



第四章 土地统计分析

本章学习重点：本章主要讲述土地统计分析的内容、种类和方法。要掌握土地综合指标分析、土地动态数列分析、土地平衡分析、土地现象相关分析、回归分析等分析方法。

第四章 土地统计分析

第一节 土地统计分析概述

第二节 土地综合指标分析

第三节 土地统计指数分析

第四节 土地动态数列分析

第五节 土地动态平衡分析

第六节 土地相关回归分析

第一节 土地统计分析概述

- 一、土地统计分析的概念
- 二、土地统计分析的内容
- 三、土地统计分析的种类
- 四、土地统计分析的意义
- 五、土地统计分析的方法

一、土地统计分析的概念

- **土地统计分析**是对土地数量结构、利用状况、权属状况的区域分布特征、动态变化规律及其内在联系与发展趋势进行分析研究。
- **目的**：发现土地利用、开发、保护及土地管理等方面的新问题、新矛盾、新情况、新动态，为在国民经济建设和土地管理中加强宏观调控、科学决策，及时正确制定政策、法规、指令计划等提供依据。

二、土地统计分析的内容

计算	计算各种分析 指标 ，包括计算绝对数、相对数和平均数等，并作进一步研究
评价	对统计和分析指标进行土地现象的 判断和评价 ，即对各种统计数据及分析指标显示的直观与表面现象作出科学的、客观的判断和评价。
对动态变化进行分析	对土地数量、质量、权属和分布等动态变化进行分析，从总体的特殊表现过渡到总体的一般表现，从而对土地现象形成 规律性的认识 。
对土地动态变化趋势进行分析	对土地动态变化趋势进行分析、 预测和推论 ，即从对土地现状的认识过渡到土地 未来 状况的认识。
撰写报告	撰写分析报告，形成各种分析数据、图表及文字材料等成果

三、土地统计分析的种类

进度分析	主要以定期的报表为依据反映土地管理部门各项工作进展情况，检查各项计划执行情况，又称计划执行情况分析。
专题分析	针对土地管理部门的某项专门问题进行分析。
发展水平分析	对土地问题发展水平及发展速度状况进行分析。
综合分析	是对土地问题及相关诸方面进行比较全面的分析，为土地利用总体规划、土地管理决策等提供有力依据。
预测分析	统计预测，合理推测土地利用结构、变化趋势

四、土地统计分析的意义

- 1、是实现统计优质服务的关键
- 2、是实施土地统计监督的基础
- 3、是促进土地统计事业发展的重要环节

五、土地统计分析方法

- 土地统计分析工作，是统计上所特有的分析方法在土地管理工作中的应用。
- 常用土地统计分析方法：主要有综合指标分析法、动态数列分析法、统计指数分析法、相关分析和平衡分析法等，并构成统计分析方法体系。

综合指标分析	总量指标——	绝对指标	
		结构相对指标	
		比较相对指标	
相对指标		强度相对指标	
		比例相对指标	
		动态相对指标	
计划完成相对指标		计划完成相对指标	
		算术平均数	简单平均数
			加权平均数
平均指标		调和平均数	
		中位数	
		众数	
标志变动度		标志变动度	
动态数列分析	绝对数动态数列		
	相对数动态数列		
	平均数动态数列		
动态变化水平分析指标	动态变化水平分析指标		
	动态变化速度分析指标		
统计指数分析	综合指数		
	平衡指数		
	指数体系因素分析		
相关分析	直线相关系数的测定		
	直线回归方程法		
	估计标准误差		
平衡分析	平衡表的编制	简单平衡数	
		并列平均数	
		棋盘式平衡表	
平衡法		收付平衡法	
		投入产出法	

