

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

第四节 土地动态分析

- 一、土地动态分析含义与内容
- 二、土地动态数列
- 三、土地动态分析指标
- 四、土地现象长期变动趋势的分析

一、土地动态分析

含义：运用一系列的动态分析指标对土地现象进行分析，以反映土地现象发展过程及其规律。

土地动态分析内容	发展水平	通过计算发展水平、增长量、平均发展水平等指标，反映土地现象在各个不同时期达到的规模水平或在某一时期内发展的一般水平。
	变动速度	通过计算发展速度、增长速度、平均发展速度、平均增长速度等指标，反映土地现象变动方向和变动程度。
	变动长期趋势	通过对原有的土地动态数列加工整理，采用时期扩大法、移动平均法、最小平方法等尽可能地排除各种偶然因素对土地现象变动趋势的影响，把土地现象的变动总趋势明显地表现出来，为土地统计预测提供依据。

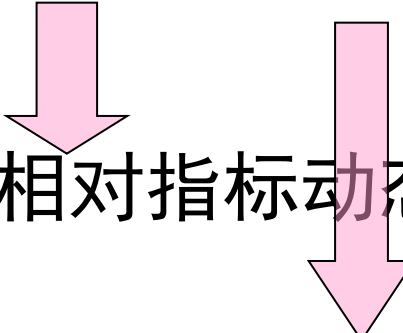
二、土地动态数列

- 按时间顺序排列的反映土地现象特征的统计指标的数列。

土地统计指标	1991	1992	1993	1994	1995	1996	1997	1998
年末耕地面积（公顷）	92000	94000	98000	88000	92000	90000	89000	86000
年内耕地增减总数（公顷）	1680	2000	4000	-10000	4000	-2000	-1000	-3000
人均耕地（公顷）	1.78	1.82	1.93	1.72	1.78	1.76	1.75	1.74
垦殖系数（%）	39.76	41.21	42.36	37.76	39.76	38.14	37.82	35.86

- 土地动态数列的两个基本要素：一是时间；二是指标值（现象发展水平）
- 作用：1) 确定土地现象在不同时期的发展状况，从土地现象的量变过程中，反映其发展变化过程的方向程度和趋势，研究其质量变化过程的阶段性和规律性
- 2) 把有关动态数列结合起来分析反映现象间客观联系和关系
- 3) 揭示土地现象发展长期趋势，预测未来发展
- 4) 作为整理、积累历史资料的方法和手段。

（二）土地动态数列的种类

- 按照指标的不同可以把土地动态数列分为：
 - 土地总量指标动态数列
 - 土地相对指标动态数列
 - 土地平均指标动态数列
- 

（二）土地动态数列的种类

□ 1. 土地总量指标动态数列

- 是指土地总量指标在不同时间上的数量表现按时间先后顺序排列而形成的动态数列。
- （1）**时期数列**。凡数列中的每一项指标，都是反映某种现象在一段时期内的发展过程总量时，这种数列就称为时期数列。
- （2）**时点数列**。凡数列的每一个指标都是反映现象在某一时点上的状态或达到的水平时，这种数列称为时点数列。

（二）土地动态数列的种类

□ 2. 土地相对指标动态数列

- 把一系列同类土地相对指标按时间先后顺序排列而成的动态数列。
- 反映土地现象数量对比关系的发展变化。

□ 3. 土地平均指标动态数列

- 把一系列土地平均指标按时间先后顺序排列而成的动态数列。
- 反映土地现象代表性（一般性）水平的发展变化。

（三）土地动态数列的编制原则

- 1. 时间上的可比性。
- 2. 总体空间范围的可比性。
- 3. 指标经济内容上的可比性。
- 4. 计算方法、价格、单位上的可比性。

三、土地动态分析指标

□ （一）绝对数分析指标

□ 1、发展水平/发展量

□ 指动态数列（时间数列）中各个指标的数值，反映现象在各个不同时期达到的规模水平。

□ （1）最初发展水平： a_0

□ （2）最末发展水平： a_n

□ （3）中间发展水平： $a_1, a_2, a_3, \dots, a_{n-1}$ 。

□ （4）基期发展水平

□ （5）报告期发展水平

发展水平

数列中的具体指标数值为**发展水平**，可以是绝对数、相对数或平均数。可以分为：最初水平、最末水平、中间水平、基期水平、报告期水平等。

我国人均GDP发展情况 单位：元/人

年份	1997	1998	1999	2000	2001	2002
人均GDP	6054	6307	6547	7084	7543	7972
符号	a0	a1	a2	a3	a4	a5



$$\text{动态相对数} = \frac{\text{分子：报告期水平}}{\text{分母：基期水平}}$$

比如： $\frac{a_4}{a_2}$ 或 $\frac{a_5}{a_1}$

2、增长量

- 报告期与基期水平之差。增长量=报告期水平-基期水平= $a_n - a_0$
- 1) 逐期增长量：以报告期的前一期作为基期计算的增长量，说明土地现象在两个相邻时期或时点上发展水平增减的数量。
- 逐期增长量= $a_1 - a_0, a_2 - a_1, \dots, a_n - a_{n-1}$
- 2) 累计增长量：报告期水平与某一固定基期水平的差额，说明事物某一时期内的总增长量。
- 累计增长量= $a_1 - a_0, a_2 - a_0, \dots, a_n - a_0$
- 3) 关系
- ① 累计增长量等于各个逐期增长量之和。
- $a_n - a_0 = (a_1 - a_0) + (a_2 - a_1) + (a_3 - a_2) + \dots + (a_n - a_{n-1})$
- ② 两个累计增长量之差等于相应的逐期增长量。
- $(a_n - a_0) - (a_{n-1} - a_0) = a_n - a_{n-1}$

某地非法占地情况

年份	1992	1993	1994	1995	1996	1997	1998
占地面积（公顷）	26	31	40	46	52	61	79
逐期增长量（公顷）	-	5	9	6	6	9	18
累计增长量（公顷）	-	5	14	20	26	35	53

- 非法占地的各逐期增长量之和为：
- $5+9+6+6+9+18=53$ （公顷）
- 结果与1998年与1992年相比的累计增长量相等。
- **（4）年距增长量**
- 为消除季节变动的影响，增加可比性，计算本期发展水平与上年同期水平的增减数量，称为年距增长量。
- $\text{年距增长量} = \text{本期发展水平} - \text{上年同期发展水平}$

（二）动态数列的相对数分析指标

□ 1、发展速度

- 两个不同时期发展水平指标对比的结果。说明报告期水平是基期水平的百分之几或若干倍。

$$\text{发展速度} = \frac{\text{报告期发展水平}}{\text{基期发展水平}}$$

- 1) 定基发展速度：报告期水平与某一固定基期水平之比，说明土地现象在一个较长时期内的变动程度，也叫总速度。

$$\frac{a_1}{a_0}, \frac{a_2}{a_0}, \frac{a_3}{a_0}, \dots, \frac{a_n}{a_0}$$

- 2) 环比发展速度：报告期水平与前期水平之比，表明报告期水平对比前期水平的逐期发展变化的情况。

$$\frac{a_1}{a_0}, \frac{a_2}{a_1}, \frac{a_3}{a_2}, \dots, \frac{a_n}{a_{n-1}}$$

