

“十三五”国家重点出版物出版规划项目

CENGAGE  
Learning

经济科学译丛

# 统计学

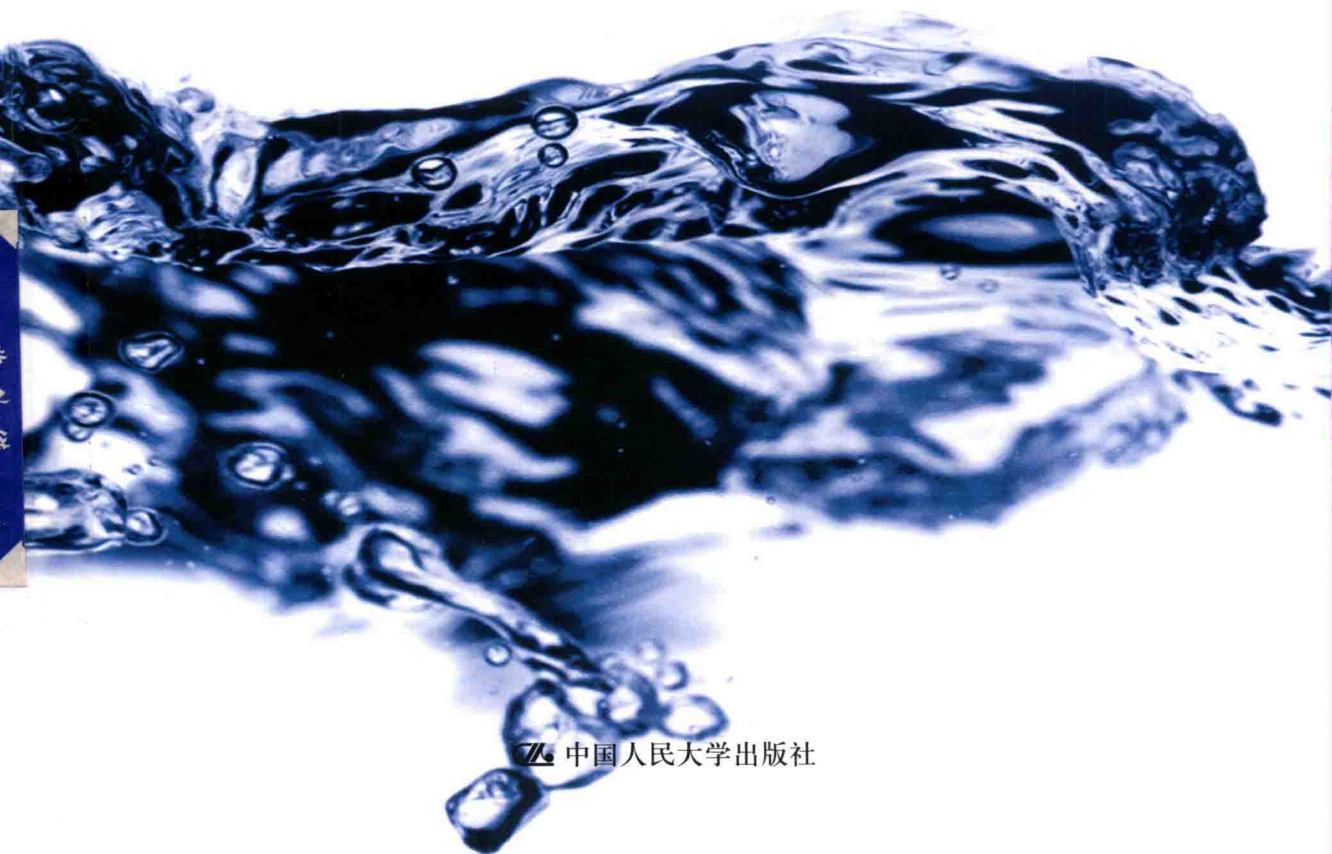
在经济和管理中的应用 (第10版)

下册

杰拉德·凯勒 (Gerald Keller) 著

Statistics for Management  
and Economics

(Tenth Edition)



