



21 世纪高等院校经济类与管理类教材

21SHIJI GAODENG YUANXIAO JINGJILEI YU
GUANLILEI JIAOCAI

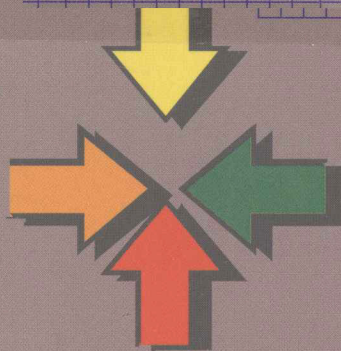
JILIANG JINGJIXUE JICHU LILUN YU FANGFA

计量经济学

——基础理论与方法

森栋公夫 赵国庆 著

21SHIJI GAODENG YUANXIAO JINGJILEI YU GUANLILEI JIAOCAI



21SHIJI GAODENG
YUANXIAO JINGJILEI
YU GUANLILEI
JIAOCAI

21SHIJI GAODENG
YUANXIAO JINGJILEI
YU GUANLILEI
JIAOCAI



中国金融出版社
CHINA FINANCIAL PUBLISHING HOUSE

21 世纪高等院校经济类与管理类教材

计量经济学

——基础理论与方法

森栋公夫 赵国庆 著

责任编辑：孔德蕴

责任校对：张志文

责任印制：尹小平

图书在版编目 (CIP) 数据

计量经济学：基础理论与方法 (Jiliang Jingjixue: Jichu Lilun Yu Fangfa) /
(日) 森栋公夫, 赵国庆著. —北京：中国金融出版社, 2010. 2

21 世纪高等院校经济类与管理类教材

ISBN 978 - 7 - 5049 - 5388 - 9

I. 计… II. ①森…②赵… III. 计量经济学的一高等学校—教材 IV.
F224. 0

中国版本图书馆 CIP 数据核字 (2010) 第 009192 号

出版

发行

中国金融出版社

社址 北京市丰台区益泽路 2 号

市场开发部 (010) 63272190, 66070804 (传真)

网上书店 <http://www.chinafph.com>
(010)63286832, 63365686 (传真)

读者服务部 (010) 66070833, 82672183

邮编 100071

经销 新华书店

印刷 北京松源印刷有限公司

装订 平阳装订厂

尺寸 170 毫米×228 毫米

印张 17. 25

字数 297 千

版次 2010 年 2 月第 1 版

印次 2010 年 2 月第 1 次印刷

印数 1—5060

定价 32. 00 元

ISBN 978 - 7 - 5049 - 5388 - 9/F. 4948

如出现印装错误本社负责调换

作者简介

森栋公夫先生为日本京都大学教授，历任日本京都大学经济学部长，日本经济学会会长（2006 年），日本统计学会、日本证券·计量·金融工程学会会长（2007~2009 年），日本文部科学省经济与管理学视学委员等职务。曾获美国斯坦福大学（Stanford University）和日本京都大学经济学双博士学位。

森栋先生长期从事计量经济学、统计学领域的理论研究，其研究成果始终处于本领域的前沿，其中多项研究成果为国际领先水平，森栋先生在国际上享有盛誉的 *Econometrica*、*Journal of American Statistical Association*、*International Economic Review*、*Review of Economics and Statistics*、*Journal of Econometrics*、*Econometric Theory*、*Econometric Reviews*、*Japanese Economic Review*、*Mathematics and Computers in Simulation* 等杂志上发表论文 50 余篇，其中多篇论文被计量经济学专著（*The Theory and Practice of Econometrics*、*Advanced Econometrics*、*The Practice of Econometrics-Classic and Contemporary*、*Estimation and Inference in Econometrics*、*Time-Series-Based Econometrics*）和国际杂志反复引用。

森栋先生近年出版计量经济学专著和教材十余部，其中《统计学》和《计量经济学》作为日本各大学的通用教材被广泛使用。

森栋先生自 20 世纪 80 年代起历任十余种国际著名经济学、计量经济学杂志的编委、国际顾问和主编等职，同时，担任伊利诺伊大学（University of Illinois）、斯坦福大学、西澳大利亚大学（University of Western Australia）等著名学府的访问教授。

