

第 1 章 统计

一、学习指导

统计学是处理和分析数据的方法和技术，它几乎被应用到所有的学科检验领域。本章首先介绍统计学的含义和应用领域，然后介绍统计数据类型及其来源，最后介绍统计中常用的一些基本概念。本章各节的主要内容和学习要点如下表所示。

二、主要术语

1. 统计学：收集、处理、分析、解释数据并从数据中得出结论的科学。
2. 描述统计：研究数据收集、处理和描述的统计学分支。
3. 推断统计：研究如何利用样本数据来推断总体特征的统计学分支。
4. 分类数据：只能归于某一类别的非数字型数据。
5. 顺序数据：只能归于某一有序类别的非数字型数据。
6. 数值型数据：按数字尺度测量的观察值。
7. 观测数据：通过调查或观测而收集到的数据。
8. 实验数据：在实验中控制实验对象而收集到的数据。
9. 截面数据：在相同或近似相同的时间点上收集的数据。
10. 时间序列数据：在不同时间上收集到的数据。
11. 抽样调查：从总体中随机抽取一部分单位作为样本进行调查，并根据样本调查结果来推断总体特征的数据收集方法。
12. 普查：为特定目的而专门组织的全面调查。
13. 总体：包含所研究的全部个体（数据）的集合。
14. 样本：从总体中抽取的一部分元素的集合。
15. 样本容量：也称样本量，是构成样本的元素数目。
16. 参数：用来描述总体特征的概括性数字度量。
17. 统计量：用来描述样本特征的概括性数字度量。
18. 变量：说明现象某种特征的概念。
19. 分类变量：说明事物类别的一个名称。
20. 顺序变量：说明事物有序类别的一个名称。
21. 数值型变量：说明事物数字特征的一个名称。
22. 离散型变量：只能取可数值的变量。
23. 连续型变量：可以在一个或多个区间中取任何值的变量。

第 2 章 数据的图表展示

一、学习指导

数据的图表展示是应用统计的基本技能。本章首先介绍数据的预处理方法，然后介绍不同类型数据的整理与图示方法，最后介绍图表的合理使用问题。本章各节的主要内容和学习要点如下表所示。

二、主要术语

- 24. 频数：落在某一特定类别（或组）中的数据个数。
- 25. 频数分布：数据在各类别（或组）中的分配。
- 26. 比例：一个样本（或总体）中各个部分的数据与全部数据之比。
- 27. 比率：样本（或总体）中各不同类别数值之间的比值。
- 28. 累积频数：将各有序类别或组的频数逐级累加起来得到的频数。
- 29. 数据分组：根据统计研究的需要，将原始数据按照某种标准划分成不同的组别。
- 30. 组距分组：将全部变量值依次划分为若干个区间，并将这一区间的变量值作为一组。
- 31. 组距：一个组的上限与下限的差。
- 32. 组中值：每一组的下限和上限之间的中点值，即组中值 = (下限值 + 上限值) / 2。
- 33. 直方图：用矩形的宽度和高度（即面积）来表示频数分布的图形。
- 34. 茎叶图：由“茎”和“叶”两部分组成的、反应原始数据分布的图形。
- 35. 箱线图：由一组数据的最大值、最小值、中位数和两个四分位数 5 个特征值绘制而成的、反应原始数据分布的图形。

第 3 章 数据的概括性度量

一、学习指导

数据分布的特征可以从三个方面进行描述：一是分布的集中趋势，反映各数据向其中心值靠拢或聚集的程度；二是分布的离散程度，反映各数据远离其中心值的趋势；三是分布的形状，反映数据分布偏斜程度和峰度。本章将从数据的不同类型出发，分别介绍集中趋势测度值的计算方法、特点及其应用场合。本章各节的主要内容和学习要点如下表所示。