第二节 土地综合指标分析

- 一、土地总量指标
- 二、土地相对指标
- 三、土地平均指标
- 四、土地标志变异指标

基本概念

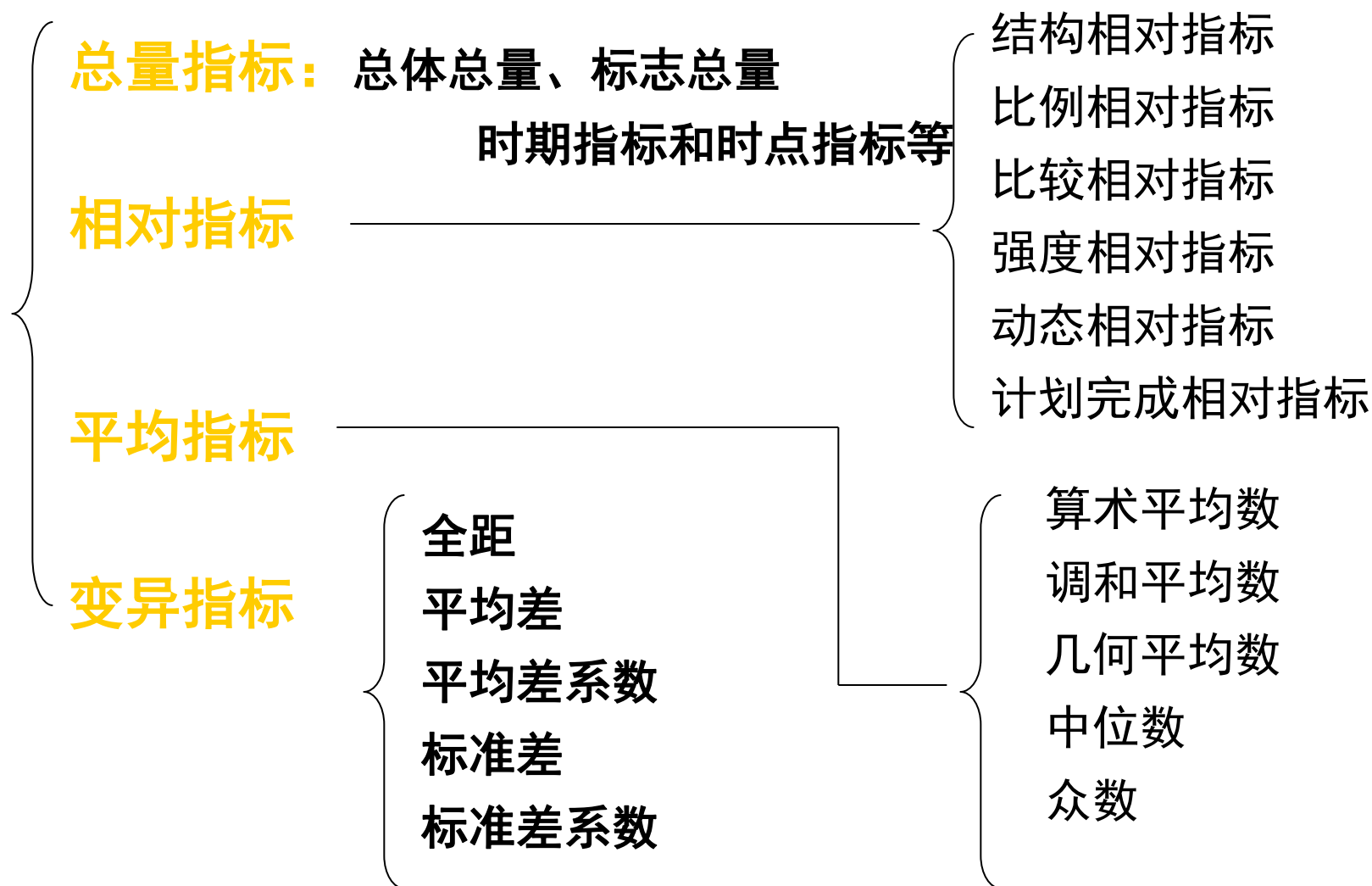
统计指标：是用来说明现象总体的特征，概括、分析和反映现象总体的数量特征和数量关系的综合性指标。

土地综合指标：土地统计调查得到的资料，经过整理、汇总后所形成的指标。主要包括土地总量指标、土地相对指标、土地平均指标和土地标志变异指标等。

土地综合指标分析：是在将土地指标划分类别的基础上，从个别到一般、从个性到共性进行概括的土地数量分析方法。

土地综合指标分析法：运用土地综合指标对土地各种现象和过程（土地的自然、经济状况及其变化过程）进行数量分析的方法。

土地综合指标的种类



一、土地总量指标

- 反映土地在一定时间地点条件下土地的总**规模****总水平**或**增减**变化，其表现形式为绝对数，因此也称为土地的**绝对指标**。如土地调查总面积、建设用地当年增加面积、农村、城镇土地面积等。
- 可以反映全国和地方各行政区、各单位不同地类土地的数量、分布及利用状况
- 是计算土地相对指标和平均指标，进行土地统计分析的基础
- 是编制土地利用计划、规划、检查其执行情况对土地实行宏观管理的依据

土地总量指标的种类

- 1. 按指标反映的具体内容划分
- **总体单位总量指标**：是用来反映总体中单位数的多少，说明总体本身规模大小的总量指标。
研究某地区居民粮食消费情况，该地区的居民人口数是总体单位总量指标。
- **总体标志总量指标**：是用来反映总体中标志值总和的总量指标。粮食消费总量是总体标志总量指标。

2. 按指标反映的时间状况划分

- **土地时期指标**：又称为土地流量指标，反映土地现象在**一定时期内**发展**变化过程**总量的指标。
- **土地时点指标**：又称为土地存量指标，是反映土地现象在某一**时刻**上状况的总量指标。
- **时期指标和时点指标的特点（区别）**：
 - a. 性质相同的时期指标的数值可以相加，时点指标相加则无意义。
 - b. 同类时期指标数值的大小与时期长短有直接关系，时点指标则没有这种关系。
 - c. 时期指标数值是经常登记取得，时点指标不是。区分时期指标和时点指标决定了统计处理与应用上的不同，在运用时期和时点指标时，注意同一指标若从不同的角度考虑则总量指标的性质也不同。如：年末耕地面积和年初耕地面积是时点指标，但年末耕地面积一年初耕地面积=年耕地净增数，就成为了时期指标。

- 3) 按指标采用的计量单位划分
- 价值量指标：是以货币为共同尺度，过渡到能彼此加总，具有广泛的综合和概括的性能。
- 实物量指标：是以实物单位计量的总量指标，其特点是直接反映产品的使用价值或现象的具体内容，是计算价值指标的基础。如人口总数，职工人数、粮食产量等
- 劳动量指标：是以劳动单位计量的总量指标，其特点是能将不同劳动者所用的时间汇总劳动消耗总量，为核算成本和计算劳动生产率等提供依据。如耕地用工