日本统计学会为表彰森栋先生在计量经济学和统计学方面的贡献，在 2004 年度授予森栋先生统计学大奖。森栋先生现为日本学术会议会

员、世界计量经济学会荣誉会员 (Econometric Society Fellow)。

赵国庆为日本京都大学经济学博士，中国人民大学经济学院教授、博士生导师。

研究方向为计量经济学理论与应用。

前 言^{*}

经济理论的数量化研究是计量经济学的目的。计算机的日益更新和计量软件的多样化给计量经济分析方法的利用提供了强有力的工具。本书旨在使读者掌握常用的计量经济基础理论与方法，强调以解决实际问题为导向，训练学生的数量化基本技能。众所周知，习题的计算对于理解计量经济学是必不可少的重要环节，在计算过程中要用到 Excel、EViews 等软件，本书的例题和习题充分考虑了这些软件的功能。我们希望通过计算过程逐步培养读者对经济问题的理解力，使读者今后能够更好地运用计量经济分析方法解决实际问题。

书中复杂的数学推导均放在每章的附录部分，这是为方便欲进一步深入学习的读者。初学者跳过这部分，对于学习本书没有任何的影响。

本书共由九章构成。第 1 章为数据分析，主要介绍基本的统计分析和检验方法，包括均值、方差、相关性和假设检验等内容。第 2 章为一元线性回归模型，讨论最简单的单变量回归中所使用的最小二乘法的含义、计算方法、估计结果的解释和检验等基本内容。第 3 章为偏相关系数与回归分析，说明三变量数据的分析方法，包括相关系数矩阵、偏相关系数以及三变量回归分析的最小二乘法与检验。本章既是第 2 章两变量分析方法的扩展，也是第 4 章多元回归分析的基础。第 4 章为多元线性回归分析，讨论包含 K 个变量的线性回归模型，主要包括模型的系数估计、估计结果的评价指标以及对估计结果的检验

* 本成果受到中国人民大学“985”工程“中国经济研究哲学社会科学创新基地”项目的支持。

等内容,同时,对实证分析的标准化解释变量、虚拟变量及结构变化检验等方法加以说明。第5章讨论模型误差项存在的诸问题,如果误差项的分布相互不独立,标准的假设条件不成立,会给最小二乘估计与检验结果带来很大影响。本章将讨论不相关和同方差的假设条件不成立时如何进行估计,同时给出序列相关和异方差性的检验法。第6章为模型估计方法的扩展,本章介绍在实证分析中经常使用的一些模型,包括分布滞后模型、二元选择模型、受限因变量模型、面板数据模型以及包含多个回归方程的估计问题。第7章为统计分析的基础,将对与前面相关一些定理及内涵作进一步的讨论。首先,以离散型随机变量中的二项分布为例,讨论其数学期望,同时给出其样本均值等统计量的性质;其次,讨论连续型随机变量、多维随机变量以及样本均值的分布特性;最后,给出最小二乘估计量BLUE性质的证明。第8章为基于中国数据的实证分析。本章利用《中国高科技产业统计年鉴》的相关数据,借助面板数据分析方法考察科技创新投入、人力资源的积累以及FDI引致的技术外溢效应对中国高科技产业技术进步的影响。第9章讨论如何运用EViews对经济数据进行建模。EViews是一种计量经济分析的软件包,本章主要介绍EViews软件包的一些基本功能,同时给出运用EViews对经济模型进行估计和预测的一些简单实例。

在充分掌握计量经济学理论基础之上,再对模型进行估计,我们司空见惯的这种学习计量经济学的时代可能已经过去。今天是在进行模型估算的同时学习少量的计量理论,这是和过去运用计量经济分析方法正好相反的流程。我们认为对于计量经济学本身而言这是一个必然的发展趋势。不需要难懂的计量理论,掌握必要的计量估算方法的时代可能已经来临。

本书是在森栋公夫先生日文版的《计量经济学》一书基础上改写的著作。原书一共7章。本书在对原书内容进行修改的基础之上增加了第8章和第9章的内容。具体分工如下:第1章、第2章:宋智(日本横滨国立大学);第3章:王文杰(日本京都大学);第4章、第5章:刘庆丰(日本小樽商科大学);第6章、第7章、第8章、第