(3) 定基发展速度与环比发展速度的关系

- ①定基发展速度等于相应各时期环比发展速度的连乘积
- ②相邻时期定基发展速度相除之商等于相应环比发展速度

$$\frac{a_1}{a_0} \times \frac{a_2}{a_1} \times \frac{a_3}{a_2} \times \dots \times \frac{a_n}{a_{n-1}} = \frac{a_n}{a_0}$$

$$\frac{a_n}{a_0} \div \frac{a_{n-1}}{a_0} = \frac{a_n}{a_{n-1}}$$

年份	1990	1991	1992	1993	1994	1995	1996
国民生产总值 (万元)	8557.6	9696.3	11301.0	14068.2	17695.3	20236.3	24036.2
环比发展速度 (%)	-	113.3	116.6	124.5	125.7	114.4	118.8
定基发展速度 (%)	100	113.3	132.1	164.4	206.8	236.5	280.9
环比增长速度 (%)	-	13.3	16.6	24.5	25.9	14.4	18.8
定基增长速度 (%)	-	13.3	32.1	64.4	106.8	136.5	180.9

- 1990-1996年环比发展速度的连乘积等于定基发展速度：
 $113.3\% \times 116.6\% \times 124.5\% \times 125.7\% \times 114.4\% \times 118.8\% = 280.9\%$
- 1996与1995年两定基发展速度之商等于相应环比发展速度
 $280.9\% / 236.5\% = 118.8\%$

2、增长速度

- 增长量与基期水平之比，说明报告期发展水平比基期发展水平增加了若干倍或百分之多少。

$$\text{增长速度} = \frac{\text{增长量}}{\text{基期发展水平}} = \frac{\text{报告期发展水平} - \text{基期发展水平}}{\text{基期发展水平}}$$

$$= \frac{\text{报告期发展水平}}{\text{基期发展水平}} - 1 \quad = \text{发展速度} - 1$$

- 增长速度=发展速度—1（或100%）

2、增长速度

□ (1) 环比增长速度

- 逐期增长量与前期发展水平之比，表明土地现象逐期增减程度。

$$\text{环比增长速度} = \frac{\text{逐期增长量}}{\text{前期发展水平}} = \text{环比发展速度} - 1$$

$$\frac{a_n - a_{n-1}}{a_{n-1}} = \frac{a_n}{a_{n-1}} - 1$$

□ 2) 定基增长速度

- 累计增长量与某一固定基期的发展水平之比，表明土地现象在一定时期内总的增减程度。

$$\text{定基增长速度} = \frac{\text{累计增长量}}{\text{固定基期发展水平}} = \text{定基发展速度} - 1$$

$$\frac{a_n - a_0}{a_0} = \frac{a_n}{a_0} - 1$$

3、增降1%绝对值

$$\text{增长1\%绝对值} = \frac{\text{报告期的逐期增长量}}{\text{报告期的环比增长速度} \times 100}$$

$$\square = \frac{\text{逐期增长量}}{\frac{\text{逐期增长量}}{\text{前一期水平}} \times 100} = \text{前一期水平} / 100$$

$$\frac{\frac{a_n - a_{n-1}}{\frac{a_n - a_{n-1}}{a_{n-1}} \times 100}}{a_{n-1}} = \frac{a_{n-1}}{100}$$

□ 例，某单位1995年生产总值3000万元，1996年为4000万元，则

$$\square \text{ 增长1\%绝对值} = \frac{4000 - 3000}{\frac{4000 - 3000}{3000} \times 100} = 3000 / 100 = 30 \text{ 万元}$$

（三）动态数列的平均分析指标

1、序时平均指标的概念和作用

序时平均指标又称序时平均数或动态平均数，是将整个时间数列作为一个整体，反映这个整体的一般水平。

它是动态数列中各项指标数值的平均数，用以反映土地现象在某一时期内的平均变动或发展变化的一般水平。包括：平均发展水平、平均增长量、平均发展速度和平均增长速度。它与一般算术平均数的异同：

相同点：都是将个别差异抽象化，概括反映一般水平。

不同点	序时平均数	算术平均数
(1) 说明的问题不同	从动态上说明某一事物在某一时期内的一般水平	静态、同一事物总体不同单位、同一时间
(2) 计算的依据不同	根据动态数列（时间数列）计算	变量数列
(3) 计算的方法不同	是事物在不同时间上的数量差异	总体各单位某一数量标志在同一时间上的数量差异

序时平均指标的作用

- (1) 可反映现象在某一时期内发展变化的一般水平
- (2) 可消除现象在短时期内的偶然波动，便于广泛对比，观察现象的发展趋势
- (3) 可把时间长短不等的指标变为可比。产品日产量不可比，日平均产量可比
- (4) 具有广泛的对比性。如不同地区某一时期现象发展的一般水平比较

2、平均发展水平（序时平均数/动态平均数）

对不同时期的发展水平加以平均得到的平均数。说明现象在某一段时间内发展的一般水平

1) 由绝对数动态数列计算平均发展水平

(1) 由时期数列计算平均发展水平

$$\bar{a} = \frac{a_1 + a_2 + a_3 + \cdots + a_n}{n} = \frac{\sum a}{n} \quad \text{例}$$

(2) 由时点数列计算平均发展水平

① 由连续时点数列计算平均发展水平

连续时点数列指的是总期数小于1个月，以天表示。

连续变动时点数列
间隔相等

$$\bar{a} = \frac{a_1 + a_2 + a_3 + \cdots + a_n}{n} = \frac{\sum a}{n} \quad \text{例}$$

非连续变动时点数列
间隔不等

$$\bar{a} = \frac{a_1 f_1 + a_2 f_2 + \cdots + a_n f_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum af}{\sum f} \quad \text{例}$$

② 由间断时点数列计算序时平均数

间断时点数列一般以月为期数单位。涉及概念：

期初：每期起始时点，如月初、季初、年初。

期末：每期结束时点，如月末、季末、年末。

显然某一期初数值等于上期期末数值，如4月初等于三月末1月初等于上年末。每期的平均数为： $(\text{期初} + \text{期末}) / 2$ 。间断时点时点数列的序时平均数其实是各期平均数的平均。

由间隔相等的间断时点数列计算序时平均数

$$\begin{aligned}\bar{a} &= \frac{\frac{a_1 + a_2}{2} + \frac{a_2 + a_3}{2} + \frac{a_3 + a_4}{2} + \cdots + \frac{a_{n-1} + a_n}{2}}{n-1} \\ &= \frac{\frac{a_1}{2} + a_2 + \cdots + a_{n-1} + \frac{a_n}{2}}{n-1}\end{aligned}$$

例

“首末折半法”

由间隔不等的间断时点数列计算序时平均数

$$\bar{a} = \frac{\frac{a_1 + a_2}{2} f_1 + \frac{a_2 + a_3}{2} f_2 + \cdots + \frac{a_{n-1} + a_n}{2} f_{n-1}}{\sum_{i=1}^{n-1} f}$$

例

“两两平均法”

2) 由相对数或平均数动态数列计算平均水平

基本公式: $\bar{c} = \frac{\bar{a}}{\bar{b}}$, 其中分子分母均为序时平均数

① a、b都为时期数列

$$\bar{c} = \frac{\bar{a}}{\bar{b}} = \frac{\frac{\sum a}{n}}{\frac{\sum b}{n}} = \frac{\sum a}{\sum b}$$

