



项目四

数据资料整理

任务
1

常用术语及其含义



数据资料整理

目的要求

掌握**统计学**上几个基本概念，准确制作次数分布表、图，计算平均数、极差、方差、标准差、变异系数。

主要内容

▶ 1.次数分布

▶ 2.平均数和变异数

目录

01

常用术语及其含义

02

次数分布



一、变数与观测值



一、变数与观测值

- 1、**变数**：在统计分析中，为了研究方便，将具有变异性的各种研究要素称作变数。
- 2、**观测值**：在同一研究要素下，不同单元或个体的数据称观测值。

番茄单果重: (g)
201 , 225 , 227 , 301 , 278 , 293

观测值

变数



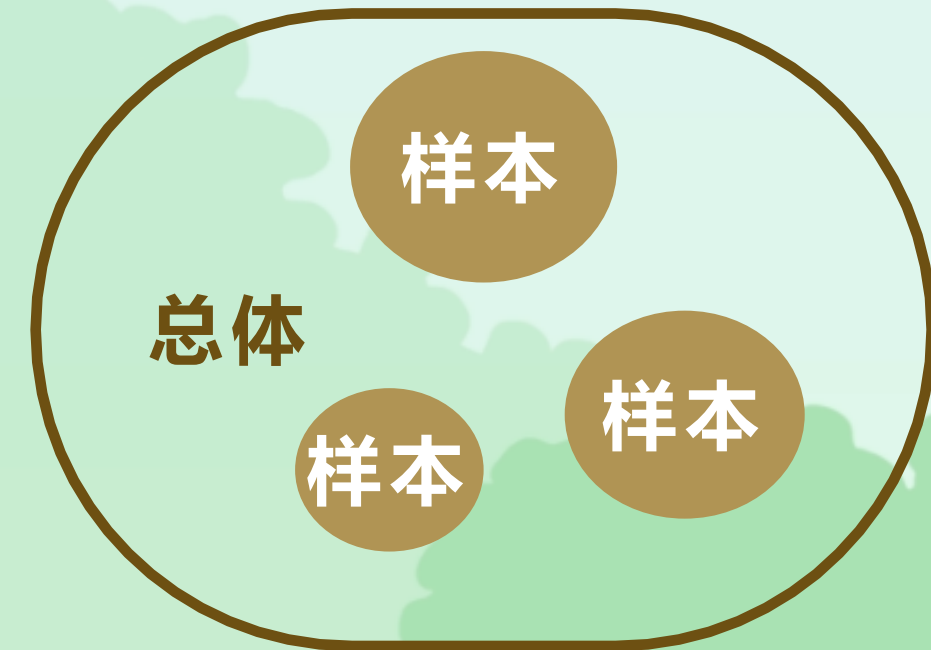


一、变数与观测值

3、**总体**：具有共同性质的个体所组成的集团。分有限总体和无限总体。

4、**样本**：从总体中抽样取出来的若干个体。

二者的关系





二、随机变数的性质与类型



二、随机变数的性质与类型

(一) 随机变数的性质

① 取值不确定，只知取值范围。

② 观测值有一定的规律性，服从一定的概率分布。



二、随机变数的性质与类型

(二) 随机变数的类型

▶ (1) 数量性状变数

- ① **连续性变数**：取值是某一区间的任何数值。如产量、单果重等性状。
番茄单果重大约在30g ~ 800g之间。
- ② **间断性变数（不连续变数）**：由计数方法获得，如株数、果数等以1为单位。