中国人民大学出版社

“十三五”国家重点出版物出版规划项目

经济科学译丛

# 统计学

在经济和管理中的应用（第10版）

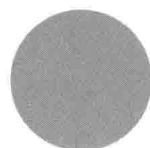
下册

杰拉德·凯勒（Gerald Keller）著

夏利宇 韩松涛 李君 张伟等译

Statistics for Management  
and Economics

(Tenth Edition)



中国人民大学出版社  
·北京·

## 图书在版编目 (CIP) 数据

统计学：在经济和管理中的应用：第 10 版/（ ）杰拉德·凯勒著；夏利宇等译．—北京：中国人民大学出版社，2019.4

（经济科学译丛）

ISBN 978-7-300-26771-5

I. ①统… II. ①杰… ②夏… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字 (2019) 第 032483 号

“十三五”国家重点出版物出版规划项目

经济科学译丛

**统计学：在经济和管理中的应用（第 10 版）**

杰拉德·凯勒 著

夏利宇 韩松涛 李君 张伟 等译

Tongjixue: Zai Jingji he Guanli zhong de Yingyong

---

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）

010-62511770（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 涿州市星河印刷有限公司

规 格 185 mm×260 mm 16 开本

版 次 2019 年 4 月第 1 版

印 张 67 插页 4

印 次 2019 年 4 月第 1 次印刷

字 数 1 614 000

定 价 158.00 元（上下册）

---

版权所有 侵权必究

印装差错 负责调换

# 《经济科学译丛》

## 编辑委员会

### 学术顾问

高鸿业 王传纶 胡代光 范家骧 朱绍文 吴易风

### 主 编

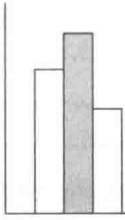
陈岱孙

### 副主编

梁 晶 海 闻

### 编 委 (按姓氏笔画排序)

王一江 王利民 王逸舟 贝多广 平新乔 白重恩  
朱 玲 刘 伟 许成钢 李 扬 李晓西 李稻葵  
杨小凯 汪丁丁 张宇燕 张维迎 林毅夫 易 纲  
金 碚 姚开建 钱颖一 徐 宽 高培勇 盛 洪  
梁小民 樊 纲

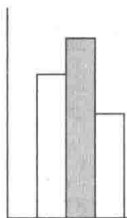


# 目 录

<b>第 14 章 方差分析</b> .....	551
14.1 单因素方差分析 .....	552
14.2 多重比较 .....	572
14.3 方差分析试验设计 .....	583
14.4 随机化区组（双因素）方差分析 .....	585
14.5 双因素方差分析 .....	595
14.6 在运营管理中的应用：发现并降低变异（可选） .....	612
附录 回顾第 12 章至第 14 章 .....	626
<b>第 15 章 卡方检验</b> .....	632
15.1 拟合优度的卡方检验 .....	633
15.2 列联表的卡方检验 .....	640
15.3 定类数据检验小结 .....	653
15.4 正态分布的卡方检验（可选） .....	655
附录 回顾第 12 章至第 15 章 .....	664
<b>第 16 章 简单线性回归和相关分析</b> .....	671
16.1 模型 .....	673
16.2 估计系数 .....	674
16.3 误差变量：必要条件 .....	684
16.4 评估模型 .....	686
16.5 回归方程的应用 .....	702
16.6 回归诊断 I .....	707