例

② a、b都为时点数列

$$\bar{c} = \frac{\bar{a}}{\bar{b}} = \frac{\frac{\frac{a_1}{2} + a_2 + a_3 + \cdots + \frac{a_n}{2}}{n}}{\frac{\frac{b_1}{2} + b_2 + b_3 + \cdots + \frac{b_n}{2}}{n}}$$

例

③ a为时期数列， b为时点数列

$$\bar{c} = \frac{\bar{a}}{b} = \frac{\sum a}{\frac{b_1}{2} + b_2 + b_3 + \cdots + b_{n-1} + \frac{b_n}{2}}$$

例

当a未知时:

$$\bar{c} = \frac{\sum bc}{\sum b}$$

当b未知时:

$$\bar{c} = \frac{\sum a}{\sum \frac{a}{c}}$$

例

实际计算时，只要先计算分子的序时平均值，再计算分母的序时平均数，然后两者相除即得到结果。

3、平均增长量

平均增长量是动态数列中各指标逐期增长量的序时平均数

$$\text{平均增长量} = \frac{\text{逐期增长量之和}}{\text{逐期增长量个数}} = \frac{\text{累计增长量}}{\text{动态数列项数} - 1}$$

4、平均发展速度和平均增长速度

- **平均发展速度**：各时期环比发展速度的序时平均数，说明现象在较长时期中逐年平均发展变化的程度。
- **平均增长速度**：各个环比增长速度的序时平均数，说明土地现象在一个较长时期内逐年平均增减变化的程度。
- **关系**：平均增长速度=平均发展速度-1（或100%）
- **平均发展速度**：
$$\bar{X} = \sqrt[n]{\frac{a_1}{a_0} \cdot \frac{a_2}{a_1} \cdot \frac{a_3}{a_2} \cdot \dots \cdot \frac{a_n}{a_{n-1}}} = \sqrt[n]{\frac{a_n}{a_0}}$$
- 我国国民收入1986年为7859亿元，1992年为19845亿元，则：
- 平均发展速度= $\bar{X} = \sqrt[6]{\frac{19845}{7859}} = 116.7\%$
- 平均增长速度=平均发展速度-100%=116.7%-100%=16.7%

四、土地现象长期变动趋势的分析

- **基本因素**：对土地现象各期的发展水平起着普遍的、长期的、决定性作用的因素。
- **偶然因素**：对土地现象的发展只起局部的、暂时的、非决定性作用的因素。
- **土地变动的趋势分析**：对动态数列进行加工、分析，尽量排除各种偶然因素对土地变动趋势的影响，把土地变动的总趋势呈现出来。
- 土地现象长期变动趋势分析方法：时距扩大法、移动平均法、最小平方法等。

（一）时期（时距）扩大法

- 又称间隔扩大法、时期合并法，是将原来时间间隔较小的动态数列加工为时间间隔较大的动态数列，以便消除前种情况受偶然因素影响，从而显示土地现象发展总趋势的方法。
- 1. 总数扩大法
- 简便易行，但只适用于时期数列，不适用于时点数列。时距必须相等，否则不可比。
- 2. 平均数扩大法
- 既适用于时期数列，又适用于时点数列，还能使总量时距不等情况下资料可比较，较准确地反映现象的长期变动趋势。

我国历年某农产品产量 单位：万吨

年份	产量	年份	产量	年份	产量	年份	产量	年份	产量
1969	196.9	1976	209.8	1983	195.8	1990	220.7	1997	354.0
1970	170.9	1977	233.7	1984	256.2	1991	270.7	1998	419.0
1971	106.3	1978	235.4	1985	246.1	1992	296.8		
1972	80.0	1979	235.5	1986	238.1	1993	359.8		
1973	75.0	1980	208.0	1987	205.6	1994	463.7		
1974	120.0	1981	227.7	1988	204.9	1995	625.9		
1975	166.3	1982	210.3	1989	216.7	1996	415.0		

□ 时距扩大法：

年份	1969-1973	1974-1978	1979-1983	1984-1988	1989-1993	1994-1998
产量	629.1	965.2	1077.3	1150.9	1364.7	2277.6

□ 平均数扩大法

年份	1969-1973	1974-1978	1979-1983	1984-1988	1989-1993	1994-1998
产量	125.82	193.04	215.50	230.18	272.94	455.52

□ （二）最小平方法/最小二乘法

- 根据原动态数列总的变化趋势配合一条最为理想的趋势线来揭示土地现象发展的趋势。趋势线模型有直线型和曲线型等。

□ （三）移动平均数

- 将原动态数列时距扩大，并逐项依次移动计算序时平均数，再将结果编成一个新的动态数列。把原动态数列中的偶然因素引起的变动削弱或消除，从而明显地反映出现象发展趋势，并基本保留与原动态数列所对应的项目，使之与实际水平更为接近。

第五节 土地动态平衡分析

- 一、土地平衡分析的概念和作用
- 二、土地平衡分析的特点
- 三、土地平衡分析表
- 四、土地平衡分析的原则
- 五、土地平衡分析方法

一、土地平衡分析的概念和作用

- **土地平衡分析：**根据土地要素的变化状况编制平衡表，反映土地这个自然与社会综合体的运动状态、变化方向与程度，进而为达到合理用地结构来满足国民经济发展需要提供依据。
- 年初耕地+年内增加耕地-年内减少耕地=年末耕地
- 增加：开荒+围垦+废弃地利用+其他
- 减少：居民点建设+交通建设+水利建设+农业结构+灾毁

二、土地平衡分析的特点

- 1、 反映土地内部的地类构成、等级分布、权属性质等因素间的数量对等关系；
- 2、 通过有联系的土地统计指标数值之间的数量对等关系，反映土地现象间的平衡关系。
- 3、 运用若干有联系的会引起土地变化的要素，以及地类变化指标之间的数量关系来分析其存在的平衡比例关系；
- 4、 进行平衡分析必须以全面资料为依据，只有这样才能正确反映数量对等关系。
- 通过土地平衡分析，还可以反映地区之间、部门之间、局部之间、局部与全局之间的平衡关系。

三、土地平衡分析表

- **平衡表**是将土地现象各指标按增减对等的顺序列成表格形式，它可以把客观存在的错综复杂的土地现象之间的联系，用平衡表的格式，比较简明、完整、全面地反映出来，以便对各土地现象间的比例关系、相互关系和平衡图等进行对比分析。
- **平衡图**是对平衡表的补充，只有将两者结合才能深入直观地进行土地平衡分析。
- 平衡表的形式：增减平衡表、并列式平衡表、棋盘式平衡表。

（一）增减式平衡表

- 增减式平衡表是根据土地经济活动中有增必有减，有减必有增，增减相等的原则来编制的。适用于反映一个部门、单位、地区或全国某一要素的平衡关系。也称为单项平衡分析。

- （1）增减左右排列 **期初实有+本期增加=本期减少+期末实有**

项目	面积	项目	面积
1.期初耕地	1531	3.期内减少耕地	66
2.期内新增耕地	42	4.期末耕地	1507
合计	1573	合计	1573

- （2）增减上下排列

项 目	本年面积	上年面积	比去年增长 (%)
一、增长项目合计	1573	1550	-1.46
1.年初耕地	1531	1510	-1.37
2.年内增加	42	40	-4.76
二、减少项目合计	1573	1550	-1.46
3.年内减少耕地	66	50	-24.24
4.年末耕地	1507	1500	-4.64