二、随机变数的性质与类型

(二) 随机变数的类型

▶ (2) 质量性状变数：指能观察而不能测量的性状变数。如花色、果色等。

➤ 常见的数量化的方法有：

① 分级法：
将变异的性状分成几种级别，每一级别指定以适当的数值代表。

② 应用统计次数的方法：
统计某种性状出现及不出现的次数。

200株豌豆中，
有140株紫花占70%
，60株白花占30%



三、参数与统计数



三、参数与统计数

参数：

由总体全部数据计算得到的总体特征数。习惯上参数用希腊字母表示。如 μ 表示总体平均数。

统计数：

由样本全部数据计算得到的样本特征数。习惯上统计数用英文字母表示。如 \bar{x} 表示样本平均数。

\bar{x}



次数分布



数据资料的整理

| | | | | |
|----|----|----|----|----|
| 17 | 15 | 18 | 14 | 11 |
| 17 | 11 | 11 | 20 | 19 |
| 10 | 17 | 11 | 17 | 15 |
| 12 | 19 | 11 | 12 | 13 |
| 16 | 12 | 19 | 12 | 15 |





数据资料的整理

次数分布表

将**观测值**按大小进行分组统计次数，编成表格形式即为次数分布表。

次数分布图

用**图形**表示资料的分布情况。





一、次数分布表



一、次数分布表

1. 连续性变数资料的整理

黄瓜果实长度（单位：cm）

| | | | | | |
|------|------|------|------|------|------|
| 21.4 | 21.2 | 25.1 | 21.6 | 26.9 | 25.9 |
| 20.4 | 24.6 | 21.2 | 30.2 | 25.8 | 28.3 |
| 20.7 | 17.1 | 17.7 | 20.6 | 21.1 | 20.9 |



一、次数分布表

1. 连续性变数资料的整理

▶ **极差**：指数据中的最大值与最小值之差。

▶ **组数**：指整个数据范围内分成的组区间数。

▶ **组限**：每个组区间的上下界限。数值小的一端叫下限，或低限。数值大的一端叫上限，或高限。

教材34页
表3-1样本大小与分组数



一、次数分布表

1. 连续性变数资料的整理

组距：每个组区间上限与下限之差，即组区间长度。组距可由极差和组数近似求得，即：**组距 = 极差 / 组数**。

组中值：也称中价、组值等，是每一个组内最中间的一个数，它可由每个组的上、下限的平均数得到，即**组中值 = (上限 + 下限) / 2**



一、次数分布表

连续性变数资料的制表

- ▶ **1.正列：由小到大排列**
- ▶ **2.求极差：最大数-最小数**
- ▶ **3.确定组数与组距：按照表确定，一般选奇数。组距 = 极差 / 组数 5 组**
- ▶ **4.确定第一组组中值与组限：人为确定组中值，下限 = 组中值 - 组距 / 2 上限**
- ▶ **5.列出全部组中值与组限。**

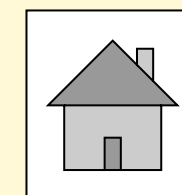


一、次数分布表

连续性变数资料的制表

黄瓜果实长度的次数分布表

| 组限 | 组中值 | 次数 |
|-------------|-----|----|
| 16.5 ~ 19.5 | 18 | 2 |
| 19.5 ~ 22.5 | 21 | 9 |
| 22.5 ~ 25.5 | 24 | 2 |
| 25.5 ~ 28.5 | 27 | 4 |
| 28.5 ~ 31.5 | 30 | 1 |
| 总计 | | 18 |





一、次数分布表

2.非连续性变数资料的整理

| | | | | |
|----|----|---|----|----|
| 10 | 10 | 7 | 8 | 10 |
| 8 | 9 | 9 | 8 | 10 |
| 9 | 8 | 7 | 10 | 8 |

番茄单株结果数（单位：个）



一、次数分布表

2.非连续性变数资料的整理

番茄单株结果数次数分布表

| 结果数 | 次数 |
|-----|----|
| 7 | 2 |
| 8 | 5 |
| 9 | 3 |
| 10 | 5 |



课堂练习

调查某番茄品种的花序坐果数，调查20个花序：

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 0 | 1 | 2 | 4 | 6 | 4 | 3 | 5 | 5 | 0 |
| 3 | 4 | 3 | 4 | 5 | 6 | 1 | 3 | 6 | 1 | |

