的区间估计	113
第四节 样本容量的确定	118
一、在重置抽样条件下样本容量的确定	118
二、在不重置抽样条件下样本容量的确定	119
本章小结	120
复习思考题	120
第六章 假设检验	123
第一节 假设检验的过程和逻辑	123
一、假设检验的基本思想	123
二、假设检验的步骤	125

三、假设检验中的两类错误	127
第二节 均值的检验	128
一、单样本总体均值检验	128
二、两样本总体均值差检验	133
第三节 比例的检验	135
一、单样本总体比例检验	135
二、两样本总体比例差检验	136
第四节 方差的检验	137
一、单样本总体方差检验	137
二、两样本总体方差检验	139
本章小结	141
复习思考题	141
第七章 相关与回归	143
第一节 现象之间的数量关系	143
一、相关关系的概念	143
二、相关关系的种类	144
三、相关图(散点图)	145
第二节 相关分析	146
一、相关系数	146
二、相关系数的计算	147
三、相关系数的显著性检验	149
第三节 一元线性回归分析	149
一、一元线性回归模型	150
二、回归模型参数的估计	151
三、线性回归方程拟合优度的测定	152
四、线性回归方程的显著性检验	155
第四节 多元线性回归分析	157
一、二元线性回归方程的拟合	159
二、判定系数、估计标准误差	159
第五节 可转化为线性模型的曲线回归分析	162
一、指数函数	163
二、幂函数	163
三、双曲线函数	164

四、对数函数	165
本章小结	165
复习思考题	165
第八章 时间序列	168
第一节 时间序列概述	168
一、时间序列的概念	168
二、时间序列的种类	168
三、时间序列的编制原则	170
第二节 时间序列的动态指标分析	171
一、时间序列的水平指标分析	171
二、时间序列的速度指标分析	175
第三节 时间序列成分分析	179
一、时间序列构成与分解	179
二、长期趋势分析	181
三、季节变动分析	187
四、周期变动分析	191
本章小结	191
复习思考题	192
第九章 统计指数	194
第一节 指数概述	194
一、指数的概念及特点	194
二、指数的种类	195
三、指数的作用	196
第二节 总指数的编制原理	197
一、总指数的基本原理	197
二、综合指数的编制原理	197
三、平均指数的编制原理	203
第三节 指数体系与因素分析	207
一、指数体系的意义	207
二、总量变动的因素分析	209
三、平均指标变动的因素分析	212
第四节 常用经济指数	215

一、居民消费价格指数 (CPI)	215
二、工业生产指数	218
三、股票价格指数	219
四、农副产品收购价格指数	221
本章小结	221
复习思考题	222
练习题	222
第十章 统计决策	225
第一节 统计决策的意义、种类和原则	225
一、统计决策的概念和种类	225
二、统计决策的程序与原则	226
第二节 确定型统计决策方法	228
一、比较择优法	229
二、微分法	229
三、线性规划法	229
四、目标规划方法	229
第三节 风险型决策方法	232
一、风险型决策方法的特点及表述方式	232
二、风险型决策的种类与分析过程	234
三、关于试验与否的决策问题	242
第四节 统计决策例析	244
一、几种不同决策准则的分析过程	244
二、最大损益期望值及最小方差决策案例分析	246
三、利用后验信息 (概率) 决策分析	248
本章小结	250
复习思考题	250
附录	252
主要参考文献	274

第一章 绪 论

重点提示：统计学是处理统计数据的通用方法论学科。通过阅读本章内容，可以了解：①统计学的特点；②统计学的作用；③统计学的历史发展脉络。

在介绍统计学知识之前，先请读者回忆一些经常接触到的反映现实情况的数字，如人口数、产品销售量、品牌知名度、天气预报等，它们通过对某些现象数量特征的描述使我们了解了一些事物，我们可以把这些数量信息称为统计数据。再请读者回忆一些“规律”，如“吸烟有害健康”、“着凉后容易感冒”、“恩格尔系数（食品支出占家庭总支出的比重）低于25%的家庭生活很富裕”，它们描述了一些“关系”——我们常把这样比较稳定的“关系”视为“规律”，这些“规律”是存在的，但并不绝对可靠。请读者从这两类现象中总结一些你认为是共同的东西。

我们每天都要接触许多数字或数量化的现象，有些是日常生活中反映“小”问题的数字，如去某处要花多少时间、今天会不会下雨，有些是反映一些“大”现象的数字，如股票指数、失业率，这些现象都与统计学有一定关系。人们从对数量的感知、计数开始，发展到现在的数学、统计学等学科，都是对数量关系、数量规律的认识与探索。