（二）并列式平衡表

- 并列式平衡表是把若干个增减平衡表并列在一张平衡表中。本表适用于反映不同部门对同一项目或是不同部门的项目增减平衡关系。表中的每一项目都是一个增减式平衡表。

农资名称	单位	年初结存 (1)	本年收入		本年支出		年末结存 (6)
			本地生产 (2)	外地调入 (3)	本地使用 (4)	调入外地 (5)	
化肥	千克						
农药	千克						
地膜	米						

（三）棋盘式平衡表

- 是一种特殊设计的平衡表，由于该表横行与纵行交叉如同棋盘，因而得名。它常用来反映各部门、各单位之间的经济联系与平衡关系。

		调整后土地面积				合计
		A 村	B 村	C 村	D 村	
调整前土地 面积	A 村	1540	1120	1380	1450	5490
	B 村	195	1310	1430	1760	5595
	C 村	1380	1220	1380	1650	5630
	D 村	1410	1710	1460	1230	5810
合计		5425	5360	5650	6090	22525

四、土地平衡分析的原则

- 1、土地总体统一平衡的原则。
- 只有从总体上联系各要素进行分析，才能了解其内在的平衡关系。
- 2、相关指标对等平衡原则。
- 进行平衡分析时，设计的统计指标要相互对等。例：
- 期初耕地实有面积+新垦荒地面积+外地区划入面积+征地后退还面积+非法占用退还面积+丈量核实盈余面积+其他增加面积
- =期末实有面积+退耕面积+划归外地区面积+征地减少面积+非法占用面积+丈量核实亏损面积+其他减少面积

五、土地平衡分析方法（单项、综合、联系）

□ （一）土地单项平衡分析

- 又称专题平衡分析。土地单项平衡分析仅反映一个单位或一个地区的一个要素的平衡关系。通常用增减式平衡表的形式来反映。

项 目	面 积	项 目	面 积
一、期初耕地	1660	三、期内减少耕地	220
1. 水田	600	1. 国家基建占地	90
2. 旱地	1060	2. 农村集体与农民建设占地	60
二、期内增加面积	60	3. 退耕还林	50
1. 开荒	10	4. 固定废弃	10
2. 基建还耕	40	5. 其他	10
3. 其他	10	四、期末耕地	1500
		1. 水田	550
		2. 旱地	950
合计	1720	合计	1720

- 期初+期内增加=期内减少+期末 $1660 + (10 + 40 + 10) = (90 + 60 + 50 + 10 + 10) + 1500$
- 单项平衡表的时间长度可以是一个时期，也可以是连续几个时期。连续的便于从较长时间了解土地的增减规律

（二）土地综合平衡分析

- **土地综合平衡分析**是指对土地现象的**整体或全局**进行的平衡关系分析，从土地现象的几个侧面综合反映土地总体的综合平衡关系。

	现有面积	规划面积	现有面积与规划面积之差
耕地面积	1800	1800	0
一、按用途分			
1.粮食作物面积	1000	900	100
2.经济作物面积	800	900	-100
二、按土地等级分			
1.一等	800	1020	-220
2.二等	1000	780	220
三、按权属分			
1.全民所有面积	650	680	-30
2.集体所有面积	1150	1120	30

单项平衡
单项平衡
单项平衡

- **综合平衡分析**能反映土地现象**整体**各种分组之间的相互关系和比例关系，因此能判断总体内各要素的协调程度。单项平衡通过在综合平衡指标下进行，综合平衡又以单项平衡为基础和补充

（三）土地联系平衡分析

- 土地联系平衡分析是从土地的整体出发研究各地区、各部门、各地类之间联系的一种分析方法。采用这种方式可以比较清楚地表明来源与去向。
- 土地联系平衡分析法研究的**重点**是有关土地的**结构比例的合理性**，进而采取促进土地利用结构合理的办法，发现薄弱环节和潜力所在，从而加强薄弱环节，挖掘潜力。

年内地类变化平衡表 (二〇 年)

[illegible]

- 《年内地类变化平衡表》为土地联系平衡分析提供了多种比例、结构，有关比例、结构主要有：
- 1、土地利用分类构成；
- 2、不同类别土地之间的比例；
- 3、各类土地的增减比例；
- 4、总减少规模的分类构成；
- 5、总增加规模的分类构成；
- 6、各类土地的增加来源构成；
- 7、各类土地的减少去向构成。

某地区各类用地调整联系平衡表

单位：公顷

项目		调整后土地面积				合计
		耕地	园地	林地	牧草地	
调整前土地 面积	耕地	540	120	380	450	1490
	园地	95	310	430	760	1595
	林地	380	220	380	650	1630
	草地	410	710	460	230	1810
合计		1425	1360	1650	2090	6525

现有

对角
线上
为保
留的

- 表中各横行表示每一用途占地的调整去向。
- 表中各纵栏则表示调整后的土地来源。
- 表式纵横交错，数字相互制约，可相互验证。既可以用于全局分析，也可用于局部分析，还常用于投入产出分析。

年 市 区耕地增减平衡表

单位：公顷

项目	第一季度	第二季度	第三季度	第四季度	年度
一、期初耕地					
1. 水田					
2. 旱地					
其中：水浇地					
二、期内增加耕地					
1. 开荒					
2. 围垦					
3. 废弃地改造					
4. 其他					
三、期内减少耕地					
1. 国家建设					
2. 集体建设					
3. 农村个人建房					
4. 农业结构调整					
5. 灾害毁地					
四、期末耕地					
1. 水田					
2. 旱地					
其中：水浇地					

年 市 区建设用地增减平衡表

单位：公顷

项目	第一季度	第二季度	第三季度	第四季度	年度
一、期初建设用地					
1. 商服用地					
2. 工矿仓储用地					
3. 住宅用地					
4. 公共管理和公共服务用地					
5. 交通运输用地					
二、期内增加建设用地					
1. 农用地转用					
2. 其他					
三、期内减少建设用地					
四、期末建设用地					
1. 商服用地					
2. 工矿仓储用地					
3. 住宅用地					
4. 公共管理和公共服务用地					
5. 交通运输用地					

年市 区建设用地使用情况增减平衡表

顷

单位：公

项目	第一季度	第二季度	第三季度	第四季度	年度
一、期初建设用地					
二、期内审批建设用地					
1. 实际新增建设用地					
其中：违法面积					
2. 批而未用					
其中：违法面积					
三、期内减少建设用地					
四、期末建设用地					

第六节 土地现象相关回归分析

- I 土地现象**相关**分析
 - 一、相关关系的概念
 - 二、相关关系的类型
 - 三、直线相关分析
- II 土地现象**回归**分析
 - 一、回归分析的概念
 - 二、直线回归分析