附录	回顾第 12 章至第 16 章	721
<b>第 17 章</b>	<b>多元回归</b>	729
17.1	模型及其必要条件	730
17.2	估计系数与评估模型	731
17.3	回归诊断 II	751
17.4	回归诊断 III (时间序列)	754
附录	回顾第 12 章至第 17 章	768
<b>第 18 章</b>	<b>建模</b>	775
18.1	多项式模型	776
18.2	定类自变量	784
18.3	在人力资源管理中的应用: 工资平等 (可选)	795
18.4	逐步回归 (可选)	802
18.5	建模	805
<b>第 19 章</b>	<b>非参数统计</b>	809
19.1	威尔科克森秩和检验	811
19.2	符号检验与威尔科克森符号秩和检验	826
19.3	KW 检验和弗里德曼检验	842
19.4	斯皮尔曼秩相关系数	856
附录	回顾统计推断方法: 第 12 章至第 19 章	873
<b>第 20 章</b>	<b>时间序列分析和预测</b>	888
20.1	时间序列构成因素	889
20.2	平滑方法	891
20.3	长期趋势与季节效应	904
20.4	预测概述	911
20.5	预测模型	914
<b>第 21 章</b>	<b>统计过程控制</b>	925
21.1	过程波动	926
21.2	控制图	927
21.3	变量控制图: $\bar{x}$ 图和 S 图	934
21.4	属性控制图: $p$ 图	948
<b>第 22 章</b>	<b>决策分析</b>	954
22.1	决策问题	955
22.2	附加信息的获取、利用和评估	961

第 23 章 结论 .....	977
附录 A 样本统计值 .....	979
附录 B 统计表 .....	994
附录 C 部分偶数习题答案 .....	1019
翻译说明 .....	1051



# 方差分析

## 社会综合调查——自由到保守的分类和收入

数据：GSS2012\*

美国人的政治态度受收入的影响，或者情况相反？如果这样，我们可能期望收入在根据以下测度（POLVIEW）定义的各组人群之间存在不同。

1. 极度自由
2. 自由
3. 轻微自由
4. 温和
5. 轻度保守
6. 保守
7. 极度保守

政治观点不同的 7 个人群间存在收入差异吗？问题的答案参见 14-1f 小节。

## 引 言

利用本章介绍的方法，统计人员可以比较两个或多个定距数据的总体，这种方法叫作方差分析（analysis of variance）。方差分析功能极为强大，是统计人员经常使用的一种方法。方差分析方法可用来判断总体均值之间是否存在差异。换句话说，该方法通过



分析样本方差来实现，它也是因此而得名的。我们将研究该方法的几种不同形式。

方差分析最早的应用之一是在 20 世纪 20 年代，用于判断施肥情况的不同是否会影  
响农作物产量。这一早期试验的术语目前仍在使用。不管是什么试验，其目的都在于研  
究组均值 (treatment means) 间是否存在较大差异。

## 14.1 单因素方差分析

方差分析是通过检验来判断两个或多个总体均值间是否存在差异的一个过程。这一  
方法的名称源于执行计算的方法，即方差分析方法通过分析数据的方差来确定我们能否  
推断出总体均值存在差异。如同第 13 章中一样，试验设计对选择哪种恰当的方法将起  
到决定性的作用。在这一节里，我们将讲述方差分析在独立样本情况下的应用流程。这  
种方法称为单因素方差分析 (one-way analysis of variance)。图 14—1 描绘了提取独立  
样本的抽样过程。总体  $j$  ( $j=1, 2, \dots, k$ ) 的均值和方差分别表示为  $\mu_j$  和  $\sigma_j^2$ ，这两  
个参数均是未知的。从每一个总体中抽取独立样本。对于每个独立样本，我们可以计算  
出均值  $\bar{x}_j$  和方差  $s_j^2$ 。

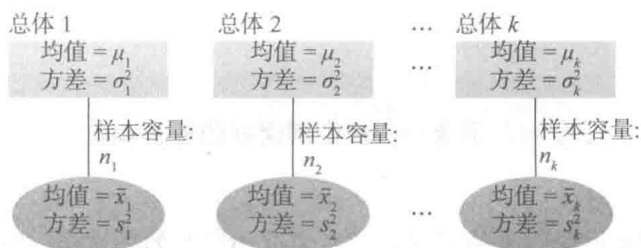


图 14—1 独立样本的抽样方案

### 例 14.1<sup>①</sup> 投资于股票市场的资产比例

数据: Xm14-01

在最近的十年里，股票经纪人的工作方式发生了戏剧性的变化。网上交易已经变得越  
来越普遍了，只需花费 7 美元。与以前相比，现在投资股票市场已经变得越来越方便和便  
宜了。这些变化带来了什么影响呢？为了回答这个问题，一位金融分析师随机抽取了 366  
个美国家庭，询问他们户主的年龄类别和投资在股票市场的资产比例。年龄类别为

- 年轻 (35 岁以下)
- 中年早期 (35~49 岁)
- 中年晚期 (50~65 岁)
- 老年 (65 岁以上)

该分析员想要判断资产比例是否会因股票所有人年龄的不同而不同。下面列出了部

<sup>①</sup> U. S. Census Bureau, "Asset Ownership of Households, May 2003," *Statistical Abstract of the United States*, 2006, Table 700.