3. 土地总量指标的统计方法

□ 1) 直接计算法

- 对研究对象用直接的测量、点数、计数等方法，登记各单位的具体数值并加以汇总，得到总量指标。统计报表或普查中的总量资料

□ 2) 间接推算法

- 采用社会经济现象之间的平衡关系、因果关系、比例关系或利用非全面调查资料推算总量。
- **常用方法：**平衡关系推算法、因果关系推算法、比例关系推算法、插值推算法、抽样推断法等。
- **平衡关系推算法：**期初耕地总面积+期内增加耕地面积=期内减少耕地面积+期末耕地总面积
- **因果关系推算法：**土地现象内部各个因素相互联系，可表现为因果关系。如农作物收获量=播种面积×单位面积产量

- (3) 比例关系推算法：根据已知某一时期、某一地区或某一单位的某种指标与其相关指标的比例关系，来推算另一时期、另一地区或另一单位的指标；或根据部分资料利用其比例关系推算总体资料。
- (4) 插值估算法：是根据若干已知项目的统计资料，估算对应关系的未知项目数值，或根据变量数列和动态数列的若干已知对应数值，估算其数列中间所缺的未知对应数值。
 - A、内插法。根据平均发展速度，估算逐年缺少的数字资料。
 - B、线性插值法。根据已知两项有关的对应关系资料，估计第三项对应的未知资料。
 - C、拉格朗日插值法。当掌握的资料多于两项时，根据已知多项对应资料可用这种方法估计某项对应的未知数值。

比例关系推算法举例

××商场2005年销售额为500万元，商品损耗为4万元。
预计2006年的销售额达到600万元，考虑下年市场变化及企业管理优化等因素，计划每百元销售额商品损耗降低0.2元。
请据以上资料推算2006年商品的损耗。

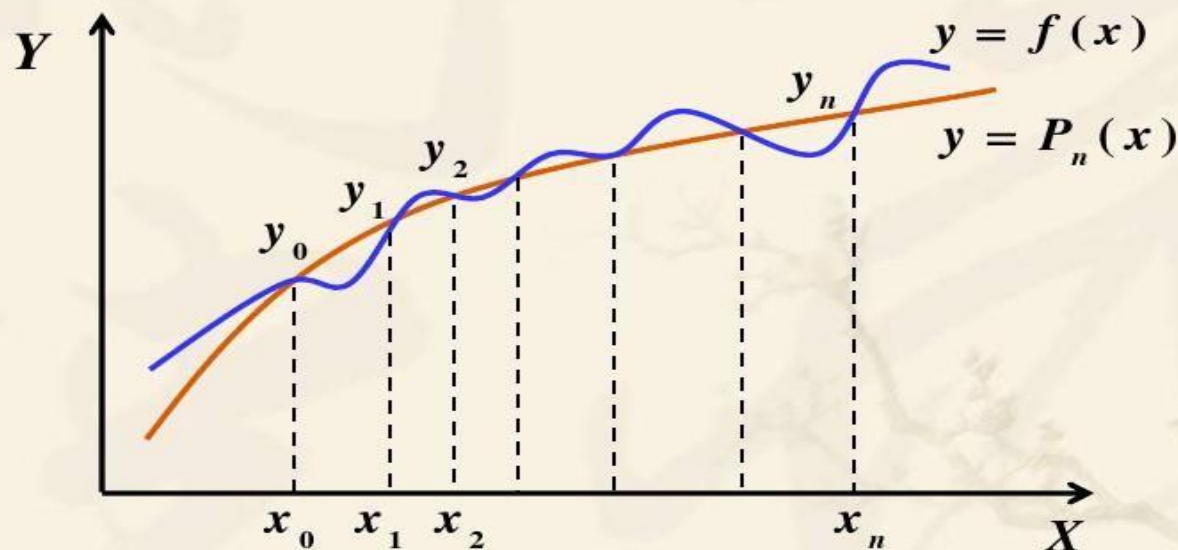
由于零售企业的商品损耗主要发生在销售环节，所以，报告期的商品损耗占销售额的百分比，即：

根据计划期的变化对报告期的百分比进行调整：计划期百分比=报告期百分比±计划期变动
 $=0.8\%-0.2\%=0.6\%$

所以，2006年的商品损耗=600×0.6%=3.6万元

即：通过比例推算法计算得2006年的商品损耗为3.6万元。

从几何上看, n 次插值多项式 $P_n(x)$, 是一条 n 次代数曲线, 它通过曲线 $y = f(x)$ 上的 $n+1$ 个点 $(x_i, y_i) (i = 0, 1, \dots, n)$.