一、相关关系的概念

□ （一）函数关系

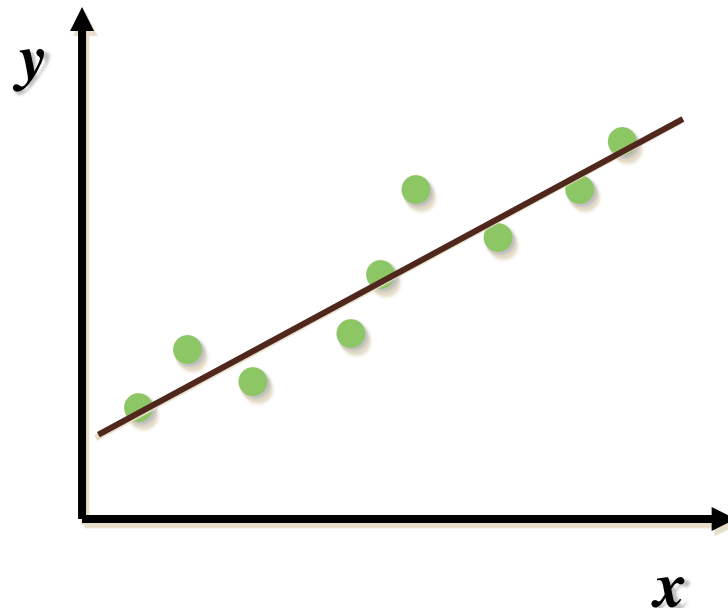
- 变量之间依一定的函数形式形成的一一对应的关系，称为函数关系。它反映着现象之间存在着严密的依存关系，在这种关系中，对于某一变量的一个数值，都有另一变量的确定的值与之对应。

□ （二）相关关系

- 它是指现象之间确实存在的，但关系值不固定的相互依存关系。即对于某一变量的每一个数值，另一变量有若干个数值与之相适应。

相关关系的特点

- (1) 变量间关系不能用函数关系精确表达
- (2) 一个变量的取值不能由另一个变量唯一确定
- (3) 当变量 x 取某个值时, 变量 y 的取值可能有几个
- (4) 各观测点分布在直线周围



相关关系举例

- 父亲身高 y 与子女身高 x 之间的关系
- 收入水平 y 与受教育程度 x 之间的关系
- 商品的消费量 y 与居民收入 x 之间的关系
- 商品销售额 y 与广告费支出 x 之间的关系
- 人口密度 y 与人均耕地占有量 x 之间的关系

相关关系与其他关系的区别与联系

□ 1. 相关关系和函数关系

- **区别：**函数关系是两个变量之间存在着相互依存关系，但它们的关系值是**固定**的；具有相关关系的变量之间关系值**不固定**。
- **联系：**由于有观察或测量误差等原因，确定的函数关系在实际中往往通过相关关系表现出来。当对事物的内部规律了解得更加深刻时，相关关系又可能转化为确定的函数关系。

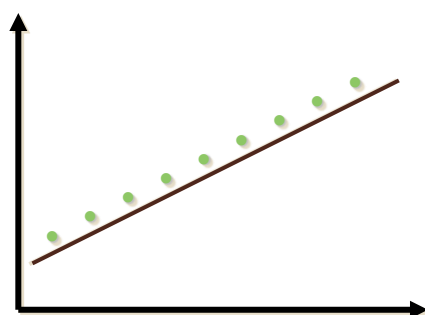
□ 2. 相关关系与因果关系

- 从相关关系的内容讲，有许多是由于因果关系而产生的，但也包括互为因果的关系。同时它还包括非直接的因果关系。所以相关关系比因果关系的概念要广泛。但这种关系必须是客观存在的真实关系。

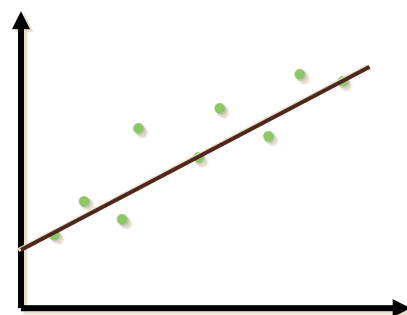
二、相关关系的类型

- (一) 按影响因素的多少分
 - **单相关**涉及两个因素，如耕地与建设用地面积的关系。
 - **复相关**涉及三个以上因素，如，施肥量、浇水量、耕作深度与产量的关系。
- (二) 按现象相关的性质分
 - **1. 数量相关**是根据现象的数量指标进行相关分析，如耕地面积与建设用地面积的相关关系。
 - **2. 质量相关**是根据质量指标进行现象间的相关分析，分为等级相关和品质相关。
 - **1) 等级相关**是按等级排列，如耕地质量等级与棉花等级进行相关分析。
 - **2) 品质相关**是两种不同品质之间的相关关系，如技术革新与产品质量的相关分析。

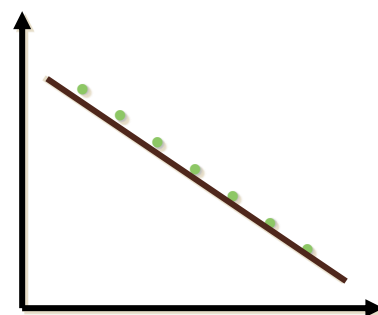
- （三）按变量之间相关关系的方向分
- **正相关**：当一个变量的值增加时，另一变量的值也相应增加；一个变量的值减少时，另一变量的值也相应减少，这样的相关关系就是**正相关**。如农业用地的投入与产出。
- **负相关**：当一个变量的值增加时另一变量的值随之减少，一个变量的值减少时，另一变量值反而增加，这种相关关系称**负相关**。如水土流失率与耕地质量。



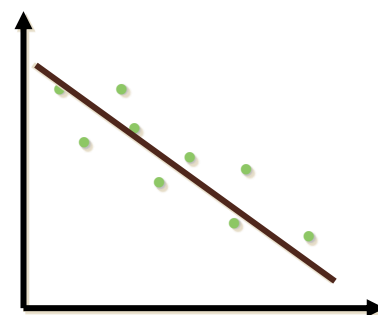
完全正线性相关



正线性相关



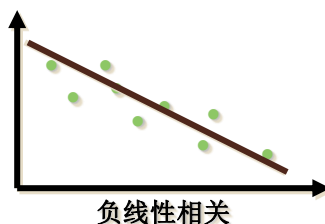
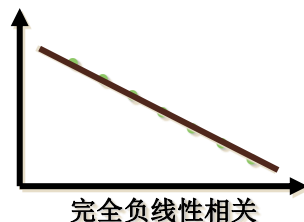
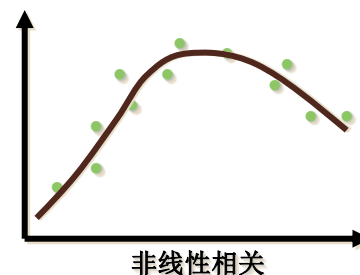
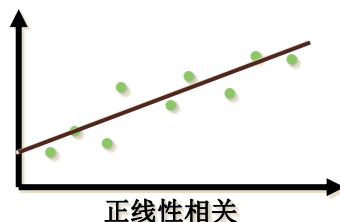
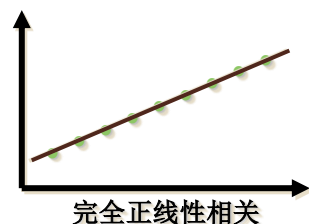
完全负线性相关