分数据。根据这些数据，分析员能否得出这四个年龄段投资股票市场的资产比例是不同的这一结论？

年轻	中年早期	中年晚期	老年
24.8	28.9	81.5	66.8
35.5	7.3	0.0	77.4
68.7	61.8	61.3	32.9
42.2	53.6	0.0	74.0
⋮	⋮	⋮	⋮

解

我们已知数据是定距数据（投资在股票市场的资产比例），且我们的目标是比较四个总体（年龄段）。参数为四个总体均值  $\mu_1$ ， $\mu_2$ ， $\mu_3$  和  $\mu_4$ 。原假设表示我们假定四个样本均值间不存在差异。因此，我们有

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

我们采用方差分析方法来确定是否存在足够的证据表明原假设是错误的。相应地，备择假设为

$$H_1: \text{至少有两个均值不相等}$$

接下来我们将确定检验统计量，这要比我们到现在为止介绍过的检验统计量稍微复杂一些。利用表 14—1 中的符号，我们可以比较容易地进行方差分析。

表 14—1 单因素方差分析的符号

	组					
	1	2	⋯	$j$	⋯	$k$
	$x_{11}$	$x_{12}$	⋯	$x_{1j}$	⋯	$x_{1k}$
	$x_{21}$	$x_{22}$	⋯	$x_{2j}$	⋯	$x_{2k}$
	⋮	⋮	⋮	⋮	⋮	⋮
	$\bar{x}_{n_1 1}$	$\bar{x}_{n_2 2}$	⋯	$\bar{x}_{n_j j}$	⋯	$\bar{x}_{n_k k}$
样本容量	$n_1$	$n_2$	⋯	$n_j$	⋯	$n_k$
样本均值	$\bar{x}_1$	$\bar{x}_2$	⋯	$\bar{x}_j$	⋯	$\bar{x}_k$

注： $x_{ij}$  = 第  $j$  个样本的第  $i$  个观测值；

$n_j$  = 从总体  $j$  中抽取的样本的观测值的个数；

$$\bar{x}_j = \text{第 } j \text{ 个样本的均值} = \frac{\sum_{i=1}^{n_j} x_{ij}}{n_j};$$

$$\bar{x} = \text{所有观测值的总平均} = \frac{\sum_{j=1}^k \sum_{i=1}^{n_j} x_{ij}}{n}, \quad \text{其中 } n = n_1 + n_2 + \cdots + n_k, \quad k \text{ 是总体的数量。}$$

我们称变量  $x$  为**因变量**（response variable），称其值为**响应**（response），并称待测单元为**试验单元**（experimental unit）。在这一例子中，因变量为投资股票市场的资产

比例，试验单元则是抽取的户主。用于划分各个总体的标准被称为**因素** (factor)，每个总体被称为**因素水平** (level)。例 14.1 中的因素即户主的年龄类别，它共有四个水平。在本章的后面部分我们还将讨论总体由两个因素划分的试验。这一节我们只研究单因素试验。

## □ 检验统计量

检验统计量将按照以下基本原理进行计算。如果原假设为真，则总体均值全部相等，因而我们预期样本均值将十分接近；如果备择假设为真，则某些样本均值之间可能存在很大差异。衡量各样本均值彼此间接近程度的统计量被称为**组间变异** (between-treatments variation)，用 SST 表示，意为**组间平方和** (sum of squares for treatments)。