统计学是一门方法论或工具性学科，从方法上看，它是收集、整理、描述、显示和分析统计数据的科学；从功能上看，它是对大量个别现象的数量特征进行概括判断的科学。

许多人简单地认为统计就是搜集数字，加加减减，形成诸如“旅游人数”、“客运量”之类的统计指标，其实这只是较规范的计数活动，是千百年来人类计数活动的延续与扩展。统计学无论在方法和理论上还是应用范围上，已远远超出了这个范围，它已经发展成为广泛应用于自然科学、政治法律、经营管理、社会生活等广阔领域的科学而有效的分析方法。在最近的100年内，由于统计学有了很大的进步，又由于统计学的广泛运用，一定程度上使物理学、生物学、医学、经济学、管理学、人类学、气象学、军事科学、语言学等许多学科都获得了实质性的进步。在色彩斑斓的现代社会，统计学是不可或缺的思维

与分析利器。

本书介绍统计学的基本方法和内容,阐述一些基本统计思想。多数统计方法看起来很平易,只涉及加减乘除、开方等简单运算,但其思想方法,不仅与数学不同,且在一定程度上与一些日常生活养成的思考方式也不同。掌握这些统计思想和统计方法有助于我们在看待事物时具有“大局观”,在叙述时能正确地使用统计数字,正确地理解和分析别人所作的统计表述,或者帮助自己阐明观点;同时,统计方法的合理运用,能使我们区分“直觉”和客观的见识。也许,由于统计学自身的特点,它比其他课程更宝贵的是,统计学能帮助我们用较明确而可靠的依据评价自己的直觉印象和简单判断。

一、统计学的特点

(一) 统计学关注于事物的数量特征

人们对世界的认识是逐步加深和拓宽的,认识逐渐丰富的过程,实际上是对“差异”(可以是属性上的差异,也可以是数量上的差异)了解更多的过程。

人们在观察自然现象时,自觉或不自觉地发现了事物之间的差异,有些是性质上的明显不同,如一棵树和一条蛇,一座山和一条河;有些相同事物具有数量上的差异,如一棵小树和一棵大树。不同的物种表现质的差异,不同的大小表现量的差异,这些都需要人们用不同的方法去表达、标记它们,也就是说,不同的事物有不同的性质和表现,为了反映这种表现上的差异性,人们找到了许多方法,其中把同类事物之间的差异用不同的数量来表现是普遍而有效的方法,这就是量化的方法。

质与量是事物的两种基本规定性。质具有唯一性,一物一质是事物存在的根本格局。量是质的表达方式,量是质的等级、规模、范围和结构的表现,是事物可以由数和形来表示的规定性。质的变化其实就是事物通过自身的数量关系和数量变化来重新表达质,生成不同于以前的新质。量变是时刻不停的,我们总是置身于无止无休的量变之中。

统计数据是使我们形成“知道”的重要手段,能使我们深入地研究事物,发现事物本身的内在规定性。比如,通过大量观察,人们发现人口男、女性别比例通常是 $103 \sim 107 : 100$,这反映了一个类别现象的根本数量规律。再如,农学家通过对农作物施肥量与产量的对比研究,探索到施肥量与产量之间的数量关系,反映了两个事物之间数量上的相互影响关系,从而揭示此二事物之间的本质联系。另如,疾病的发生受多种因素影响,医生可以使用统计方法分析许多事物之间相互影响的情况,进而探索某些生物学机理和治疗手段。

统计学观察的对象是事物的数量特征。统计学运用科学的方法搜集、整理那些反映客观事物数量特征的数据,并使用符合数学逻辑的方法分析事物的数量规律。

(二) 统计学的观察对象具有随机性

偶然现象存在于时时处处,绝对的精确只存在于理论之中。“机会”、“运气”、“不确定性”等词汇都是偶然现象的近似表达。统计学把形形色色的偶然性统称为随机性,即“随机会而定”的意思,从技术角度出发,把所有可以取不同数值的变量称为“随机变量”。例如,要知道一件物体的重量,把它放在衡器上去称,若衡器是没有误差的,则称一次就知道其确切重量。若衡器虽然有误差,但确切知道误差是多少(即误差没有随机性),则称一次也就解决问题。惟其有误差,且误差是随机而不可预测的,就有统计推断问题存在。此时,为了更准确地测量重量,通常的做法是把这物件重复称几次,以其结果的平均值去“估计”物重。这“估计”就是一种统计方法。多次平均的结果仍可能有误差,但误差缩小了,缩小多少?这都是统计问题。又如想了解全国农户的平均货币收入,若把每个农户的收入都调查清楚,再把结果加以平均,则得到确切的答案。在理想化条件下这可以归结为简单的算术问题,可现实中的统计问题却要面对以抽样来估计整体的限制条件。如果我们只抽出一部分农户来调查,就会有误差,这误差的产生是由于哪一部分农户被抽出是凭机会的,因此这误差也有随机性。常用统计方法多有“由部分推断整体”的性质。抽查农户收入是推断问题,测量重量也是推断问题。