负线性相关

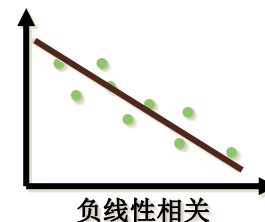
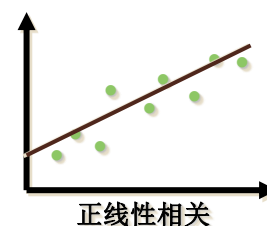
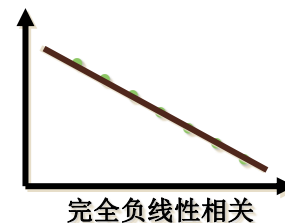
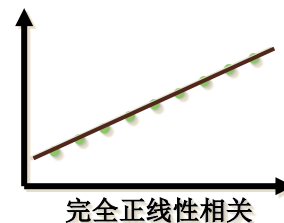
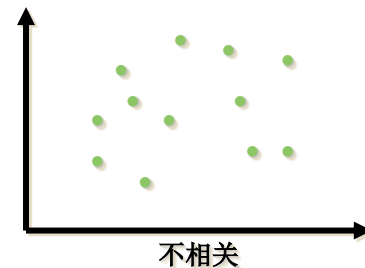
（四）按相关关系的**表现形态**分

- **直线相关**：如果变量 x 值发生变动，和它对应的另一变量 y 随之**大致均等的变动**，从散点图上看，它们相关点的分布近似地表现为**直线形式**。
- **曲线相关**：如果变量 x 值发生变动，和它对应的另一变量 y 值也变动，但**不是均等地变动**，从散点图上看，它们相关点的分布近似于某种**曲线形式**。称曲线相关，也称非线性相关。



（五）按相关的程度分

- **不相关（零相关）**是指变量间存在相关关系，只是由于运用的函数公式不符合相互影响的规律，结果为零。
- **完全相关**是指变量y数值完全随变量x数值变动而变动，在散点图上表现为所有的相关点都落在一条直线上，实际上，这种相关关系已经转化为函数关系。
- **不完全相关**是两个变量之间的相关介于完全相关和不相关之间，作为统计的相关分析一般研究不完全相关现象。



相关分析

相关分析是研究两个或两个以上**随机变量之间**，相互**依存关系紧密程度**的分析，其内容是用相关表、相关图和相关系数等方式确定现象之间**有无**相关关系及相关的表现**形式、方向和程度**。

三、直线相关分析

□ （一）概念和特点

- 直线相关分析是对现象的两个变量之间是否存在线性依存关系及其关系的密切程度进行分析。
- 特点：
 - 1、两个变量之间是对等关系。两个变量都是随机变量，不表示因果关系。
 - 2、只需计算出一个相关系数，相关系数的绝对值在0与1之间，其值大小反映两变量间相关的密切程度。改变两变量间的地位，不影响相关系数的数值。
 - 3、相关系数有正负之分。分别表示正相关和负相关。
 - 4、资料要求：相关的两个变量必须是随机的。

（二）直线相关的测定方式

□ 1. 相关表和相关图

- **相关表**：将其中一个变量的数值按一定的顺序（时间或某一变量数值的大小）排列，同时列出与之对应的其它变量的变量值，这样形成的表格称为相关表。通过相关表，观察其依存关系。
- 根据资料**是否分组**，相关表分为简单相关表和分组相关表两种。
- **1) 简单相关表**是将原始数据按顺序平均排列在一张表格上，借以大致反映变量之间的相关关系。

年份	1990	1991	1992	1993	1994	1995	1996	1997	1998
灌溉面积比率（%）	35.2	36.5	38.1	36.8	39.3	41.2	45.0	44.8	46.7
粮食总产量（万吨）	98.3	99.9	105.2	102.1	109.8	113.5	124.1	125.8	130.2

- 2) 分组相关表：将原始数据进行分组而编制成的表，由于相关表中有两个变量，因此，分组相关表又可分为：单变量分组和双变量分组。
- (2) 单变量分组相关表是按一个变量分组而制成的相关表。
- 编制步骤：①将变量分为若干组；②计算各组频数；③计算各组对应的另一变量的平均值。

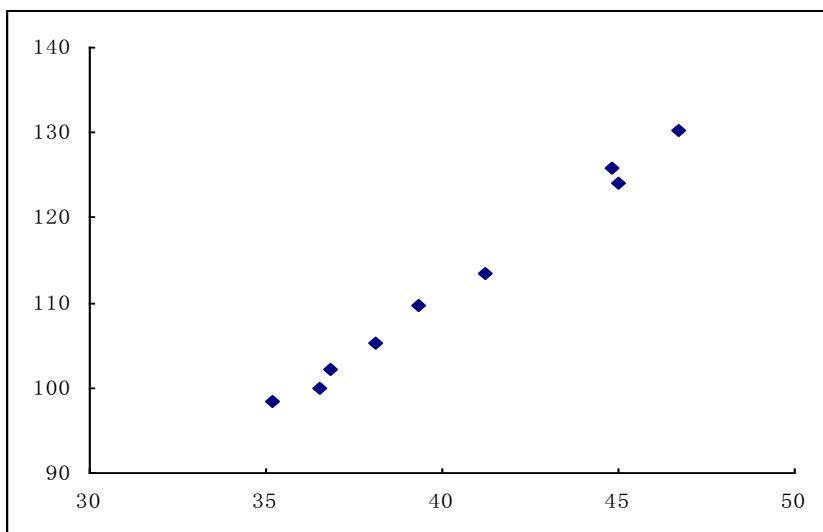
按承包面积分组（公顷）	农户数	产量（吨/公顷）	
5 以下	27	1012.5	
5-10	34	1136.2	
10-15	58	1235.5	
15-20	23	1208.6	
20-25	15	1452.3	
25 以上	8	1526.8	↓

- (2) 双变量分组相关表是自变量和因变量都进行分组而制成的相关表。其编制程序：
- ①确定自变量的组数；②按两个变量的组数设计棋盘表格；③计算各组频数置于相对应的方格之中。

人均国民生产总值		3000 以下	3000-4000	4000-5000	5000-6000	6000-7000	7000 以上	合计
出让面积	组中值 Y	2500	3500	4500	5500	6500	7500	f_y
50 以下	25 ^X	8	4					12
50-100	75	4	12	5				21
100-150	125		2	18	5			25
150-200	175			7	14	3		24
200-250	225		频数/城市个数		2	5	1	8
250 以上	275					4	6	10
合计	f_x	12	18	30	21	12	7	100

100个城市国有土地使用权有偿出让面积与人均国民生产总值

- **相关图/散点图**：根据原始数据，在直角坐标中绘制出两个变量相对应的观察值的所有点，从这些点的分布情况观察分析两个变量间的关系，这个图称为相关图。
- 利用相关图可以直观反映两个变量之间所存在的关系的性质和密切程度。



灌溉面积比率与粮食总产量相关图（直线相关）

（三）相关系数的测定和应用

- **相关系数**是说明两个变量之间**有无直线相关关系及**相关关系密切程度的统计指标。一般用r表示。

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n} \right) \left(\sum y^2 - \frac{(\sum y)^2}{n} \right)}}$$

- **相关系数性质**
- r数值范围，在+1和-1之间，即 $-1 \leq r \leq 1$ 。
- （1）r的绝对值等于1时，表明变量x与变量y之间为完全相关即函数关系。r=1，为完全正相关；r=-1，为完全负相关。
- （2）当 $0 < |r| < 1$ 时，表示变量x与变量y之间有一定相关关系，r>0表示x与y为正相关，r<0表示x与y为负相关。
- （3）r=0时，x与y没有相关关系，称为不相关或零相关。
- |r|越趋于1表示关系越密切；|r|越趋于0表示关系越不密切