编制次数分布表。





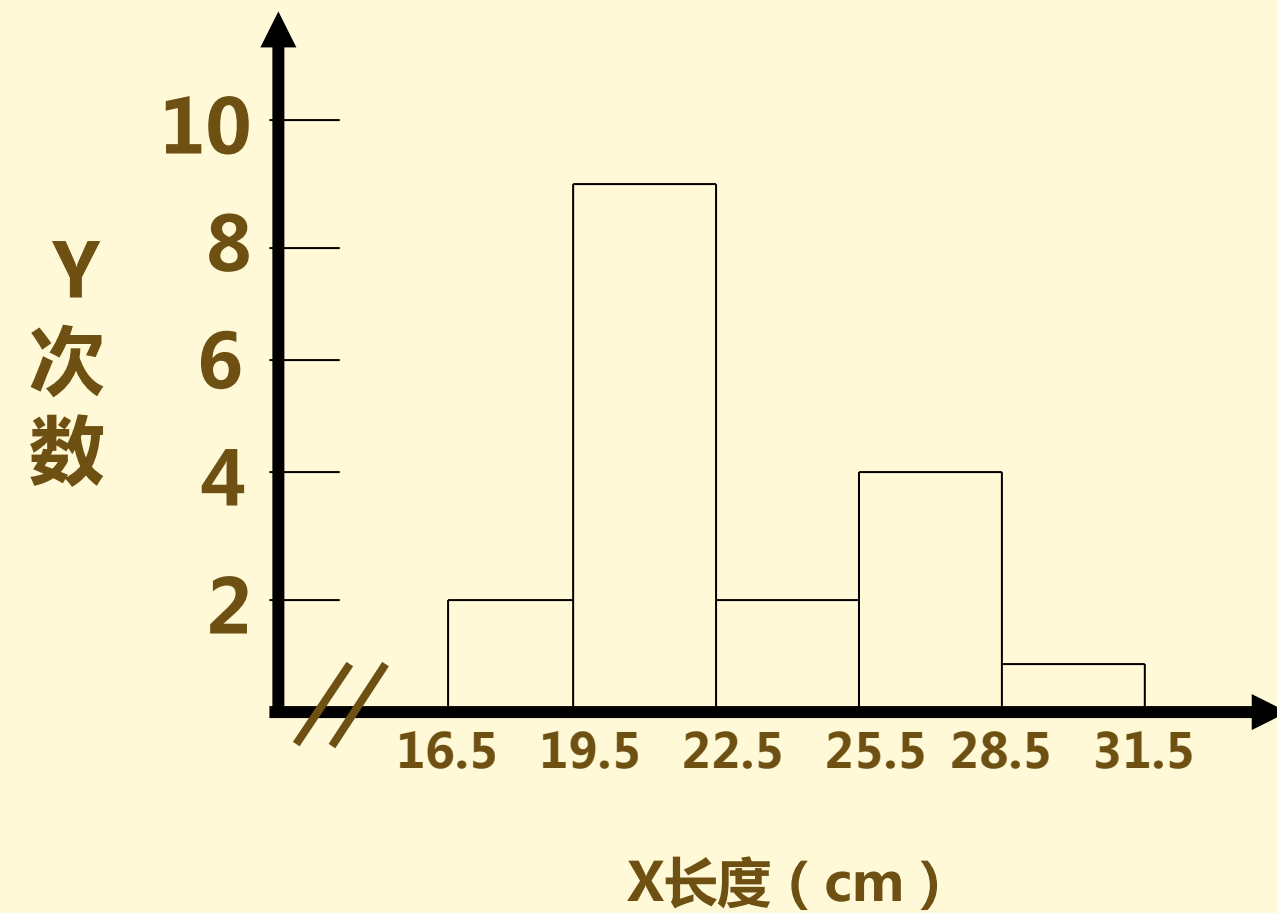
二、次数分布图



二、次数分布图

1. 柱形图