二、主要术语和公式

（一）主要术语

- 1. 众数：一组数据中出现频数最多的变量值，用 M_o 表示。

2. 中位数：一组数据排序后处于中间位置上的变量值，用 M_e 表示。
3. 四分位数：一组数据排序后处于 25%和 75%位置上的值。
4. 平均数：一组数据相加后除以数据的个数而得到的结果。
5. 几何平均数： n 个变量值乘积的 n 次方根，用 G_m 表示。
6. 异众比率：非众数组的频数占总频数的比率。
7. 四分位差：也称为内距或四分间距，上四分位数与下四分位数之差。
8. 极差：也称全距，一组数据的最大值与最小值之差。
9. 平均差：也称平均绝对离差，各变量值与其平均数离差绝对值的平均数。
10. 方差：各变量值与其平均数离差平方的平均数。
11. 标准差：方差的平方根。
12. 标准分数：变量值与其平均数的离差除以标准差后的值。
13. 离散系数：也称为变异系数，一组数据的标准差与其相应的平均数之比。
14. 偏态：数据分布的不对称性。
15. 偏态系数：对数据分布不对称性的度量值。
16. 峰态：数据分布的平峰或尖峰程度。
17. 峰态系数：对数据分布峰态的度量值。

（二）主要公式

名称	公式
中位数	$M_e = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ 为奇数} \\ \frac{1}{2} \left\{ x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)} \right\} & n \text{ 为偶数} \end{cases}$
简单样本平均数	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$
加权样本平均数	$\bar{x} = \frac{\sum_{i=1}^k M_i f_i}{n}$
几何平均数	$G_m = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$
异众比率	$V_r = \frac{\sum f_i - f_m}{\sum f_i} = 1 - \frac{f_m}{\sum f_i}$
四分位差	$Q_d = Q_U - Q_L$
极差	$R = \max(x_i) - \min(x_i)$

简单平均差	$M_d = \frac{\sum_{i=1}^n x_i - \bar{x} }{n}$
加权平均差	$M_d = \frac{\sum_{i=1}^k M_i - \bar{x} f_i}{n}$
简单样本方差	$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$
简单样本标准方差	$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$
加权样本方差	$s^2 = \frac{\sum_{i=1}^k (M_i - \bar{x})^2 f_i}{n-1}$
加权样本标准差	$s = \sqrt{\frac{\sum_{i=1}^k (M_i - \bar{x})^2 f_i}{n-1}}$
标准分数	$z_i = \frac{x_i - \bar{x}}{s}$
离散系数	$v_s = \frac{s}{\bar{x}}$
未分组数据的偏态系数	$SK = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \bar{x}}{s} \right)^3$
分组数据的偏态系数	$SK = \frac{\sum_{i=1}^k (M_i - \bar{x})^3 f_i}{ns^3}$
未分组数据的峰态系数	$K = \frac{n(n+1) \sum (x_i - \bar{x})^4 - 3 \left(\sum (x_i - \bar{x})^2 \right)^2 (n-1)}{(n-1)(n-2)(n-3)s^4}$
分组数据的峰态系数	$K = \frac{\sum_{i=1}^k (M_i - \bar{x})^4 f_i}{ns^4} - 3$

第 4 章 抽样与参数估计

一、学习指导

参数估计是推断统计的重要内容之一，它是在抽样及抽样分布的基础上，根据样本统计量来推断我们所关心的总体参数。本章首先介绍抽样分布的有关知识，然后讨论参数估计的一般问题，最后介绍一个总体参数估计的基本方法和参数估计中样本容量的确定问题。本章各节的主要内容和学习要点如下表所示。