相关程度判定的标准

- 为便于判断 $n=10$ 左右的变量相关密切程度，统计上提出了相关关系密切程度的等级划分标准（四级划分法）：
 - $0 < |r| < 0.3$ 弱相关
 - $0.3 \leq |r| < 0.5$ 低度相关
 - $0.5 \leq |r| < 0.8$ 显著相关
 - $0.8 \leq |r| < 1$ 高度相关

一元线性回归计算表

序号	施肥量 (×10 千克/公顷)	产量 (×10 ³ 千 克/公顷)	计算栏		
	x	y	x ²	y ²	xy
1	35	5.90	1225.00	34.81	206.50
2	39	6.55	1521.00	43.90	255.45
3	40	6.20	1600.00	38.44	248.00
4	42	6.99	1764.00	48.86	293.58
5	42	6.30	1764.00	39.69	264.60
6	45	6.53	2025.00	42.64	293.85
7	48	7.50	2304.00	56.25	360.00
8	50	8.72	2500.00	76.04	436.00
9	56	9.52	3136.00	90.63	533.12
10	58	9.49	3364.00	90.06	550.42
合计	455	73.70	21203.00	561.32	3441.52

- 根据表中数据 $\sum x=455$, $\sum y=73.7$, $\sum x^2=21203.00$, $\sum y^2=561.32$, $\sum xy=3441.52$, $n=10$, 代入公式, 可得

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} = \frac{3441.52 - \frac{455 \times 73.7}{10}}{\sqrt{\left(21203 - \frac{455^2}{10}\right)\left(561.32 - \frac{73.70^2}{10}\right)}} = 0.95$$

相关系数的显著性检验

- 准确判断两变量的相关密切程度的方法可利用《相关系数检验的临界值表》，对应 $n-2$ 和 α （相关显著性标准，可以为0.05、0.01等），可以从表中查出相关系数检验的临界值 $r_{\alpha}(n-2)$ ，判定如下：
 - ①当 $|r| < r_{0.05}(n-2)$ 时，表明两变量相关不显著或不相关；
 - ②当 $r_{0.05}(n-2) \leq |r| < r_{0.01}(n-2)$ 时，表明两变量相关显著；
 - ③当 $|r| \geq r_{0.01}(n-2)$ 时，表明两变量相关极显著。

相关系数检验临界值表

自由度	显著性水平(α)			自由度	显著性水平(α)		
$n-m-1$	0.10	0.05	0.01	$n-m-1$	0.10	0.05	0.01
1	0.98769	0.99692	0.99988	201	0.01823	0.01091	0.00288
2	0.90000	0.95000	0.99000	202	0.05068	0.04332	0.02581
3	0.80538	0.87834	0.95874	203	0.06874	0.06615	0.05189
4	0.72930	0.81140	0.91720	204	0.07915	0.08069	0.07253
5	0.66944	0.75449	0.87453	205	0.08573	0.09038	0.08807
6	0.62149	0.70673	0.83434	206	0.09019	0.09718	0.09986
7	0.58221	0.66638	0.79768	207	0.09337	0.10217	0.10898
8	0.54936	0.63190	0.76459	208	0.09573	0.10595	0.11618
9	0.52140	0.60207	0.73479	209	0.09752	0.10888	0.12197
10	0.49726	0.57598	0.70789	210	0.09891	0.11120	0.12670
11	0.47616	0.55294	0.68353	211	0.10001	0.11307	0.13062
12	0.45750	0.53241	0.66138	212	0.10089	0.11460	0.13390
13	0.44086	0.51398	0.64114	213	0.10160	0.11586	0.13667
14	0.42590	0.49731	0.62259	214	0.10217	0.11690	0.13903
15	0.41236	0.48215	0.60551	215	0.10264	0.11777	0.14106
16	0.40003	0.46828	0.58971	216	0.10302	0.11850	0.14281
17	0.38873	0.45553	0.57507	217	0.10332	0.11911	0.14432
18	0.37834	0.44376	0.56144	218	0.10356	0.11962	0.14564
19	0.36874	0.43286	0.54871	219	0.10376	0.12006	0.14679
20	0.35983	0.42271	0.53680	220	0.10391	0.12042	0.14780

四、时间序列自身相关分析

同一时间序列在不同时期各个数值之间存在一定的联系，称为**时间序列自身相关**。通过计算**自身相关系数**可以确定各数值之间的相互联系程度。

时间序列自身相关**特点**：一个变量自身随时间不同，其值与时间顺序表现出一定的依存关系。较多的情况是，任一具体时期的数值都和它前一期或前几期的数值之间存在一定的联系。

包括直线相关和曲线相关

直线相关系数计算公式：

$$r = \frac{n \sum y_t y_{t-m} - \sum y_t \sum y_{t-m}}{\sqrt{[n \sum y_t^2 - (\sum y_t)^2][n \sum y_{t-m}^2 - (\sum y_{t-m})^2]}}$$

计算所得 r 表示时间序列 y_t 与其**前推 m 期**得到的时间序列 y_{t-m} 之间的相关程度。

某村2006、2007年各月末耕地面积，计算2007年时间序列前移5个月的自相关系数并判断相关程度

月份	2007年面积	前移5个月面积	计算栏		
	y_t	y_{t-5}	y_t^2	y_{t-5}^2	$y_t y_{t-5}$
1	49	50	2401	2500	2450
2	48	51	2304	2601	2448
3	48	49	2304	2601	2448
4	47	50	2209	2500	2350
5	48	49	2304	2401	2352
6	47	49	2209	2401	2303
7	46	48	2116	2304	2208
8	46	48	2116	2304	2208
9	47	47	2209	2209	2209
10	46	48	2116	2304	2208
11	45	47	2025	2209	2115
12	45	46	2025	2116	2070
合计	562	584	27363	28438	26338

$$r = \frac{n \sum y_t y_{t-m} - \sum y_t \sum y_{t-m}}{\sqrt{[n \sum y_t^2 - (\sum y_t)^2][n \sum y_{t-m}^2 - (\sum y_{t-m})^2]}}$$

$$= \frac{12 \times 27363 - 562 \times 584}{\sqrt{[12 \times 26338 - 562^2][12 \times 28438 - 584^2]}}$$

$$= 0.72$$

该时间序列自相关程度显著 ($0.5 \leq |r| < 0.8$)

第六节 土地现象相关回归分析

- II土地现象回归分析
 - 一、回归分析的概念
 - 二、直线回归分析

一、回归分析的概念

- **回归分析法**: 测定现象之间数量变化上的一般关系所使用的数学方法的总称。
- **回归分析**是对具有相关关系的两个或两个以上变量之间数量变化的一般关系进行测定，建立一个相应的数学表达式；之后，从一个已知量推算另一个未知量（叫逆行推测，regression，即**回归**），为估算预测提供一个重要的方法。
- **线性回归**分析有一元线性回归和多元线性回归两种；
- **非线性回归**分析，也有一元非线性回归和多元非线性回归两种。
- **回归方程**：在回归分析中建立的相应方程。

二、直线回归分析

- (一) 直线回归分析的概念和特点
- 直线回归分析即一元线性回归分析，是只有一个自变量，而且两个变量的变化按一定比例变化，即两个变量的**变化比率为常数**的回归分析。其**特点**：
 - (1) 两个变量**不是对等**的，要区分自变量和因变量。
 - (2) 回归方程是用自变量数值推算因变量数值的根据，必须反映变量之间关系的一般变动情况。
 - (3) “y倚x回归方程”与“x倚y回归方程”这两条回归直线方程的**斜率不同，意义不同**。
 - (4) 直线回归方程系数即**斜率有正有负**。
 - (5) 资料要求：**因变量是随机的，自变量是给定的数值**。