### 组间平方和

$$SST = \sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2$$

从上述公式我们可以推出，如果样本均值彼此接近，则所有样本的均值将近似于总体均值，因此 SST 的值会很小。事实上，当各样本均值完全相等时，SST 达到其最小值 (0)，即若

$$\bar{x}_1 = \bar{x}_2 = \dots = \bar{x}_k$$

则  $SST=0$ 。

于是我们可以认为，较小的 SST 值能够支持原假设。在本例中，我们计算样本均值和总均值如下：

$$\bar{x}_1 = 44.40$$

$$\bar{x}_2 = 52.47$$

$$\bar{x}_3 = 51.14$$

$$\bar{x}_4 = 51.84$$

$$\bar{x} = 50.18$$

样本容量为

$$n_1 = 84$$

$$n_2 = 131$$

$$n_3 = 93$$

$$n_4 = 58$$

$$n = n_1 + n_2 + n_3 + n_4 = 84 + 131 + 93 + 58 = 366$$

则

$$\begin{aligned} SST &= \sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2 \\ &= 84(44.40 - 50.18)^2 + 131(52.47 - 50.18)^2 + 93(51.14 - 50.18)^2 \\ &\quad + 58(51.84 - 50.18)^2 \\ &= 3\,738.8 \end{aligned}$$

如果样本均值间存在较大差异,那么至少某些样本均值与总体均值偏差甚远,这将会导致 SST 值偏大,这时候就应该拒绝原假设而选择备择假设。这一检验(与所有其他统计检验一样)中我们要回答的关键问题是:“统计量的值大到何种程度时,我们才可以证明拒绝原假设是合理的?”在我们的例子中,  $SST=3\,738.8$ 。这一数值是否已经足够大到可以证明总体均值不同?为了回答这个问题,我们应了解每周的销量之间究竟存在多大差别,这需要用组内变异(within-treatments variation)来衡量。组内变异可用 SSE [误差平方和(sum of squares for error)] 来表示,它为我们提供了衡量我们能够预期到的、已经观测的随机变量之间的差异的方法。在本例中,我们试图判断是否投资于股票市场的资产比例会因为户主年龄的不同而不同。然而,除了年龄以外,还存在其他变量会影响因变量。我们预期可能的变量有家庭收入、户主的职业和家庭规模。所有这些变量(还有一些我们没有想到的变量)都是变异的来源,我们将这些变量划为一类,统称为误差。这部分变异由误差平方和来衡量。

### 误差平方和

$$SSE = \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

将 SSE 展开,我们可以得到

$$SSE = \sum_{i=1}^{n_1} (x_{i1} - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_{i2} - \bar{x}_2)^2 + \cdots + \sum_{i=1}^{n_k} (x_{ik} - \bar{x}_k)^2$$

如果检查 SSE 的  $k$  个组成部分,你会发现,每个部分都是对其所在样本的可变性的一个测量。如果我们令每一组成部分除以  $n_j - 1$ , 就可以获得样本方差。我们可以把 SSE 表述成

$$SSE = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2$$

式中  $s_j^2$  是样本  $j$  的方差。因此, SSE 是  $k$  个样本的联合方差,这是对 13.1 节中计算的一个扩展。在 13.1 节中,我们使用一般总体方差的联合估计( $s_p^2$ )检验并估计了两总体均值之间的差异,应用该统计方法的一个必要条件是两总体方差相等。在使用 SSE 时,该条件依然是必要的,即我们要求

$$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$$

回到我们的例子中来,我们计算出样本方差为

$$s_1^2 = 386.55$$

$$s_2^2 = 469.44$$

$$s_3^2 = 471.82$$

$$s_4^2 = 444.79$$

因此

$$\begin{aligned} SSE &= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 + (n_4 - 1)s_4^2 \\ &= (84 - 1) \times 386.55 + (131 - 1) \times 469.44 + (93 - 1) \times 471.82 + (58 - 1) \end{aligned}$$

$$\begin{aligned} & \times 444.79 \\ & = 161\,871.3 \end{aligned}$$

接下来我们计算均方 (mean squares)。组内均方 (mean square for treatments) 是用 SST 除以总体数目与 1 的差得到的。