章节	主要内容	学习要点
4.1 抽样与抽样分布	概率抽样方法	<ul style="list-style-type: none"> ▶ 概念：简单随机抽样，简单随机样本，重复抽样，不重复抽样，分层抽样，系统抽样，整群抽样。 ▶ 用 Excel 抽取简单随机样本。
	抽样分布	<ul style="list-style-type: none"> ▶ 概念：抽样分布，样本均值的抽样分布，样本比例的抽样分布，样本方差的抽样分布。 ▶ 中心极限定理。 ▶ 样本均值抽样分布的特征。 ▶ 样本均值的抽样分布与总体分布的关系。 ▶ 样本比例抽样分布的形式和特征。 ▶ 样本方差抽样分布的形式。
4.2 参数估计的一般问题	估计量与估计值	▶ 概念：估计量，估计值。
	点估计与区间估计	<ul style="list-style-type: none"> ▶ 概念：点估计，区间估计，置信区间，置信水平。 ▶ 置信区间构建的原理。 ▶ 置信区间的解释。
	评价估计量的标准	▶ 概念：无偏性，有效性，一致性。
4.3 总体均值的区间估计	正态总体、方差已知，或非正态总体、大样本	<ul style="list-style-type: none"> ▶ 正态总体、方差已知时的置信区间。 ▶ 非正态总体、大样本时的置信区间。 ▶ 用 Excel 计算给定 α 的正态分布的临界值。
	正态总体、方差未知、小样本	<ul style="list-style-type: none"> ▶ 正态总体、方差已知时的小样本置信区间。 ▶ 正态总体、方差未知时的小样本置信区间。 ▶ 用 Excel 计算给定 α 的 t 分布的临界值。
4.4 总体比例的区间估计	总体比例的区间估计	▶ 总体比例的置信区间。

计		
4.5 总体方差的区间估计	总体方差的区间估计	<ul style="list-style-type: none"> ▶ 总体方差的置信区间。 ▶ 用Excel计算给定α的χ^2分布的临界值。
4.6 样本容量的确定	估计总体均值时样本容量的确定	▶ 样本容量的计算方法。
	估计总体比例时样本容量的确定	▶ 样本容量的计算方法。

二、主要术语和公式

(一) 主要术语

36. 简单随机抽样：也称纯随机抽样，它是从含有 N 个元素的总体中，抽取 n 个元素作为样本，使得总体中的每一个元素都有相同的机会（概率）被抽中。
37. 简单随机样本：从含有 N 个元素的总体中，抽取 n 个元素作为样本，使得总体中每一个样本量为 n 的样本都有相同的机会（概率）被抽中。
38. 重复抽样：从总体中抽取一个元素后，把这个元素放回到总体中再抽取第二个元素，直至抽取 n 个元素为止。
39. 不重复抽样：一个元素被抽中后不再放回总体，而是从所剩元素中抽取第二个元素，直到抽取 n 个元素为止。
40. 分层抽样：也称分类抽样，它是在抽样之前先将总体的元素划分为若干层（类），然后从各个层中抽取一定数量的元素组成一个样本。
41. 系统抽样：也称等距抽样或机械抽样，它是先将总体中的各元素按某种顺序排列，并按某种规则确定一个随机起点；然后，每隔一定的间隔抽取一个元素，直至抽取 n 个元素形成一个样本。
42. 整群抽样：先将总体划分成若干群，然后以群作为抽样单位从中抽取部分群，随后再对抽中的各个群中所包含的所有元素进行观察。
43. 抽样分布：在重复选取样本量为 n 的样本时，由样本统计量的所有可能取值形成的相对频数分布。
44. 样本均值的抽样分布：在重复选取样本量为 n 的样本时，由样本均值的所有可能取值形成的相对频数分布。
45. 样本比例抽样分布：在重复选取样本量为 n 的样本时，由样本比例的所有可能取值形成的相对频数分布。
46. 标准误差：也称为标准误，它是样本统计量的抽样分布的标准差。
47. 估计标准误差：若计算标准误时所涉及的总体参数未知，可用样本统计量代替计算的标准误。
48. 估计量：用来估计总体参数的统计量的名称，用符号 $\hat{\theta}$ 表示。
49. 估计值：用来估计总体参数时计算出来的估计量的具体数值。
50. 点估计：用样本统计量 $\hat{\theta}$ 的某个取值直接作为总体参数 θ 的估计值。
51. 区间估计：在点估计的基础上，给出总体参数估计的一个范围。

- 52.** 置信区间：由样本统计量所构造的总体参数的估计区间。
- 53.** 置信水平：也称为置信系数，它是将构造置信区间的步骤重复多次后，置信区间中包含总体参数真值的次数所占的比率。