许多实际问题中都用函数来表示某种内在联系或规律, 而不少函数都只能通过实验和观测来了解。如对实践中的某个物理量进行观测, 在若干个不同的地方得到相应的观测值, 拉格朗日插值法可以找到一个多项式, 其恰好在各个观测的点取到观测到的值。这样的多项式称为**拉格朗日(插值)多项式**。数学上来说, 拉格朗日插值法可以给出一个恰好穿过二维平面上若干个已知点的多项式函数。

4. 计算和应用总量指标应注意的问题

- （1）明确规定每项指标的含义和范围
- （2）注意现象的同质性（同质性是由事物的性质或用途决定的）
- （3）正确确定每项指标的计量单位

二、土地相对指标

1.概念

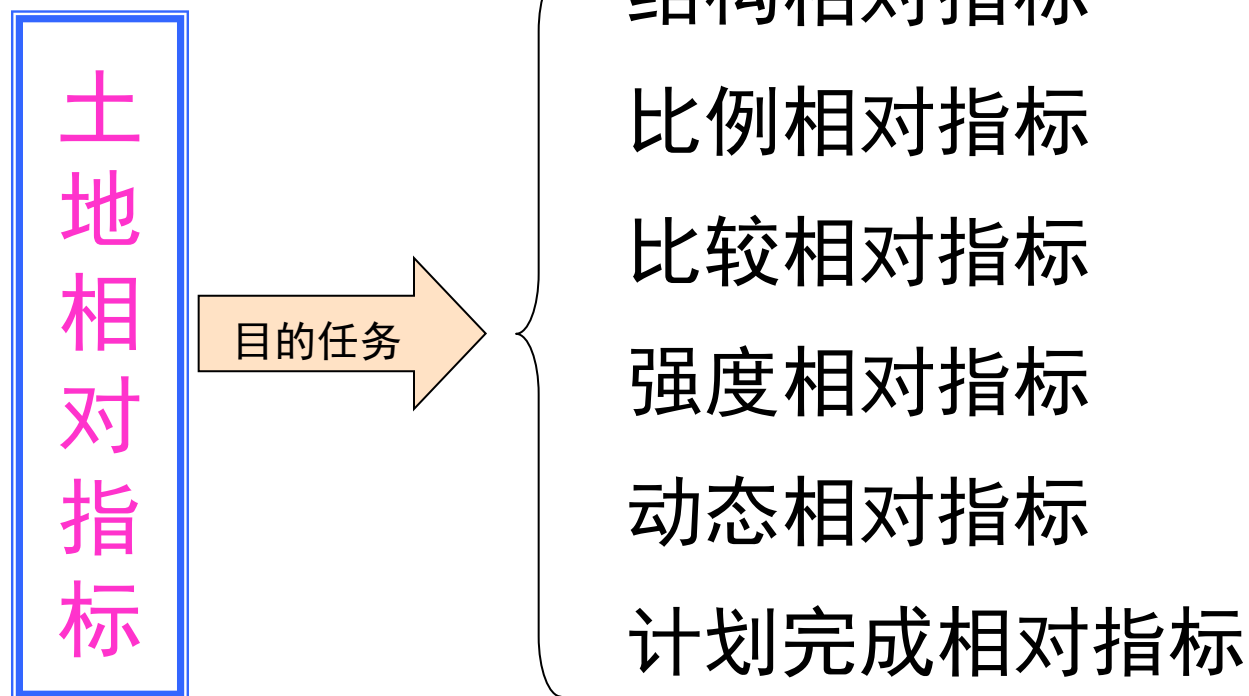
土地相对指标/相对数：两个与土地相联系的指标数值之比，反映土地各指标之间的数量联系程度。一般是没有计量单位，以相对数的形式存在，或为复名数（强度相对数如人均土地面积（公顷/人））。

2.作用

- （1）反映土地现象总体内部的结构、比例、发展状况或彼此之间的关系。
- （2）使一些不能直接对比的土地现象找到共同比较的基础。

$$\text{相对数} = \frac{\text{对比数}}{\text{基数}}$$

(二) 相对指标的种类及其计算



（二）相对指标的种类及其计算

1、结构相对指标

利用分组结果，计算各组成部分在总体中的比重，反映总体内部的构成。

$$\text{结构相对数} = \frac{\text{总体中某部分数值}}{\text{总体全部数值}} \times 100\%$$

- （1）反映总体的构成、性质及特征
- （2）通过不同时期结构相对数的变动，显示土地变化的发展规律。

年度 用地性质	1997	1998	1999
国家建设	17.7	52.6	29.0
集体建设	2.9	7.0	24.2
农村个人建房	79.4	40.4	46.8

2、比较相对指标

- 两个同类指标在不同空间的对比
- 说明土地利用或使用过程中的不平衡程度

$$\text{土地比较相对指标} = \frac{\text{某类现象的指标数值}}{\text{不同空间同类现象的指标数值}} \times 100\%$$

- 两个特点：
 - (1) 不可累加
 - (2) 分子、分母可以是绝对数，也可以是相对数或平均数，且分子分母可互换。

(3) 比例相对指标

反映总体内部各个组成部分之间的数量对比关系

$$\text{比例相对数} = \frac{\text{总体中某一部分数值}}{\text{总体中另一部分数值}}$$

（4）强度相对指标

指同一时期两个性质不同而又有联系的总量指标之比，用来说明某种现象的强度、密度、普遍程度或利用程度等。

$$\text{土地强度相对指标} = \frac{\text{某一总体中的指标数值}}{\text{另一有联系总体的指标数值}}$$

特点：

（1）既可用总量指标比较，也可用其它指标对比。若其数值与反映土地现象的强度、密度、普遍程度成正相关，即为正指标，反之为逆指标。

（2）它带有平均的意义，但又区别于平均指标。

（4）强度相对指标

- 应用：
- （1）反映土地总体的基本情况，如人均土地面积、人均耕地面积、人均住房面积等；
- （2）反映土地现象的强度、密度、普遍程度，如人口密度、路网密度、商业网点密度；
- （3）反映土地生产经营活动的条件、效果，如每公顷耕地拥有的拖拉机台数、成本利润率、耕地利用程度、土地产值率、土地利润率等。