9章：赵国庆（中国人民大学）。全书由赵国庆负责最后的总体修改。

本书的参与者多为森栋公夫先生指导过的博士生。在此，我们对森栋先生表示最诚挚的谢意！

在本书的定稿过程中，中国人民大学数量经济学博士生张中元、林梦瑶、任宇宁、肖敏奇、尹慧等参与了部分工作，在此，向他们表示感谢。

中国金融出版社的同志为本书的编辑出版付出了辛勤的劳动，值此书付梓之际，表示由衷的感谢。

赵国庆
2010年1月

目 录

第 1 章 数据分析	1
1.1 均值和方差	1
1.2 相关	9
1.3 检验	13
习题	21
数学附录	22
第 2 章 一元线性回归模型	24
2.1 线性回归模型	24
2.2 回归模型的估计	27
2.3 决定系数	30
2.4 估计结果的检验	36
2.5 最小二乘法的性质	43
习题	48
数学附录	50
第 3 章 偏相关系数与回归分析	51
3.1 三变量数据分析	51
3.2 偏协方差与偏相关系数	53
3.3 回归分析	58
3.4 三变量最小二乘法	61
3.5 估计结果的检验	65
3.6 柯布一道格拉斯生产函数	69
习题	72
数学附录	73

第4章 多元线性回归分析	77
4.1 K 变量回归分析	77
4.2 估计结果的检验	85
4.3 变量的标准化	91
4.4 虚拟变量	96
4.5 残差平方和的性质	102
4.6 利用残差进行回归诊断	104
习题	108
数学附录	109
第5章 误差项存在的诸问题	112
5.1 序列相关	112
5.2 德宾—沃森检验	116
5.3 拉格朗日乘数检验	118
5.4 序列相关回归模型的估计	121
5.5 异方差性	125
5.6 其他诊断检验	129
习题	132
数学附录	133
第6章 估计方法的扩展	134
6.1 分布滞后模型	134
6.2 离散选择模型	142
6.3 受限因变量模型	147
6.4 面板数据分析	153
6.5 SUR 模型	159
6.6 联立方程组模型	162
习题	173
数学附录	175
第7章 统计分析基础	177
7.1 二项分布随机变量	177
7.2 二项分布随机变量的样本均值	180
7.3 二项分布随机变量的样本方差	183

7.4 离散型随机变量	185
7.5 连续型随机变量	187
7.6 联合分布	191
7.7 样本均值、方差、协方差	194
7.8 样本均值的分布	199
7.9 最小二乘估计量的分布	200
7.10 序列相关	208
习题	211
第 8 章 基于中国数据的实证分析	213
8.1 技术进步的度量	213
8.2 FDI 溢出效应对技术创新活动、技术进步的影响	219
8.3 估计结果的稳健性检验	221
第 9 章 运用 EViews 对经济数据的处理与建模	225
9.1 EViews 的主要功能	225
9.2 EViews 的基本规则	228
9.3 EViews 中线性方程的估计	236
9.4 EViews 中变量的引用和显示	244
9.5 矩阵运算	248
9.6 统计量的计算	250
部分习题答案与提示	253
附表 1 t 分布表	256
附表 2 标准正态分布表	256
附表 3 χ^2 分布表	257
附表 4 F 分布表	257
附表 5 德宾—沃森 (Durbin-Watson) 检验表 (1 % 的上下界)	259
附表 6 德宾—沃森 (Durbin-Watson) 检验表 (5 % 的上下界)	260
参考文献	261

第1章 数据分析

统计学的目的是用数据来研究观测对象的性质。观测对象的全体叫总体，决定总体特征的指标叫总体参数。总体参数的例子有总体均值或总体方差等，收集与观测对象有关的数据后，用数据来决定总体参数值的方法叫估计。数据也叫样本。在本章中所说明的样本均值和样本方差是总体均值和总体方差的一种估计方法。估计方法（其相应的计算公式）也叫估计量。

1.1 均值和方差

1. 样本均值

给出与变量 x 有关的数据 $(x_1, x_2, \dots, x_n)$ ，变量 x 的样本均值的定义是

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1.1)$$

希腊字母西格玛 ($\sum$) 是对全部观测值求和的意思，

$$\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n \quad (1.2)$$

样本均值表示 n 个观测值 $(x_1, x_2, \dots, x_n)$ 的中心的位置。

在初等统计学中，有时也使用中位数 (median) 作为观测值的中心的值。中位数是说，把 n 个观测值，按从最小值到最大值的顺序排成一列后，位于中心位置的值。观测值从小到大排列，如果观测值是 $(3, 7, 8, 12, 20)$ ，那么样本均值就是

$$\frac{1}{5}(3 + 7 + 8 + 12 + 20) = 10$$

因为中位数是中心位置的值，所以中位数是 8。

如果没有最后的观测值 20，那么样本均值就是 7.5。此时，中位数是 7 和 8 的平均值 7.5。此时，样本均值和中位数一致。

当观测值的个数是偶数的时候，因为中心位置的值无法确定，所以用中心的

2 个值的平均值作为中位数。

从数据中删除 20 的结果是：样本均值从 10 变为 7.5，缩小得很大，但是中位数却仅从 8 变为 7.5，缩小得很小。由此可见，中位数受特别大的值或特别小的值的影响很小，比较稳定。观测个数比较大的时候，这个性质会变得更加明显，这个性质也叫中位数的稳健性（robustness）。