（二）主要公式

名称	公式
总体均值的置信区间（正态总体， σ 已知）	$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$
总体均值的置信区间（ σ 未知，大样本）	$\bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$
总体均值的置信区间（正态总体， σ 未知，小样本）	$\bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$
总体比例的置信区间	$p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$
总体方差的置信区间	$\frac{(n-1)s^2}{\chi_{\alpha/2}^2} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi_{1-\alpha/2}^2}$
估计总体均值时的样本容量	$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2}$
估计总体比例时的样本容量	$n = \frac{(z_{\alpha/2})^2 \cdot \pi(1-\pi)}{E^2}$

第 5 章 假设检验

一、学习指导

假设检验是推断统计的另一项重要内容，它是利用样本信息判断假设是否成立的一种统计方法。本章首先介绍有关假设检验的一些基本问题，然后介绍一个总体参数的检验方法。本章各节的主要内容和学习要点如下表所。

二、主要术语和公式

（一）主要术语

- 18.** 假设：对总体参数的具体数值所做的陈述。

19. 假设检验：先对总体参数提出某种假设，然后利用样本信息判断假设是否成立的过程。
20. 备择假设：也称研究假设，是研究者想收集证据予以支持的假设，用 H_1 或 H_a 表示。
21. 原假设：也称零假设，是研究者想收集证据予以反对的假设，用 H_0 表示。
22. 单侧检验：也称单尾检验，是指备择假设具有特定的方向性，并含有符号“>”或“<”的假设检验。
23. 双侧检验：也称双尾检验，是指备择假设没有特定的方向性，并含有符号“≠”的假设检验。
24. 第 I 类错误：当原假设为正确时拒绝原假设，犯第 I 类错误的概率记为 α 。
25. 第 II 类错误：当原假设为错误时没有拒绝原假设，犯第 II 类错误的概率通常记为 β 。
26. 显著性水平：假设检验中发生第 I 类错误的概率，记为 α 。
27. 检验统计量：根据样本观测结果计算得到的，并据以对原假设和备择假设做出决策的某个样本统计量。
28. 拒绝域：能够拒绝原假设的检验统计量的所有可能取值的集合。
29. 临界值：根据给定的显著性水平确定的拒绝域的边界值。
30. P 值：也称观察到的显著性水平，如果原假设 H_0 是正确的，那么所得的样本结果出现实际观测结果那么极端的概率。

（二）主要公式

名称	公式
总体均值检验的统计量（正态总体， σ 已知）	$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$
总体均值检验的统计量（ σ 未知，大样本）	$z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$
总体均值检验的统计量（正态总体， σ 未知，小样本）	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$
总体比例检验的统计量	$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}}$
总体方差检验的统计量	$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$

第 6 章 方差分析

一、学习指导

本章主要介绍检验多个总体均值是否相等的一种统计方法，即方差分析。它是通过对各观察数据误差来源的分析来判断多个总体均值是否相等。本章首先介绍方差分析中的一些基本问题，包括方差分析中的一些术语、方差分析的基本思想和基本假设，然后介绍单因素方差分析方法，最后介绍方差分析中的多重比较。本章各节的主要内容和学习要点如下表所示。

二、主要术语和公式

(一) 主要术语

- 31. 方差分析(ANOVA): 检验多个总体均值是否相等的统计方法。
- 32. 因素: 也称因子，是方差分析中所要检验的对象。
- 33. 水平: 也称处理，是因素的不同表现。
- 34. 组内误差: 来自水平内部的数据误差。
- 35. 组间误差: 来自不同水平之间的数据误差。
- 36. 总平方和: 反映全部数据误差大小的平方和，记为 SST 。
- 37. 组内平方和: 反映组内误差大小的平方和，记为 SSE 。
- 38. 组间平方和: 反映组间误差大小的平方和，记为 SSA 。
- 39. 单因素方差分析: 只涉及一个分类型自变量的方差分析。
- 40. 组内方差: 组内平方和除以相应的自由度。
- 41. 组间方差: 组间平方和除以相应的自由度。

(二) 主要公式

名称	公式
组间方差	$MSA = \frac{\text{组间平方和}}{\text{自由度}} = \frac{SSA}{k-1}$
组内方差	$MSE = \frac{\text{组内平方和}}{\text{自由度}} = \frac{SSE}{n-k}$
方差分析的检验统计量	$F = \frac{MSA}{MSE} \sim F(k-1, n-k)$
关系强度的测量	$R^2 = \frac{SSA(\text{组间SS})}{SST(\text{总SS})}$
多重比较的 LSD	$LSD = t_{\alpha/2} \sqrt{MSE(\frac{1}{n_i} + \frac{1}{n_j})}$