(5) 动态相对指标

- 将两个同类指标在不同时间上进行对比，反映土地利用变化的程度和规律。

$$\text{土地动态相对指标} = \frac{\text{报告期水平}}{\text{基期水平}} \times 100\%$$

(6) 计划完成相对指标

- 以某现象某一段时间内实际完成的数字与计划规定的数字对比，来反映和监督计划执行情况的指标。

$$\text{计划执行情况相对指标} = \frac{\text{实际完成数}}{\text{计划规定数}} \times 100\%$$

- 计划指标的形式：
 - (1) 用总量指标下达计划，如非农业用地计划；
 - (2) 用相对指标来做计划，如土地开发计划等。

（6）计划完成相对指标

- （1）计划规定指标用总量指标即绝对数表示
- 由于计划的性质不同分为用**最高限量和最低限量**规定。如建设用地计划、耕地保有量。
- （2）计划规定指标用相对数表示。

$$\text{计划执行情况相对指标 (\%)} = \frac{\text{实际完成相对数 (\%)}}{\text{计划完成相对数 (\%)}} \times 100\%$$

- （3）长期计划执行情况相对指标考核办法

- 水平法

$$\text{计划执行情况相对指标 (\%)} = \frac{\text{计划期末实际达到的水平}}{\text{计划期规定水平}} \times 100\%$$

- 累计法

$$\text{计划执行情况相对指标 (\%)} = \frac{\text{计划期内累计完成数}}{\text{计划期规定累计数}} \times 100\%$$

(6) 计划完成相对指标

- (4) 长期计划**提前完成**时间的计算方法。
- **水平法**：计划期规定时间—完成计划时间=提前完成时间
- **累计法**：计划全期时间—从计划开始到全部完成时间=提前完成时间

（三）对比分析中应注意的几个问题

- 1. 注意指标的可比性：总体范围可比、计算方法和单位可比、统计指标所包含的内容可比。
- 2. 相对数与绝对数结合起来运用。
- 3. 要正确地选择作为比较标准的基期。
- 4. 各种相对数相结合使用。

四、土地平均指标

(一) 概念

是土地现象在一定时间、地点条件下的一般水平或代表性水平的综合指标。其数值一般表现为平均数。 平均耕地面积、平均地价

(二) 种类及其计算 算术、调和、几何，众数、中位数

1、算术平均数

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots + X_n}{n} = \frac{\sum X}{n} \quad \bar{X} = \frac{X_1 f_1 + X_2 f_2 + \cdots + X_n f_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum Xf}{\sum f}$$

经营面积分组	农户数 (f)	组中值 (公顷)
5 以下	5	2.5
5-10	20	7.5
10-15	34	12.5
15-20	65	17.5
20-25	42	22.5
25 以上	14	27.5

2、调和平均数

- 各变量值倒数平均数的倒数。是算术平均数的变形。

$$H = \frac{\sum m}{\sum \frac{m}{x}}$$

- H: 表示调和平均数, m代表各标志值的总量, x代表各标志值。

- 简单调和平均数

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} = \frac{n}{\sum \frac{1}{x}}$$

- 加权调和平均数

$$H = \frac{m_1 + m_2 + \cdots + m_n}{\frac{m_1}{x_1} + \frac{m_2}{x_2} + \cdots + \frac{m_n}{x_n}} = \frac{\sum m}{\sum \frac{m}{x}}$$

四块玉米单产和总产量资料，求平均单产

地块编号	单产（千克） x	总产量（千克） m	播种面积（公顷） m/x
1	720	5760	8
2	680	8160	12
3	660	10560	16
4	760	16720	22
合计	-	41200	58

$$H = \frac{\sum m}{\sum \frac{m}{x}} = \frac{5760 + 8160 + \cdots + 16720}{\frac{5760}{720} + \frac{8160}{680} + \cdots + \frac{16720}{760}} = \frac{41200}{58} = 710.34 \text{ (千克/公顷)}$$

粮食作物播种面积占 播种面积比重（%）	组中值（x）	粮食作物播种面积 （公顷 m）	m/x
50-60	0.55	16000	29090.9
60-70	0.65	36000	55384.6
70-80	0.75	30000	40000.0
80-90	0.85	12000	14117.6
合计		94000	138593.1

粮食作物播种面
积平均比重：

$$H = \frac{\sum m}{\sum \frac{m}{x}} = \frac{16000 + 36000 + 30000 + 12000}{\frac{16000}{0.55} + \frac{36000}{0.65} + \frac{30000}{0.75} + \frac{12000}{0.85}} = \frac{94000}{138593.1} = 67.84\%$$

3、几何平均数

几何平均数是n个变量值连乘积的n次方根。

(1) 简单几何平均数

$$G = \sqrt[n]{X_1 \cdot X_2 \cdot \dots \cdot X_n} = \sqrt[n]{\prod X}$$

$$\text{平均发展速度} = \sqrt[5]{152\% \cdot 91\% \cdot 84\% \cdot 117\% \cdot 128\%} = 111.7\%$$