加权平均也是一种广为人知的求平均数的方法。加权平均是对每个观测值给予不同的权数（weight），用权数来计算平均值。权数是非负的值，和为 1。例如刚才的数据，其加权平均值是

$$\frac{1}{9} \times 3 + \frac{2}{9} \times 7 + \frac{3}{9} \times 8 + \frac{2}{9} \times 12 + \frac{1}{9} \times 20 = \frac{85}{9} = 9.44$$

虽然每个观测值都乘以它的权数，但权数的和是 1。离中央近的值，给予比较大的权数。样本均值无非是相同权数的加权平均。一般来说，加权平均使用非负的权数 w_i 来定义。

$$\begin{aligned} \sum_{i=1}^n w_i x_i &= w_1 x_1 + w_2 x_2 + \cdots + w_n x_n \\ \sum_{i=1}^n w_i &= 1 \end{aligned} \quad (1.3)$$

权数也可以是 0，这表明对应的观测值被忽略。

2. 样本方差

样本方差是以样本均值为中心，表示数据离散程度的指标，由

$$s_x^2 = s_{xx} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (1.4)$$

来定义。如果不用求和记号“ $\sum$ ”，那么 n 项的平方和就是

$$\sum_{i=1}^n (x_i - \bar{x})^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2 \quad (1.5)$$

各项都是观测值与样本均值 $\bar{x}$ 之差的平方，西格玛是把这些项加起来的意思。因此，样本方差是指关于各个观测值和样本均值 $\bar{x}$ 的差的平方的平均。如果数据集中在均值的周围，那么样本方差的值就小。如果数据离均值较远，那么样本方差就大。

在求均值的时候，(1.4) 式被 $(n-1)$ 除。在本书中，样本方差用此定义来统一。用 n 除的情况也很多。当 n 较大的时候，结果基本上没有差别。对观测对象全体进行调查时叫普查（全面普查），此时的分母是 n 。

3. 样本方差的分解公式

计算样本方差可以由下面的公式简化：

$$\sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = \sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \quad (1.6)$$

证明 左边是

$$(x_1^2 - 2x_1\bar{x} + \bar{x}^2) + (x_2^2 - 2x_2\bar{x} + \bar{x}^2) + \cdots + (x_n^2 - 2x_n\bar{x} + \bar{x}^2)$$

因此, 整理括号内的第1项、第2项、第3项

$$\begin{aligned} & (x_1^2 + x_2^2 + \cdots + x_n^2) - 2(x_1 + x_2 + \cdots + x_n)\bar{x} + \overbrace{(\bar{x}^2 + \bar{x}^2 + \cdots + \bar{x}^2)}^{n\uparrow} \\ &= (x_1^2 + x_2^2 + \cdots + x_n^2) - 2n\bar{x}^2 + n\bar{x}^2 \end{aligned}$$

于是推导出(1.6)式的右边。与左边相比, 右边的计算比较容易。

根据样本均值的定义, 样本对于均值的离差之和为0:

$$\sum_{i=1}^n (x_i - \bar{x}) = 0 \quad (1.7)$$

利用这个性质, 可以给出另外的证明方法。

离差的平方和可以分解为

$$\begin{aligned} \sum_{i=1}^n [(x_i - \bar{x})^2] &= \sum_{i=1}^n [(x_i - \bar{x})(x_i - \bar{x})] \\ &= \sum_{i=1}^n [(x_i - \bar{x})x_i - (x_i - \bar{x})\bar{x}] \\ &= \sum_{i=1}^n (x_i - \bar{x})x_i - \sum_{i=1}^n (x_i - \bar{x})\bar{x} \end{aligned}$$

第1项是

$$\sum_{i=1}^n (x_i - \bar{x})x_i = \sum_{i=1}^n (x_i^2 - \bar{x}x_i) = \sum_{i=1}^n x_i^2 - \bar{x} \sum_{i=1}^n x_i$$

第2项是

$$\sum_{i=1}^n (x_i - \bar{x})\bar{x} = (x_1 - \bar{x})\bar{x} + \cdots + (x_n - \bar{x})\bar{x} = \bar{x} \sum_{i=1}^n (x_i - \bar{x})$$