第7章 相关与回归分析

一、学习指导

相关与回归是研究变量之间关系的统计方法,该方法广泛应用于自然科学和社会科学的各个领域。本章首先介绍相关分析方法,然后介绍一元线性回归和多元线性回归分析方法。本章各节的主要内容和学习要点如下表所示。

二、主要术语和公式

(一) 主要术语

- 42. 相关关系: 变量之间存在的的数量关系。
- 43. 相关系数: 也称 Pearson 相关系数, 是根据样本数据计算的度量两个变量之间线性关系强度的统计量。
- 44. 因变量: 被预测或被解释的变量, 用 y 表示。
- 45. 自变量: 用来预测或用来解释因变量的一个或多个变量, 用 x 表示。
- 46. 回归模型: 描述因变量 y 如何依赖于自变量 x 和误差项 ε 的方程。
- 47. 回归方程: 描述因变量 y 的期望值如何依赖于自变量 x 的方程。
- 48. 估计的回归方程: 根据样本数据求出的回归方程的估计。
- 49. 最小二乘法: 也称最小平方法, 使因变量的观察值 y_i 与估计值 \hat{y}_i 之间的离差平方和达到最小来求得 $\hat{\beta}_0$ 和 $\hat{\beta}_1$ 的方法。
- 50. 判定系数: 回归平方和占总平方和的比例, 记为 R^2 。
- 51. 估计量的标准误差: 均方残差 (MSE) 的平方根, 用 s_e 来表示。
- 52. y 的平均值的点估计: 利用估计的回归方程, 对于 x 的一个特定值 x_0 , 求出 y 的平均值的一个估计值 $E(y_0)$ 。
- 53. y 的个别值的估计值: 利用估计的回归方程, 对于 x 的一个特定值 x_0 , 求出 y 的一个个别值的估计值 \hat{y}_0 。
- 54. y 的平均值的置信区间估计: 对 x 的一个给定值 x_0 , 求出 y 的平均值的区间估计。
- 55. y 的个别值的预测区间估计: 对 x 的一个给定值 x_0 , 求出 y 的一个个别值的区间估计。
- 56. 多元线性回归模型: 描述因变量 y 如何依赖于自变量 x_1, x_2, \dots, x_k 和误差项 ε 的方

程。

57. 多元线性回归方程：描述 y 的期望值如何依赖于 x_1, x_2, \dots, x_k 的方程。
58. 估计的多元线性回归方程：根据样本数据得到的多元线性回归方程的估计。
59. 多重判定系数：在多元回归中，回归平方和占总平方和的比例。
60. 修正的多重判定系数：用模型中自变量的个数和样本量进行调整的多重判定系数，记为

$$R_a^2。$$

(二) 主要公式

名称	公式
相关系数	$r = \frac{n \sum xy - \sum x \times \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \times \sqrt{n \sum y^2 - (\sum y)^2}}$
相关系数检验的统计量	$t = r \sqrt{\frac{n-2}{1-r^2}} \sim t(n-2)$
回归方程的截距	$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$
回归方程的斜率（回归系数）	$\hat{\beta}_1 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}$
判定系数	$R^2 = \frac{SSR}{SST} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$
估计标准误差	$s_e = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n-2}} = \sqrt{\frac{SSE}{n-2}}$
线性关系检验的统计量	$F = \frac{SSR/1}{SSE/n-2} \sim F(n-2)$
回归系数检验的统计的统计量	$t = \frac{\hat{\beta}_i}{s_{\beta_i}} \sim t(n-2)$
y 的平均值的置信区间	$\hat{y}_0 \pm t_{\alpha/2} s_e \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$

y 的个别值的预测区间	$\hat{y}_0 \pm t_{\alpha/2} S_e \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}}$
修正的多重判定系数	$R^2 = 1 - (1 - R^2) \times \frac{n-1}{n-k-1}$