(2) 加权几何平均数

$$G = \sqrt[\sum f]{\prod X^f}$$

年份	第 1-5 年	第 6-13 年	第 14-23 年	第 24-26 年	第 27-30 年
年利率	5%	8%	9%	14%	7%

平均利率：

$$G = \sqrt[\sum f]{\prod X^f} - 1 = \sqrt[30]{1.05^5 \cdot 1.08^8 \cdot 1.09^{10} \cdot 1.14^3 \cdot 1.07^4} - 1 = 1.08275 - 1 = 0.0828 \text{ 或 } 8.28\%$$

平均数应用

- 1. **算术平均数**：适用于计算普通意义下的绝对量的平均数，普通简单直观地表现中心位置。
- 2. **几何平均数**：计算相对量(如增长率)的平均数；当数据呈**倍数关系或不对称分布**时（增长率或生长率、动态发展速度），通常运用几何平均数。
- 3. **调和平均数**：社会经济统计中**资料不全或不可靠**等原因，许多问题不能直接计算算术平均数，而用调和平均数。适用于观测值是阶段性变异的资料。
- 4. **平方平均数**：应用在一些具有一定**体积**的物体的边长、直径、半径等资料上。平方平均数（quadratic mean），又名均方根（Root Mean Square），是指一组数据的平方的平均数的算术平方根。
- 正数 $x_i > 0 (i=1, 2, \dots, n)$ 的算术平均数 \bar{x} ，几何平均数 x_G ，调和平均数 x_H 间的关系 $\bar{x} \geq x_G \geq x_H$

(4) 众数

在土地现象总体中出现**次数最多**的标志值。

在**单项数列**资料条件：次数最多的组的变量值就是众数。

在**组距数列**条件下：

$$m_0 = X_0 + \frac{f_2 - f_1}{(f_2 - f_1) + (f_2 - f_3)} \times i$$

m_0 ——众数；

x_0 ——众数所在组下限

f_1 ——变量值小于众数组且与众数相邻的一组的次数；

f_2 ——众数组次数；

f_3 ——变量值大于众数组且与众数相邻的一组的次数；

i ——组距。

[例] 某地区农民家庭年人均收入资料

	A	B	C	D
1	按年人均收入分组 (元)	农民家庭数 (户)	向上累计频数	向下累计频数
2	X	F	S	S
3	1000~1200	240	240	3000
4	1200~1400	480	720	2760
5	1400~1600	1050	1770	2280
6	1600~1800	600	2370	1230
7	1800~2000	270	2640	630
8	2000~2200	210	2850	360
9	2200~2400	120	2970	150
10	2400~2600	30	3000	30
11	合 计	3000	—	—

$$M_o = 1400 + \frac{1050 - 480}{(1050 - 480) + (1050 - 600)} \times 200 = 1511.8(\text{元})$$

(5) 中位数

中位数是将总体各单位的标志值按大小顺序排列，处于数列中点位置的标志值。

首先计算中位数所在的项次，而后求出中位数。

对单项数列：

$$O_m = \frac{\sum f + 1}{2}$$

PH 值	5.5	6.0	6.5	7.0	7.5	合计
地块数 (f)	8	12	23	18	8	69

中位数所在项次为：

$$O_r = \frac{\sum f + 1}{2} = \frac{69 + 1}{2} = 35$$

(5) 中位数

对于组距数列：

$$m_e = x_1 + \frac{\frac{\sum f}{2} - S_{m-1}}{f_m} \times i$$

式中： m_e ——中位数；

x_1 ——中位数所在组下限；

S_{m-1} ——变量值比中位数小的各组累积次数；

f_m ——中位数所在组次数；

i ——组距。

经营面积分组	农户数 (f)	组中值 (公顷)
5 以下	5	2.5
5-10	20	7.5
10-15	34	12.5
15-20	65	17.5
20-25	42	22.5
25 以上	14	27.5

$$m_e = 15 + \frac{\frac{180}{2} - 59}{65} \times 5 = 17.38 \text{ (公顷)}$$

6. 应用平均指标应注意的问题

- 运用平均指标分析土地现象时，应注意：
 - （1）必须注意所研究社会经济现象的**同质性**。
 - 对同一性质总体单位构成的总体计算平均数
 - （2）必须注意**用组平均数补充说明总平均数**。
 - 当总体单位间存在各种因素影响时，需计算组平均数补充说明对总平均数的影响
 - （3）必须注意**应用分配数列补充说明平均数**。
 - 平均指标掩盖总体单位的差异，需把平均指标与具体数列结合补充说明平均数
 - （4）必须注意**一般与个别相结合**，把平均数和典型事例结合起来。
 - （5）平均指标要与变异指标结合运用

五、土地标志变异指标

- （一）概念和作用
- **标志变异**：总体各单位标志值与平均指标之间存在的差异。
- **标志变动度**即标志变异指标，又称离散程度或离中程度，指总体中各单位标志值差别大小的程度。
- **土地标志变异指标**
- （1）是衡量土地平均指标代表性大小的尺度。一般来说，土地标志变异指标愈大，平均指标的代表性愈小。反之，标志变异指标愈小，平均指标的代表性就愈大。
- （2）反映土地现象的均衡性、节奏性和稳定性。
- （3）进行抽样推断、相关分析、统计预测的依据之一。
- **测定标志变异的指标**：极差（全距）、平均差、方差、标准差（均方差）、变异系数等。