因此, 根据(1.7)式, 该项为0。(证毕)

例如, 数据(3, 7, 8, 12, 20), 样本均值为10。该组数据的方差是, 根据对于均值10的离差的平方计算而得

$$\frac{1}{4}(7^2 + 3^2 + 2^2 + 2^2 + 10^2) = 41.5$$

或者也可以由观测值的平方和减去均值平方的5倍来计算:

$$\frac{1}{4}[(3^2 + 7^2 + 8^2 + 12^2 + 20^2) - 5 \times 10^2] = 41.5$$

如果不用平均的 5 倍去减, 而用总和的平方的 $\frac{1}{5}$ 去减也可以得到相同的答案:

$$\frac{1}{4} \left[(3^2 + 7^2 + 8^2 + 12^2 + 20^2) - \frac{1}{5} \times 50^2 \right] = 41.5$$

4. 样本标准差

样本方差正的平方根 s_x 称为样本标准差 (standard error)。样本方差是观测值和样本均值之间差的平方的平均值。因此, 样本方差是以样本均值为中心的。对于观测值的分布情况, 一般无法从样本方差得到直观的信息, 同时, 测定单位是原单位的平方, 因此有必要以样本标准差来表示数据的性质。

样本标准差的值, 在理解以均值为中心的数据的分散程度这一概念时非常有用。样本标准差, 很多时候写做 σ , 特别是下面要说明的 σ 区间是直观理解数据分散程度的重要工具。比如以均值为中心, 1σ 区间

$$(\text{均值} - \text{样本标准差}, \text{均值} + \text{样本标准差})$$

内, 含有一半以上的观测值, 2σ 区间

$$(\text{均值} - 2 \times \text{样本标准差}, \text{均值} + 2 \times \text{样本标准差})$$

内, 可以理解为含有大部分的观测值。更进一步, 3σ 区间

$$(\text{均值} - 3 \times \text{样本标准差}, \text{均值} + 3 \times \text{样本标准差})$$

内, 通常认为含有所有的观测值。

对于数据 (3, 7, 8, 12, 20) 而言, 标准差是 6.4, 因此 1σ 区间是

$$(10 - 6.4, 10 + 6.4) = (3.6, 16.4)$$

5 个观测值中有 3 个位于此区间内。所有的观测值都包含于 2σ 区间中。

当观测对象是来自服从正态分布的总体时, 包含于 1σ 区间内的观测值约占 70%, 2σ 区间内的观测值约占 95%, 3σ 区间内观测值占 99% 以上。对于判断数据的分散程度, 切比雪夫不等式虽然比较粗略, 但却是一个经常使用的标准。

切比雪夫不等式

k 为正整数, 以均值为中心, $k\sigma$ 区间由

$$(\bar{x} - k \times \text{样本标准差}, \bar{x} + k \times \text{样本标准差}) \quad (1.8)$$

来确定。在所有的观测值中, 不包含于此区间的观测值所占的比例在 $(1/k^2)$ 以下。

这个不等式的证明将在本章的最后给出。切比雪夫不等式虽然可以说对所有的数据都适用, 但是比如像超出 2σ 区间的观测值所占比例, 根据切比雪夫不等式所得结果是 $1/4 = 0.25$, 因此, 它只能够提供像小于 25% 这样非常粗略的信息。同样, 超出 3σ 区间的观测值所占比例小于 $1/9$ 。超出 1σ 区间的观测值所占比例小于 1, 这个结果基本上没有意义。虽然这个不等式较为粗略, 但对于观测

值分布的直观印象而言,却非常有用。

表 1-1

均值和标准差

SUM X ✓ f_x =SUM(h3:h20)								
	F	G	H	I	J	K	L	M
1								
2		国家或地区	G	G ²	L	L ²	G x L	
3		老挝	1.53	2.34	54	2916	83	
4		孟加拉国	1.65	2.72	61	3721	101	
5		印度	2.39	5.71	63	3969	151	
6		印度尼西亚	2.84	8.07	66	4356	187	
7		巴西	7.32	53.58	67	4489	490	
8		菲律宾	4.22	17.81	69	4761	291	
9		土耳其	7.03	49.42	69	4761	485	
10		泰国	6.33	40.07	69	4761	437	
11		中国	3.94	15.52	70	4900	276	
12		伊朗	5.90	34.81	71	5041	419	
13		马来西亚	8.36	69.89	72	5184	602	
14		沙特阿拉伯	11.05	122.10	72	5184	796	
15		韩国	17.34	300.68	73	5329	1266	
16		美国	34.26	1173.75	77	5929	2638	
17		新加坡	24.97	623.50	78	6084	1948	
18		以色列	19.32	373.26	78	6084	1507	
19		中国香港	25.66	658.44	80	6400	2053	
20		日本	26.46	700.13	81	6561	2143	
21		和	210.57	4251.80	1270	90430	15871	