#### 组内均方

$$MST = \frac{SST}{k-1}$$

均方误 (mean square for error) 则由 SSE 除以样本总量 (式中用  $n$  表示) 与组数的差值确定。

#### 均方误

$$MSE = \frac{SSE}{n-k}$$

最后, 定义检验统计量为两个均方之比。

#### 检验统计量

$$F = \frac{MST}{MSE}$$

### □ 检验统计量的抽样分布

如果因变量服从正态分布, 则检验统计量  $F = \frac{MST}{MSE}$  服从自由度为  $k-1$  和  $n-k$  的  $F$  分布。在 8.4 节中, 我们介绍了  $F$  分布; 在 13.4 节中, 我们利用  $F$  分布来检验和估计了两个总体方差——在该应用中,  $F$  统计量是两个样本方差  $s_1^2$  和  $s_2^2$  之比。回顾一下 SST 和 SSE 的定义, 你会发现两者都与本书中用于计算样本方差  $s^2$  的公式的分子相似。将 SST 除以  $k-1$  并将 SSE 除以  $n-k$ , 就可以分别计算出 MST 和 MSE, 这时我们实际上计算的是一般总体方差的无偏估计量。因此, 比率  $F = MST/MSE$  就是两个样本方差之比。其自由度为上述两个样本方差的分母, 即

$$\begin{aligned} v_1 &= k-1 = 4-1 = 3 \\ v_2 &= n-k = 366-4 = 362 \end{aligned}$$

在例 14.1 中, 我们得到

$$\begin{aligned} MST &= SST/(k-1) = 3\,738.8/3 = 1\,246.27 \\ MSE &= SSE/(n-k) = 161\,871.3/362 = 447.16 \\ F &= MST/MSE = 1\,246.27/447.16 = 2.79 \end{aligned}$$

### □ 拒绝域和 $p$ 值

计算  $F$  统计量 ( $F$ -statistic) 的目的在于确定 SST 的值是否大得足以拒绝原假设。

正如我们所看到的, SST 越大,  $F$  值也越大。因此, 我们仅当  $F > F_{\alpha, k-1, n-k}$  时才拒绝原假设。

如果我们令  $\alpha = 0.05$ , 则例 14.1 的拒绝域为

$$F > F_{\alpha, k-1, n-k} = F_{0.05, 3, 362} \approx F_{0.05, 3, \infty} = 2.61$$

我们发现该检验统计量的值  $F = 2.79$ 。因此, 有足够证据表明四个年龄段投资者投资于股票市场的平均比例不同。

该检验的  $p$  值为

$$P(F > 2.79)$$

计算该值需要通过计算机完成, 计算结果为 0.040 5。

图 14—2 描绘了例 14.1 的抽样分布。

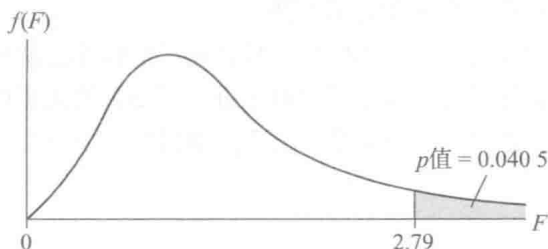


图 14—2 例 14.1 的抽样分布

我们通常用方差分析表 [analysis of variance (ANOVA) table] 给出方差分析的结果。表 14—2 给出了方差分析表的一般结构, 表 14—3 则给出了例 14.1 的方差分析表。

表 14—2 单因素方差分析的方差分析表

差异源	自由度	平方和	均值平方	$F$ 统计量
组内	$k-1$	SST	$MST = SST / (k-1)$	$F = MST / MSE$
误差	$n-k$	SSE	$MSE = SSE / (n-k)$	
总计	$n-1$	SS (总)		

表 14—3 例 14.1 的方差分析表

差异源	自由度	平方和	均值平方	$F$ 统计量
组内	3	3 738.8	1 246.27	2.79
误差	362	161 871.3	447.16	
总计	365	165 610.1		

方差分析表中使用的术语 (也就是检验本身使用的术语) 是以平方和的分解为基础的。这种分解源于以下等式 (其可变性可由求和规则来说明):

$$\sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x})^2 = \sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2 + \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2$$