回归分析与相关分析的区别与联系

区别：

- 1、相关关系用来度量变量之间关系的紧密程度，**本质**上只是对客观存在的关系的**测度**。回归分析是根据所拟合的回归方程研究自变量与因变量一般关系值的方法，可由自变量数值推算因变量数值，具有**推理**的性质。
- 2、相关关系不需要确定哪个是自变量，哪个是因变量；回归分析的首要问题就是确定哪个是自变量，哪个是因变量。
- 3、相关关系研究只能计算**一个**相关系数；回归系数可能有**两个**。

联系：

- 两者**相辅相成**，相关分析测定变量间相关的密切程度，决定着回归分析的必要性和意义的大小。相关系数还是检验回归系数的标准，回归分析的结果也可推算相关系数。相关分析需要回归分析来表明现象数量关系的具体形式，回归分析应建立在相关分析基础上。

（二）简单直线回归方程的建立和求解

- 配合回归线相应的方程称为**直线回归方程**，即一元一次方程。即：

- $y=a+bx$ 实际上应为： $y=a+bx +u_t$ —— u_t 是误差项即后面的 S_{yx}

- 式中， y 为因变量， x 为自变量， a 与 b 是待定参数。 a 为直线的截距， **b 为直线斜率又称回归系数**。

- 确定参数 a 和 b

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

- 则

$$a = \bar{y} - b\bar{x}$$

- 式中： \bar{x} 、 \bar{y} 为资料中自变量和因变量的平均值。
- 直线回归方程的检验可用相关系数和方差分析检验，并计算估计误差进行预测

（三）计算估计标准误差

- 估计标准误差的计算公式如下：

$$S_{yx} = \sqrt{\frac{\sum (y - y_c)^2}{n - 2}}$$

- 公式中 S_{yx} 代表估计标准误差，即 x 为自变量， y 为因变量时的估计标准误差。 y —因变量实际观察值； y_c —相对应于自变量 x 的回归值。计算较繁琐

- 简捷计算方法

$$S_{yx} = \sqrt{\frac{\sum y^2 - a \sum y - b \sum xy}{n - 2}}$$

（四）直线回归分析的步骤

- 直线回归分析在相关分析基础上进行，包括：
- 1. 判定简单线性相关及相关的密切程度。一般通过相关表、相关图和相关系数判定，其中相关系数最重要。
- 2. 给出两变量的直线回归方程，并列参数 a 、 b 的计算表，即回归分析计算表。
- 3. 计算参数 a 、 b 并代入回归方程。
- 4. 计算估计标准误差。
- 5. 解释回归分析结果的含义。

（五）运用回归方程分析时注意的问题

- 用回归方程分析变量之间的变动关系，是一种科学的方法，在计算和应用时，应注意以下几点：
- （一）在定性分析的基础上进行定量分析，是保证正确运用回归分析的必要条件。
- （二）在回归方程中，回归系数的绝对值只能表示自变量与因变量之间的联系程度，以及两变量间的变动比例。因为其值大小直接取决于变量所用计算单位的大小。
- （三）具体问题具体分析，推算或预测时要注意条件的变化。
- （四）最好要与相关系数、估计标准误差结合使用。
- （五）注意社会经济现象的复杂性。

某农场利润率与劳动生产率回归计算表

序号	生产率 (千元) x	利润率 (%) y	x ²	y ²	xy
1	8	5.5	64	30.25	44
2	10	6	100	36	60
3	10	6.5	100	42.25	65
4	10	7	100	49	70
5	12	8	144	64	96
6	12	8.5	144	72.25	102
7	14	9	196	81	126
8	16	10.5	256	110.25	168
9	18	12.5	324	156.25	225
10	20	14	400	196	280
合计	130	87.5	1828	837.25	1236

- 1. 首先计算两者的相关系数，计算得

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} = \frac{1236 - \frac{130 \times 87.5}{10}}{\sqrt{\left(1828 - \frac{130^2}{10}\right) \left(837.25 - \frac{87.5^2}{10}\right)}} = 0.907$$

- 查相关系数检验表： $r_{0.05}(8)=0.632$ ， $r_{0.01}(8)=0.765$ ， $r=0.907$ 均大于 $r_{0.05}(8)$ 和 $r_{0.01}(8)$ ，两变量呈极显著直线相关。

- 2. 因利润率与劳动生产率呈直线相关，所以提出简单直线回归方程式：
- $y=a+bx$
- 3. 用最小二乘法计算参数a和b：

$$\square b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = \frac{10 \times 1236 - 130 \times 87.5}{10 \times 1828 - 130^2} = 0.714$$

□ 则

$$\square a = \bar{y} - b\bar{x} = 8.75 - 0.714 \times 13 = -0.532$$

□ 回归方程：

$$\square y = 0.714x - 0.532$$

- 4. 得出回归方程后，还应计算估计标准误差：

$$s_{yx} = \sqrt{\frac{\sum y^2 - a \sum y - b \sum xy}{n-2}} = \sqrt{\frac{837.25 + 0.532 \times 87.5 - 0.714 \times 1236}{10-2}} = 0.405$$

- 估计标准误差平均占因变量实际值的比例为：

$$s_{yx} / (\sum y / n) = \frac{0.405}{\frac{87.5}{10}} = 4.62\%$$

- 因此，估计值可在一定概率上确定区间，即

- $y_c \pm 0.405$ ，概率为68.3%
- $y_c \pm 0.810$ ，概率为95.45%
- $y_c \pm 1.215$ ，概率为97.73%

$$y=0.714x-0.532$$

- 以上回归分析结果表明：
- 在劳动生产率为8-20千元或其邻近范围内，该农场利润率、劳动生产率呈**极显著直线相关**，劳动生产率每提高或降低1千元时，将使利润率平均提高或降低0.714%；
- 如果根据所建回归方程计算**一定劳动生产率所对应的利润率的估计值**，则该劳动生产率对应的实际值落在估计值上下0.405个百分点范围内的可能性为68.3%，落在估计值上下0.810个百分点范围内的可能性是95.45%，落在估计值上下1.215个百分点范围内的可能性是97.73%；
- 估计标准误差平均占实际值的比例为4.62%，小于一般的误差要求5%，说明用此回归方程进行估计较为准确。

三、多元线性回归分析

- y 为因变量， x_1x_2 为自变量，其回归方程：
- $y_c=a+bx_1+cx_2$
- y_c ——回归估计值， a 常数项， b 、 c —— y_c 分别依 x_1x_2 的回归系数， y_c 对某一自变量固定时，该自变量变动一个单位使 y_c 平均变动的数值，所以也称为偏回归系数
- 求二元回归方程中 a 、 b 、 c 的正规（标准）方程组如下：
- $\sum y=na+b\sum x_1+c\sum x_2$
- $\sum x_1y=a\sum x_1+b\sum x_1^2+c\sum x_1x_2$
- $\sum x_2y=a\sum x_2+b\sum x_1x_2+c\sum x_2^2$

解正规方程组，可求得 a 、 b 、 c 的值，代入 $y_c=a+bx_1+cx_2$ 即得二元线性回归方程