



第4章：基本估计技术

学习目标

- ❖ (4.1) 建立并解释简单线性回归模型；
- ❖ (4.2) 利用最小二乘法估计回归参数的截距和斜率；
- ❖ (4.3) 利用t检验，或者参数估计相应的p值，决定估计的参数是否统计显著；
- ❖ (4.4) 利用R²统计检验来评价回归模型与数据吻合的程度，利用F检验来评价整个回归模型是否统计显著；
- ❖ (4.5) 建立并解释多解释变量的多元回归模型；
- ❖ (4.6) 用线性回归技术估计两个常用非线性模型的参数：二次回归模型和对数-?线性回归模型。

基本估计

$$C = a + bQ + cQ^2 + dQ^3$$

❖ 参数

~ 一个方程中的系数，用来决定方程中各变量间精确的数学关系。

❖ 参数估计

~ 对一个方程中的参数的具体数值做出估计的过程。

回归分析

❖ 回归分析

~ 一种统计技术，包括估计参数值和检验统计显著性

❖ 因变量

~ Variable whose variation is to be explained

❖ 自变量

~ 指那些被认为会影响因变量取值的经济学变量

简单线性回归

- 简单线性回归模型反映了因变量 Y 与解释变量 X 之间的关系

$$Y = a + bX$$

- 截距参数 (a) 给出了回归线与 Y -轴的交点处 Y 的值。(当 X 为零时 Y 的值)
- 斜率参数 (b) 表示 X 变化一单位引起的 Y 的变化。

$$b = \Delta Y / \Delta X$$

简单线性回归

❖ 通过选择使残差平方和最小的 a 和 b 的值来得到参数估计。

~ 残差是 Y 的真实值和拟合值之间的差： $Y_i - \hat{Y}_i$

❖ 样本回归直线是真实回归直线的估计

$$\hat{Y} = \hat{a} + \hat{b}X$$

拟合回归曲线

❖ 时间序列数据

~ 一个数据组，其中自变量与因变量的数据是在一段时间内从某个特定的公司（或特定的行业）中收集到的

❖ 横截面数据

~ 一个数据组，其中自变量与因变量的数据是在某个给定的时间内，从几个不同的公司或行业中收集到的

拟合回归曲线

❖ 最小二乘法

~ 用于估计线性回归方程参数值的一种方法。
它所确定的样本回归线使得各个样本数据点到直线的距离平方和最小。

拟合回归曲线

- ❖ 回归分析法确定的直线（即选定的a和b的值）使得残值的平方和（ $\sum e^2_i$ ）最小
 - ~ 实际值与拟合值（预测值）之间的差叫作残值： $Y_i - \hat{Y}_i$
 - ~ 它等于数据点到拟合的回归直线的垂直距离