至于随机性之所以广泛存在,可以从哲学角度理解。事物的互相联系是普遍存在的,各种事物又有各自的运行规律。一件事物的变化要受到许多现象的影响,现实当中,我们无法穷尽地知道影响一个事物变化的所有因素及这些因素的具体状态,当然也就无法绝对精确地判断这个事物。即使是在理想条件下观察物体,因为观察的瞬时性不能概括具有多种状态的事物,我们仍然难以真正地把握这个事物。

(三) 统计方法的逻辑过程

所有研究过程的基本逻辑是:始于材料,用既有理论和方法进行分析,最终得到结论。在所有这些阶段,都要使用统计方法,各阶段又有一定的特点。使用统计方法,首先要通过大量观察,了解事物的数量表现,获取有关基础数据,这是对事物形成认识的重要基础。其次是有条理地简化数据。我们无法通过庞杂无绪的原始数据形成条理化的认识,必须对原始数据进行整理、加工,

形成抽象、明晰、有代表性的数量特征值。随后是分析与归纳，这是统计活动的目的所在——寻找客观事物内在的确定性，使我们能够触类旁通，执简驭繁。

简单地讲，统计方法的逻辑过程就是：大量观察→分组划类→归纳推断。

为什么首先要进行大量观察呢？这是因为：

1. 通过大量观察，可以了解许多个别事物的共同特征 客观事物都是必然性与偶然性的对立统一，必然性通过偶然性表现出来。通过大量观察，我们会发现，大多数个别现象都是围绕着某个中心左右波动，随着时间的流逝或观察范围的扩大，这种波动能体现出较稳定的趋势，这种趋势就经常被我们视为规律性或必然性。比如，美国的“人寿保险之父”伊莱泽·赖特在 19 世纪中叶说：“没有任何东西比一个人寿命的长短更难确定，然而，也没有任何东西比一千个人的平均寿命更为肯定”。概率论的大数定律从数量方面刻画了大量观察的效能。

2. 大量观察也是社会发展的要求 人们视野之所及，要求相应地有认识该范围的手段。人类早期的计数活动只能局限于比较小的范围，随着社会的发展，交流日益密切，全社会的系统化程度逐渐提高，对更大范围内事物的充分认识是社会活动范围扩大的必要条件，2 600 多年前中国春秋战国时期的齐相管仲就有“不明于计数，而欲举大事，犹无舟楫而欲经于水险也”的思想。现代社会是要能“用数目字管理的社会”（历史学家黄仁宇语），各国政府、各种机构都在获取数字信息上花费了很大的财力，不进行大量观察是无法得到较全面的统计信息的。

分组划类的必要性体现在：人们的认识能力和认识模式不足以平行地、一次性地深入认识大量信息，或者，事物的复杂程度相对于我们的认识能力而言是无穷大的，有效的认识方法只能是简约化，对大量信息进行抽象和归类，形成金字塔式的有序信息结构之后，才能把握其较主要、较根本的特征。比如生物学理论认为“一切生物体都是由细胞组成的”，但这不足以认识生物的多样性，因此就把生物类别按等级划分为“界门纲目科属种”，按照某种共性划分为不同级别的类型，使阵营庞大、姿态万千的生命世界中的生物各有归属，每一个等级所属的生物都有相同的性质。对同类事物还可以按照某种数量特征进行分组划类，以便更清晰、深入地认识其数量表现。

由于事物的相似性使对事物的分组划类成为可能，即事物之间所存在的相似性，是对众多事物进行分组划类的客观基础。为了最经济地使用大脑资源，我们总习惯于把某些差异较小的性质定义为“相似”，再把若干具有相似性的事物划为同一类型，由其中的典型事例来代表整个类别，而不是记住所有事物的性质。相似性因为抽象和概括而具有代表性。按照不同的差异级别，我们就