五、土地标志变异指标

1、极差/全距

$$R = X_{\max} - X_{\min}$$

2. 平均差

各个标志值与其算术平均数离差绝对值的平均数。

(未分组资料)

$$A \bullet D = \frac{\sum |X - \bar{X}|}{n}$$

(分组资料)

$$A \bullet D = \frac{\sum |X - \bar{X}|f}{\sum f}$$

3. 方差和标准差（均方差）

方差

□ 未分组资料

$$\sigma^2 = \frac{\sum (X - \bar{X})^2}{n}$$

□ 分组资料

$$\sigma^2 = \frac{\sum (X - \bar{X})^2 f}{\sum f}$$

标准差（均方差）

□ 未分组资料

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{n}}$$

□ 分组资料

$$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2 f}{\sum f}}$$

4. 标志变异系数

□ (1) 全距系数

$$V_R = \frac{X_{\max}}{X_{\min}}$$

□ (2) 平均差系数

$$V_A = \frac{A \bullet D}{\bar{X}} \times 100\%$$

□ (3) 标准差系数

$$V_\sigma = \frac{\sigma}{\bar{X}} \times 100\%$$

例题： 已知甲乙两地年降雨量资料 单位： 10mm

年份序号	甲地区	乙地区
1	108	106
2	175	138
3	79	125
4	76	103
5	142	128
6	99	132
7	85	118
8	100	117
9	143	120
10	139	114
11	79	130
12	190	104
13	92	144
14	115	108
15	89	152
16	168	119
17	229	135
18	145	155
19	160	134
20	85	116
合计	2498	2498

□ 则两地的标志变异指标：

□ (1) 全距

□ $R_{\text{甲}} = X_{\text{max}} - X_{\text{min}} = 229 - 76 = 153 \text{ (10mm)} = 1530 \text{ mm}$

□ $R_{\text{乙}} = X_{\text{max}} - X_{\text{min}} = 155 - 103 = 52 \text{ (10mm)} = 520 \text{ mm}$

□ (2) 平均差

□ 甲地 $\bar{x} = 2498 / 20 = 124.9 \text{ (10mm)} = 1249 \text{ mm}$

□ 平均差 =

$$\begin{aligned}
 A \cdot D_{\text{甲}} &= \frac{\sum |x - \bar{x}|}{n} \\
 &= \frac{|108 - 124.9| + |175 - 124.9| + \dots + |85 - 124.9|}{20} \\
 &= 35.69 \text{ (10mm)} = 356.9 \text{ mm}
 \end{aligned}$$

例题： 已知甲乙两地年降雨量资料 单位： 10mm

□ (2) 平均差

$$\text{乙地区 } \bar{x}_z = 2498 / 20 = 124.9 \text{ (10mm)} = 1249 \text{ mm}$$

平均差 =

$$\begin{aligned} A \cdot D_z &= \frac{\sum |x - \bar{x}|}{n} \\ &= \frac{|106 - 124.9| + |98 - 124.9| + \dots + |116 - 124.9|}{20} \\ &= 12.4 \text{ (10mm)} = 124 \text{ mm} \end{aligned}$$

年份序号	甲地区	乙地区
1	108	106
2	175	138
3	79	125
4	76	103
5	142	128
6	99	132
7	85	118
8	100	117
9	143	120
10	139	114
11	79	130
12	190	104
13	92	144
14	115	108
15	89	152
16	168	119
17	229	135
18	145	155
19	160	134
20	85	116
合计	2498	2498

例题： 已知甲乙两地年降雨量资料 单位： 10mm

□ (3) 标准差

$$\text{甲地区 } \sigma_{\text{甲}} = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

$$= \sqrt{\frac{(108 - 124.9)^2 + \dots + (85 - 124.9)^2}{20}}$$

$$= 42.21 \text{ (10mm)}$$

$$= 422.1 \text{ mm}$$

$$\text{乙地区 } \sigma_{\text{乙}} = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

$$= \sqrt{\frac{(106 - 124.9)^2 + \dots + (116 - 124.9)^2}{20}}$$

$$= 14.79 \text{ (10mm)}$$

$$= 147.9 \text{ mm}$$

年份序号	甲地区	乙地区
1	108	106
2	175	138
3	79	125
4	76	103
5	142	128
6	99	132
7	85	118
8	100	117
9	143	120
10	139	114
11	79	130
12	190	104
13	92	144
14	115	108
15	89	152
16	168	119
17	229	135
18	145	155
19	160	134
20	85	116
合计	2498	2498

例题： 已知甲乙两地年降雨量资料 单位： 10mm

□ （4）标志变异系数

甲地区

$$\text{全距系数 } V_R = \frac{R_{\max}}{R_{\min}} = 229/76 = 3.01$$

$$\text{平均差系数 } V_A = \frac{A \cdot D}{\bar{X}} \times 100\% = \frac{355.9}{1249} \times 100\% = 28.58\%$$

$$\text{标准差系数 } V_\sigma = \frac{\sigma}{\bar{X}} \times 100\% = \frac{422.1}{1249} \times 100\% = 33.80\%$$

乙地区

$$\text{全距系数 } V_R = \frac{R_{\max}}{R_{\min}} = 155/103 = 1.50$$

$$\text{平均差系数 } V_A = \frac{A \cdot D}{\bar{X}} \times 100\% = \frac{124}{1249} \times 100\% = 9.93\%$$

$$\text{标准差系数 } V_\sigma = \frac{\sigma}{\bar{X}} \times 100\% = \frac{147.9}{1249} \times 100\% = 11.84\%$$