等式左侧项表示全部数据的**总变异** (total variation), 用  $SS$  (总) 表示。将  $SS$  (总) 除以样本容量与 1 的差值 (即除以  $n-1$ ), 我们就得到样本方差 (假设原假设为真)。等号右侧的第一项为  $SST$ , 第二项为  $SSE$ 。我们看到, 总变异 [ $SS$  (总)] 被划分为两种变异: 组间平方和 ( $SST$ ) (源于组间均值的差异) 以及误差平方和 ( $SSE$ ) (对样本内的变异的度量)。因此, 上一等式可重新表示为

$$SS(\text{总}) = SST + SSE$$

故该检验也就由  $SST$  和  $SSE$  的比较而决定。

回忆一下在 13.3 节中讨论的配对试验的优势与劣势。我们指出, 统计人员经常千方百计地减少或解释随机变量的变异。在本节介绍的方差分析中, 组间平方和解释了组间的变异 (广告战略), 误差平方和则度量了不可预知的变异。若  $SST$  解释了总变异的大部分, 我们就可以得出总体均值不同的结论。在 14.4 节和 14.5 节中, 我们还将介绍方差分析的其他试验设计——减少或解释更多变异的方法。

如果在前面的章节中你用计算机和统计软件代替了统计学方法中的手工计算, 那么现在它们将更加有用, 因为计算机可以使我们避免进行需要耗费大量时间并且十分枯燥乏味的方差分析的手工计算。跟以前一样, 我们还是利用 Excel 和 Minitab 来解答例 14.1, 其结果如下所示。

## 计 算

### Excel

	A	B	C	D	E	F	G
1	方差分析:单因素方差分析						
2							
3	SUMMARY						
4	组	观测数	求和	平均	方差		
5	Young	84	3 279.5	44.40	386.55		
6	Early Middle Age	131	6 873.9	52.47	469.44		
7	Late Middle Age	93	4 755.9	51.14	471.82		
8	Senior	58	3 006.6	51.84	444.79		
9							
10							
11	方差分析						
12	差异源	SS	df	MS	F	P-value	F crit
13	组间	3 741.4	3	1 247.12	2.79	0.040 5	2.629 6
14	组内	161 871.0	362	447.16			
15							
16	总计	165 612.3	365				

## 计算机命令

1. 在相邻的两列中键入或者导入数据（打开 Xm14-01）。
2. 点击“数据”—“数据分析”，然后选择“方差分析：单因素方差分析”。
3. 确定“输入区域”（A1：D132）和 $\alpha$ 的值（0.05）。

## Minitab

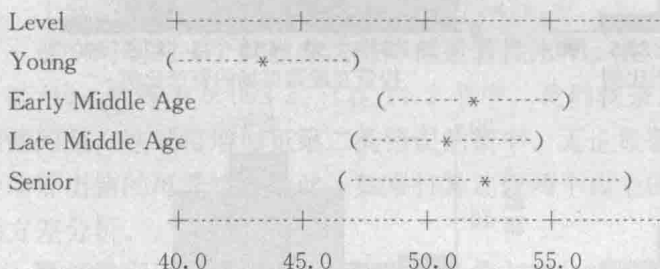
### One-way ANOVA: Young, Early Middle Age, Late Middle Age, Senior

Source	DF	SS	MS	F	P
Factor	3	3 741	1 247	2.79	0.041
Error	362	161 871	447		
Total	365	165 612			

S=21.15 R-Sq=2.26% R-Sq (adj) =1.45%

Level	N	Mean	StDev
Young	84	44.40	19.66
Early Middle Age	131	52.47	21.67
Late Middle Age	93	51.14	21.72
Senior	58	51.84	21.09

Individual 95% CIs For Mean Based on Pooled StDev



Pooled StDev=21.15

## 计算机命令

如果数据为非累积格式：

1. 键入或者导入数据（打开 Xm14-01）。
2. 点击“Stat”（统计）—“ANOVA”（方差分析）—“Oneway(Unstacked) ...” [单因子（非堆叠存放）...]。