SUM

Number1: h3:h20 = {1.53;1.65;2.39;2.84;7.32;4.22;7.03;6.33;3.94;5.90;8.36;11.05;17.34;34.26;24.97;19.32;25.66;26.46}

Number2: =

计算单元格区域中所有数值的和

Number1: number1, number2, ... 1 到 30 个待求和的数值。单元格中的逻辑值和文本将被忽略。但当作为参数键入时,逻辑值和文本有效

计算结果 = 210.57

有关该函数的帮助 (H)

确定 取消

G:通过购买力平价 (PPP) 来评价每个人的平均GNI(单位:1000美元)

GNI (国民总收入) 是GNP加上海外的一次性收入

L:出生时的平均寿命

资料来源:世界银行:《世界发展报告》2003年第2期

例 1.1 在表1-1中,以 Excel 的形式给出不同国家和地区 2002 年度的寿命和收入数据。 G 是收入, L 是寿命,此表表示了均值和方差的计算过程。表中 18 个国家和地区的平均收入是 11700 美元,方差是

$$\frac{1}{17} \times (4251.8 - \frac{1}{18} \times 210.6^2) = 105.16(\text{千美元})^2$$

标准差为 10.25 (千美元) = 10250 美元 (表中的和是 210.57)。注意方差的单位是千美元的平方。同样, 平均年龄是 70.6 岁, 方差是

$$\frac{1}{17} \times (90430 - \frac{1}{18} \times 1270^2) = 48.5 (\text{岁})^2$$

标准差是 7.0 岁。收入的 $1\hat{\sigma}$ 区间为

$$(11700 - 10250, 11700 + 10250) = (1450 \text{ 美元}, 21950 \text{ 美元})$$

18 个国家和地区中有 14 个国家和地区包含于此区间。 $2\hat{\sigma}$ 区间为

$$\begin{aligned} & (11700 - 2 \times 10250, 11700 + 2 \times 10250) \\ & = (-8800 \text{ 美元}, 32200 \text{ 美元}) \end{aligned}$$

在该区间内含有负收入。18 个国家和地区中只有美国超出了此区间。关于年龄的 $1\hat{\sigma}$ 区间为

$$(70.6 - 7, 70.6 + 7) = (63.6 \text{ 岁}, 77.6 \text{ 岁})$$

有 11 个国家和地区包含于此区间。 $2\hat{\sigma}$ 区间为

$$(70.6 - 2 \times 7, 70.6 + 2 \times 7) = (56.6 \text{ 岁}, 84.6 \text{ 岁})$$

只有老挝不在此区间。无论哪一组数据, 不包含于 $2\hat{\sigma}$ 区间的观测值所占比例只有 $1/18$, 也就是 5.6%。这个比例当然要比 $1/4$ 小。

例 1.2 表 1-1 中, 为了计算均值, 用 Excel 来求变量 G 列的总和。

(1) 将鼠标指针放在 H21 上, 点击函数 f_x , 选择 [数学与三角函数] 中的 [SUM]。

(2) 用鼠标, 将单元格 H3 到 H21 指定 [Number1] 的范围, 然后点击 [OK] 即可求出总和。

(3) 如果在工具栏中出现了 Σ 符号, 通过点击它就可求出列的总和。

(4) 均值是总和除以 18 求得的。或者在步骤 (1) 中, 点击函数 f_x , 选择 [统计] 中的 [AVERAGE], 然后按照步骤 (2) 进行。

计算方差时, 需要 G 列的值平方, 它是按以下的步骤进行的。

(1) 在 I3 中输入 [=H3^2]。

(2) 使用文件句柄, 计算直到单元格 I20 的平方值。文件句柄是将鼠标指针置于单元格 I3 右下角时出现的 [+] 记号。然后将这个 [+] 记号拖曳到单元格 I20 就可以了。在这里根据拖曳, 单元格的坐标发生变化是利用了 Excel 循环引用的性质。如果在单元格的纵坐标或者横坐标的前面加上 [\$], 无论怎么拖曳, 单元格的坐标都不发生变化。