六、分布偏态与峰度的测度

- 集中趋势和离散程度是数据分布的两个重要特征，但要全面了解数据分布的特点，还需要知道数据分布的形状是否对称、偏斜的程度以及分布的扁平程度等。偏态和峰度就是对这些分布特征的进一步描述。

一、偏态及其测度

- 偏态是对分布偏斜方向及程度的测度。
- 利用众数、中位数和均值之间的关系就可以判断分布是左偏还是右偏。显然，判别偏态的方向并不困难，但要测度偏斜的程度就需要计算偏态系数。偏态系数的计算方法有很多，这里仅介绍其中比较常用的一种。

六、分布偏态与峰度的测度

偏态系数是对分布偏斜程度的测度，其计算公式为

$$a = \frac{\sum_{i=1}^K (X_i - \bar{X})^3 F_i}{N\sigma^3}$$

式中： a 表示偏态系数， σ^3 是标准差的三次方。

- 从公式可以看到，偏态系数是根据离差三次方的平均数再除以标准差的三次方。当分布对称时，离差三次方后正负离差可以相互抵消，因而 a 的分子等于0，则 $a=0$ ；当分布不对称时，正负离差不能抵消，就形成了正或负的偏态系数 a 。当 a 为正值时，表示正偏离差值较大，可以判断为正偏或右偏；反之，当 a 为负值时，表示负离差数值较大，可判断为负偏或左偏。在计算 a 时，将离差三次方的平均数除以 σ^3 是将偏态系数转化为相对数， a 的绝对值越大，表示偏斜的程度就越大。

六、分布偏态与峰度的测度

例：已知1997年我国农村居民家庭按纯收入分组的有关数据如下表。试计算偏态系数。

农村居民家庭按纯收入分组的数据

1997年农村居民家庭纯收入数据	
按纯收入分组（元）	户数比重（%）
500以下	2.28
500~1000	12.45
1000~1500	20.35
1500~2000	19.52
2000~2500	14.93
2500~3000	10.35
3000~3500	6.56
3500~4000	4.13
4000~4500	2.68
4500~5000	1.81
5000以上	4.94

$$a_3 = \frac{\sum_{i=1}^K (X_i - \bar{X})^3 F_i}{N\sigma^3} = \frac{\sum_{i=1}^K (X_i - 21.429)^3 F_i}{1 \times 12.089^3} = \frac{1689.25}{1766.7339} = 0.956$$

由计算结果可以看出，偏态系数为正值，而且数值较大，说明农村居民家庭纯收入的分布为右偏分布，即收入较少的家庭占多数，而收入高的家庭占少数，而且偏斜程度较大。

六、分布偏态与峰度的测度

二、峰度及其测度

峰度是分布集中趋势高峰的形态。它通常是与正态分布相比较来说的，在归化到同一方差时，若分布的形状比正态分布更瘦更高，则称为尖峰，若比正态分布更矮更胖，则称为平峰分布。

峰度系数是离差四次方的平均数，再除以标准差的四次方，其计算公式为式中：

$$\beta = \frac{\sum_{i=1}^K (X_i - \bar{X})^4 F_i}{N\sigma^4}$$

式中： β 表示峰度系数， σ^4 是标准差的四次方。

六、分布偏态与峰度的测度

$$\beta = \frac{\sum_{i=1}^K (X_i - \bar{X})^4 F_i}{N\sigma^4}$$

- 公式中将离差的四次方除以 σ^4 ，是为了将峰度系数转化成相对数。用峰度系数说明分布的尖峰和扁平程度，是通过与正态分布的峰度系数进行比较而言的。
- 由于正态分布的峰度系数为3，当 $\beta > 3$ 时为尖峰分布，当 $\beta < 3$ 时为扁平分布。

六、分布偏态与峰度的测度

例：根据例中的数据，计算农村居民家庭纯收入分布的峰度系数。

解：根据上表的计算结果，代入公式得

$$\beta = \frac{\sum_{i=1}^K (X_i - \bar{X})^4 F_i}{N\sigma^4} = \frac{72521.25}{1 \times 12.089^4} = 3.4$$

由于 $\beta = 3.4 > 3$ ，说明我国农村居民家庭纯收入的分布为尖峰分布，说明低收入家庭占有较大的比重。

1997年农村居民家庭纯收入数据

按纯收入分组（元）	户数比重（%）
500以下	2.28
500~1000	12.45
1000~1500	20.35
1500~2000	19.52
2000~2500	14.93
2500~3000	10.35
3000~3500	6.56
3500~4000	4.13
4000~4500	2.68
4500~5000	1.81
5000以上	4.94

七、土地综合指标分析内容

1. 土地总面积及各类用途占地面积；
2. 反映土地的开发程度和利用程度；
垦殖系数、森林覆盖率、土地利用率、复种指数、建筑密度等
3. 反映土地有关指标的计划完成程度；
土地整治计划、建设用地计划
4. 反映土地内部构成； 各用途占总面积比重
5. 反映土地集约经营程度； 机械化、水利化程度
6. 反映土地质量状况； 各等级占比
7. 反映土地生产率水平。 单位面积产值、产量、收入等