3. 在“Responses (in separate columns)” [响应（在单独列中）] 框中，键入或选择处理的变量名称（Young, Early Middle Age, Late Middle Age, Senior）。点击“OK”（确定）。

如果数据是累积的：

1. 在两列中输入或导入数据。
2. 点击“Stat”（统计）—“ANOVA”（方差分析）—“Oneway...”（单因子...）。
3. 输入“Response”（响应）变量名称和“Factor”（因子）变量名称。点击“OK”（确定）。



检验统计量的值为  $F=2.79$ ， $p$  值为 0.040 5，这意味着我们有足够的证据可推出，至少有两个年龄段的投资者投资在股票市场的资产比例是不同的。

注意，本例中的数据是观测得到的。我们不能组织受控试验，因为组织这类试验要求金融分析师将每个家庭随机指派到某个年龄组中，而这是不可能的。

顺便介绍一下，在单因素方差分析中，如果数据是从受控试验中获得的，则我们将该试验设计称为完全随机化方差分析设计（completely randomized design of the analysis of variance）。

### □ 14-1a 检查必要条件

方差分析的  $F$  检验要求随机变量必须服从等方差的正态分布。通过绘出各样本的直方图，我们可以很方便地看出样本是否满足该要求。从图 14—3 所示的 Excel 直方图可知，我们没有理由认为这一条件没有得到满足。

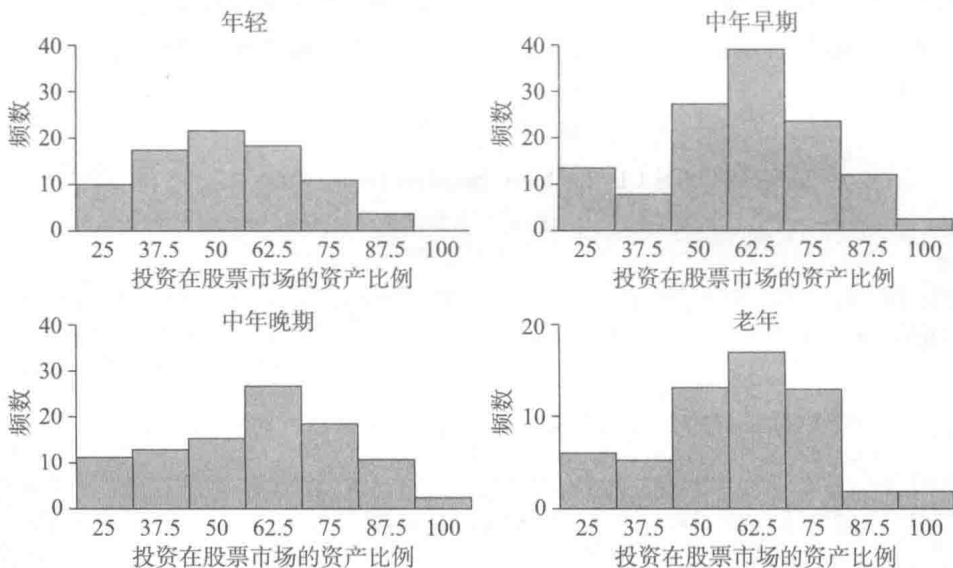


图 14—3 例 14.1 的直方图

是否等方差则通过输出样本标准离差或方差来判断。Excel 的输出结果含有方差，Minitab 可计算出标准离差。样本方差的近似性允许我们假定总体方差相等。在凯勒网站的附录中，我们介绍了巴特莱特 (Bartlett) 检验，它是专为检验方差是否相等而设计的。

### □ 14-1b 不满足必要条件

如果数据不服从正态分布，我们就可以用非参数检验方法取代单因素方差分析，我们称这种方法为 KW 检验，请参阅 19.3 节。<sup>①</sup>

如果总体方差不等，我们就可以利用几种方法来校正。不过，这些方法不在本书的

<sup>①</sup> 当不满足正态性时，为了检验两个或多个均值的差异，可采用非参数方法，想讲授这种方法的教师可参见凯勒网站附录——KW 检验和弗里德曼检验。