第 8 章 时间序列分析和预测

一、学习指导

分析时间序列数据的主要目的是对未来的观测值进行预测。本章在给出时间序列概念及分类的基础上，首先介绍了时间序列的描述性分析方法，然后介绍了平稳序列和非平稳序列的一些简单预测方法。本章各节的主要内容和学习要点如下表所示。

二、主要术语和公式

（一）主要术语

- 61.** 时间序列：同一现象在不同时间上的相继观察值排列而成的序列。
- 62.** 平稳序列：基本上不存在趋势的序列。
- 63.** 非平稳序列：包含趋势性、季节性或周期性的序列。
- 64.** 趋势：也称长期趋势，是指时间序列在长时期内呈现出来的某种持续向上或持续下降的变动。
- 65.** 季节性：也称季节变动，是指时间序列在一年内重复出现的周期性波动。
- 66.** 周期性：也称循环波动，是指时间序列中呈现出来的围绕长期趋势的一种波浪形或振荡式变动。
- 67.** 随机性：也称不规则波动，是指时间序列中除去趋势、周期性和季节性之后的偶然性波动。
- 68.** 增长率：也称增长速度，是指时间序列中报告期观察值与基期观察值之比减 1 后的结果，用%表示。
- 69.** 平均增长率：也称平均发展速度，是指时间序列中各逐期环比值（也称环比发展速度）的几何平均数减 1 后的结果。
- 70.** 增长 1%绝对值：增长率每增长一个百分点而增加的绝对数量。
- 71.** 简单平均法预测：根据过去已有的 t 期观察值通过简单平均来预测下一期的数值。
- 72.** 移动平均法预测：通过对时间序列逐期递移求得平均数作为预测值的一种预测方法。
- 73.** 指数平滑法预测：对过去的观察值加权平均进行预测的一种方法，该方法使得第 $t+1$ 期的预测值等于 t 期的实际观察值与第 t 期指数预测值的加权平均值。

（二）主要公式

名称	公式
----	----

环比增长率	$G_i = \frac{Y_i}{Y_{i-1}} - 1$
定基增长率	$G_i = \frac{Y_i - Y_0}{Y_0} = \frac{Y_i}{Y_0} - 1$
平均增长率	$\bar{G} = \sqrt[n]{\frac{Y_1}{Y_0} \times \frac{Y_2}{Y_1} \times \cdots \times \frac{Y_n}{Y_{n-1}}} - 1 = \sqrt[n]{\frac{Y_n}{Y_0}} - 1$
年度化增长率	$G_A = \left(\frac{Y_i}{Y_{i-1}}\right)^{m/n} - 1$
平均预测误差	$ME = \frac{\sum_{i=1}^n (Y_i - F_i)}{n}$
平均绝对预测误差	$MAD = \frac{\sum Y_i - F_i }{n}$
均方预测误差	$MSE = \frac{\sum_{i=1}^n (Y_i - F_i)^2}{n}$
平均百分比预测误差	$MPE = \frac{\sum \left(\frac{Y_i - F_i}{Y_i} \times 100 \right)}{n}$
简单平均法预测	$F_{t+1} = \frac{1}{t} (Y_1 + Y_2 + \cdots + Y_t) = \frac{1}{t} \sum_{i=1}^t Y_i$
移动平均法预测	$F_{t+1} = \bar{Y}_t = \frac{Y_{t-k+1} + Y_{t-k+2} + \cdots + Y_{t-1} + Y_t}{k}$
指数平滑法预测	$F_{t+1} = \alpha Y_t + (1 - \alpha) F_t$
线性趋势方程的截距和斜率	$\begin{cases} b = \frac{n \sum tY - \sum t \sum Y}{n \sum t^2 - (\sum t)^2} \\ a = \bar{Y} - b \bar{t} \end{cases}$
二次曲线的标准方程组	$\begin{cases} \sum Y = na + b \sum t + c \sum t^2 \\ \sum tY = a \sum t + b \sum t^2 + c \sum t^3 \\ \sum t^2 Y = a \sum t^2 + b \sum t^3 + c \sum t^4 \end{cases}$