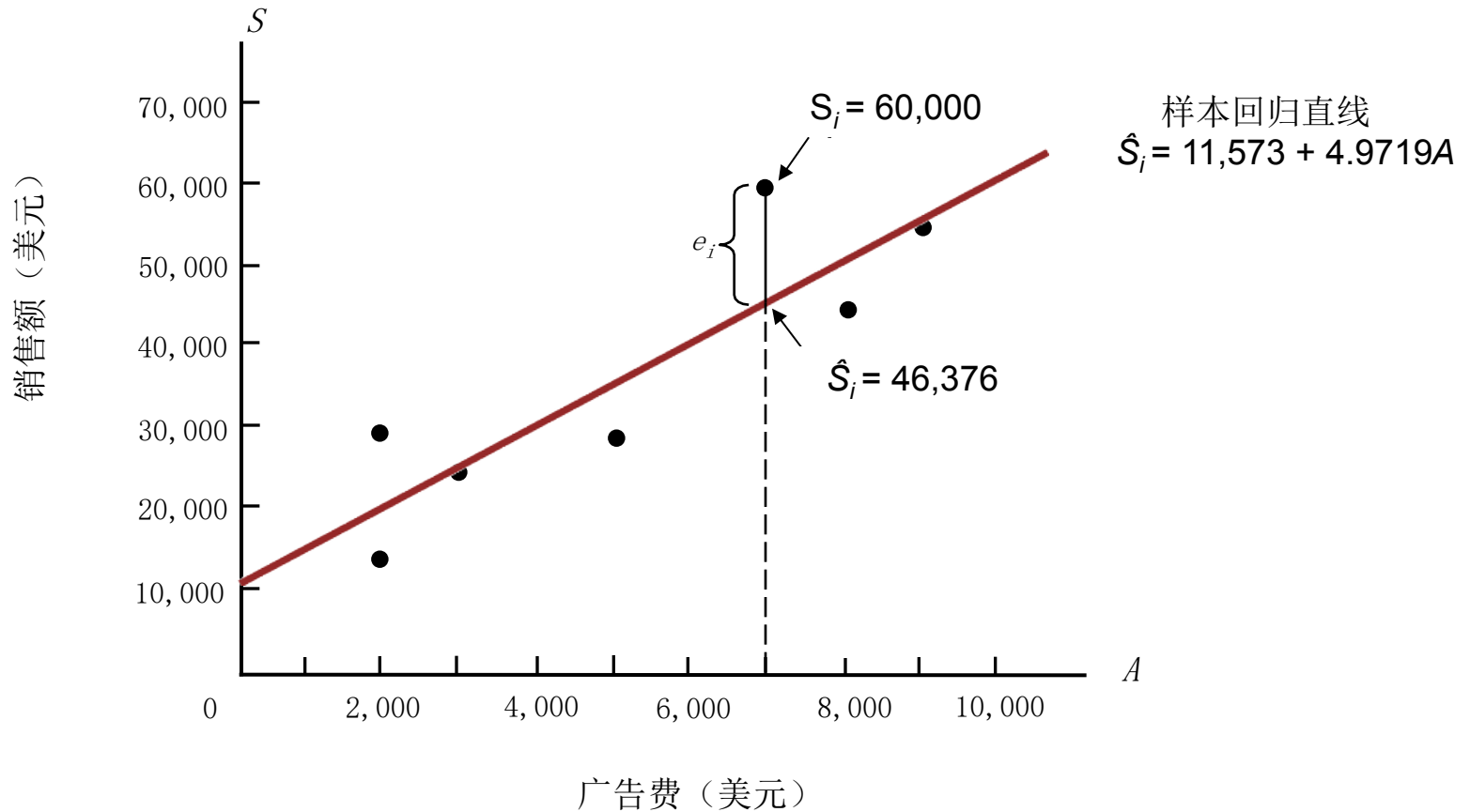
拟合回归线

❖ 样本回归线是对真实回归线的估计

$$\hat{Y} = \hat{a} + \hat{b}X$$

~ \hat{a} 和 \hat{b} 是对真实（人口）参数 a 和 b 用最小二乘法估计出的值

样本回归直线 (图 4.2)



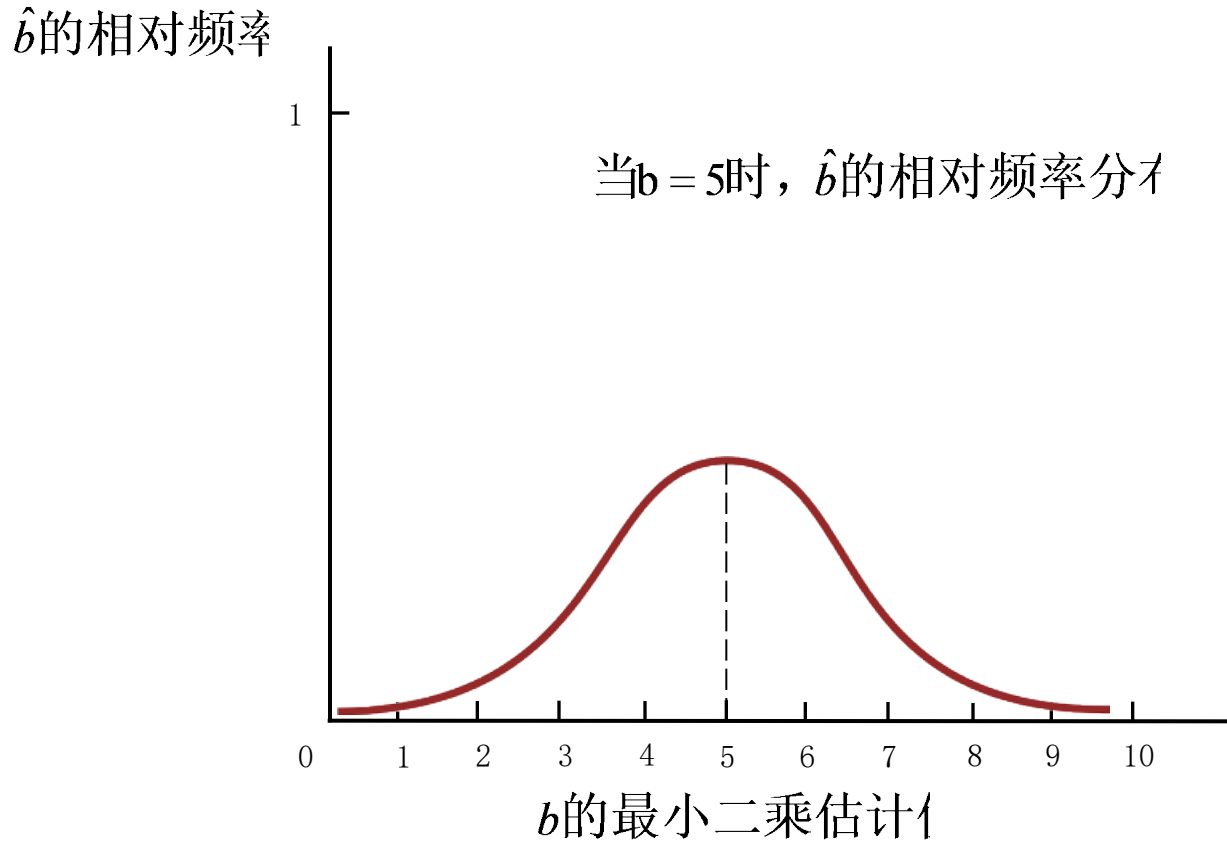
无偏估计量

- 估计的 \hat{a} 和 \hat{b} 通常不等于 a & b 的真值
 - \hat{a} 和 \hat{b} 是根据随机抽取的样本数据计算出的随机变量。
- ❖ 估计量取值的分布是以参数的真值为中心。

无偏估计量

- ❖ 一个估计量的均值（或期望值）等于该参数的真值时，该估计量是*无偏的*

相对频率分布* (图 4.3)



*也称为概率密度函数 (pdf)

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/138125000073006117>