指数曲线的标准方程组	$\begin{cases} \sum \lg Y = n \lg a + \lg b \sum t \\ \sum t \lg Y = \lg a \sum t + \lg b \sum t^2 \end{cases}$
修正指数曲线的未知数	$\begin{cases} b = \left(\frac{S_3 - S_2}{S_2 - S_1} \right)^{\frac{1}{m}} \\ a = (S_2 - S_1) \frac{b-1}{b(b^m - 1)^2} \\ K = \frac{1}{m} \left(S_1 - \frac{ab(b^m - 1)}{b-1} \right) \end{cases}$
龚铂茨曲线的未知数	$\begin{cases} b = \left(\frac{S_3 - S_2}{S_2 - S_1} \right)^{\frac{1}{m}} \\ \lg a = (S_2 - S_1) \frac{b-1}{b(b^m - 1)^2} \\ \lg K = \frac{1}{m} \left(S_1 - \frac{b(b^m - 1)}{b-1} \times \lg a \right) \end{cases}$
Logistic 曲线未知数	$\begin{cases} b = \left(\frac{S_3 - S_2}{S_2 - S_1} \right)^{\frac{1}{m}} \\ a = (S_2 - S_1) \frac{b-1}{b(b^m - 1)^2} \\ K = \frac{1}{m} \left(S_1 - \frac{ab(b^m - 1)}{b-1} \right) \end{cases}$

第 9 章 指数

一、学习指导

指数是应用于经济领域的一种特殊统计方法。本章首先介绍指数的概念和分类，然后介绍加权指数的编制方法，最后介绍实际中几种常用的价格指数。本章各节的主要内容和学习要点如下表所示。

章节	主要内容	学习要点
9.1 引言	指数的概念和分类	► 概念：指数，个体指数，综合指数，简单

		指数，加权指数，时间性指数，区域性指数。
9.2 加权指数	加权综合指数	<ul style="list-style-type: none"> ▶ 概念：加权综合指数。 ▶ 加权综合指数的编制。
	加权平均指数	<ul style="list-style-type: none"> ▶ 概念：加权平均指数。 ▶ 加权平均指数的编制。
	价值指数与指数体系	<ul style="list-style-type: none"> ▶ 概念：价值指数，指数体系。 ▶ 指数体系的应用。
9.3 几种常用的价格指数	几种常用的价格指数	<ul style="list-style-type: none"> ▶ 概念：零售价格指数，居民消费价格指数，生产价格指数，股票价格指数。 ▶ 零售价格指数和居民消费价格指数的区别。 ▶ 居民消费价格指数的作用。

二、主要术语和公式

（一）主要术语

- 74.** 指数：测定多个项目在不同场合下综合变动的相对数。
- 75.** 加权综合指数：通过加权来测定一组项目的综合变动状况的指数。
- 76.** 加权平均指数：以某一时期的价值总量为权数对个体指数加权平均计算的指数。
- 77.** 价值指数：由两个不同时期的价值总量对比形成的指数。
- 78.** 指数体系：由价值指数及其若干个因素指数构成的数量关系式。
- 79.** 零售价格指数：反映城乡商品零售价格变动趋势的一种经济指数。
- 80.** 居民消费价格指数：反映一定时期内城乡居民所购买的生活消费品价格和服务项目价格的变动趋势和程度的一种相对数。
- 81.** 生产价格指数：测量在初级市场上所售货物（即在非零售市场上首次购买某种商品时）价格变动的一种价格指数。
- 82.** 股票价格指数：是反映某一股票市场上多种股票价格变动趋势的一种相对数。

（二）主要公式

名称	公式
加权综合价格指数	$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$
加权综合销售量指数	$I_q = \frac{\sum p_0 q_1}{\sum p_0 q_0}$
加权平均价格指数	$I_p = \frac{\sum p_1 q_1}{\sum \frac{1}{p_1/p_0} p_1 q_1}$

加权平均销售量指数	$I_q = \frac{\sum \frac{q_1}{q_0} p_0 q_0}{\sum p_0 q_0}$
价值指数	$v = \frac{\sum p_1 q_1}{\sum p_0 q_0}$