黄瓜果实长度次数分布柱形图

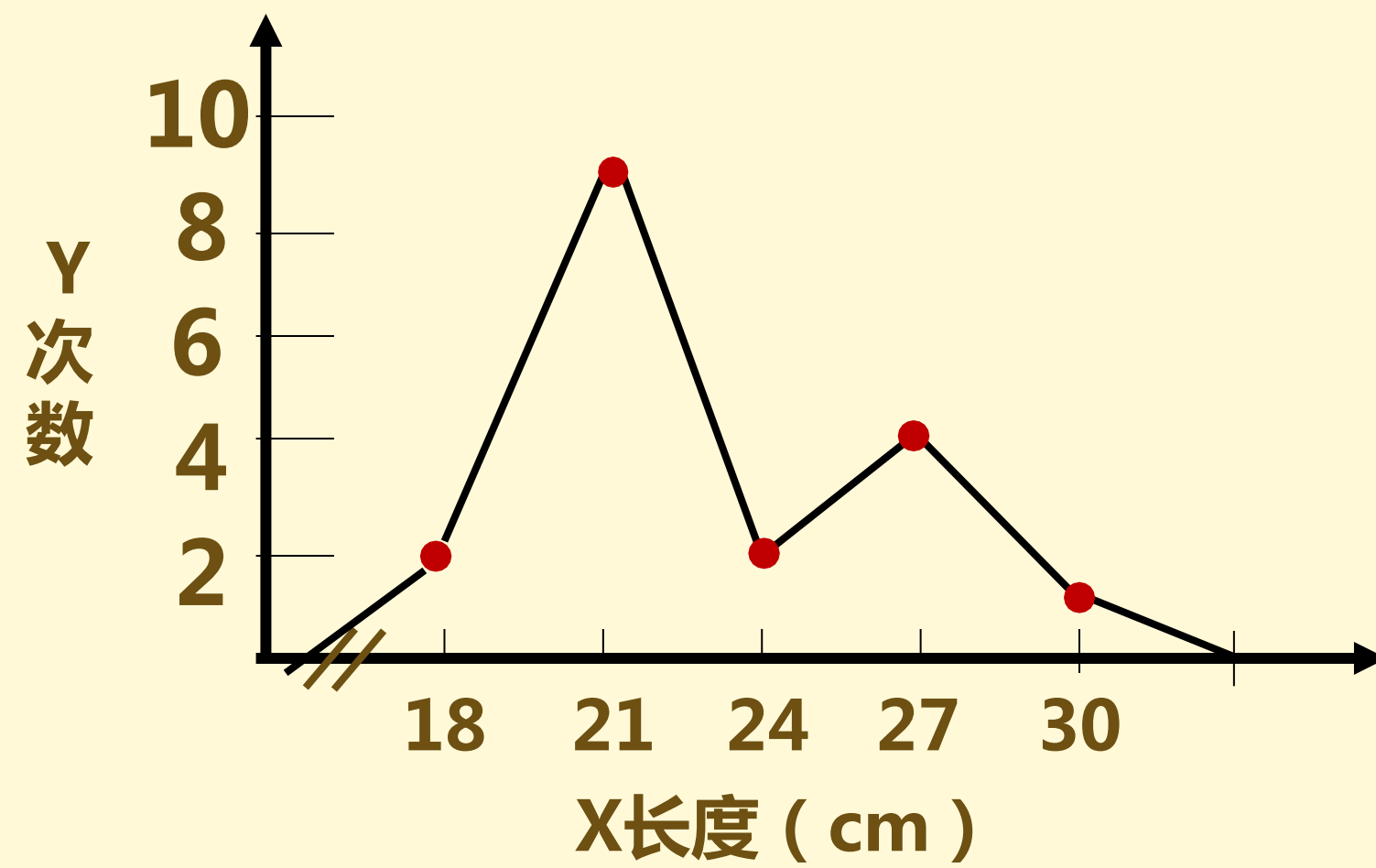




二、次数分布图

2. 多边形图

黄瓜果实长度次数分布多边形图

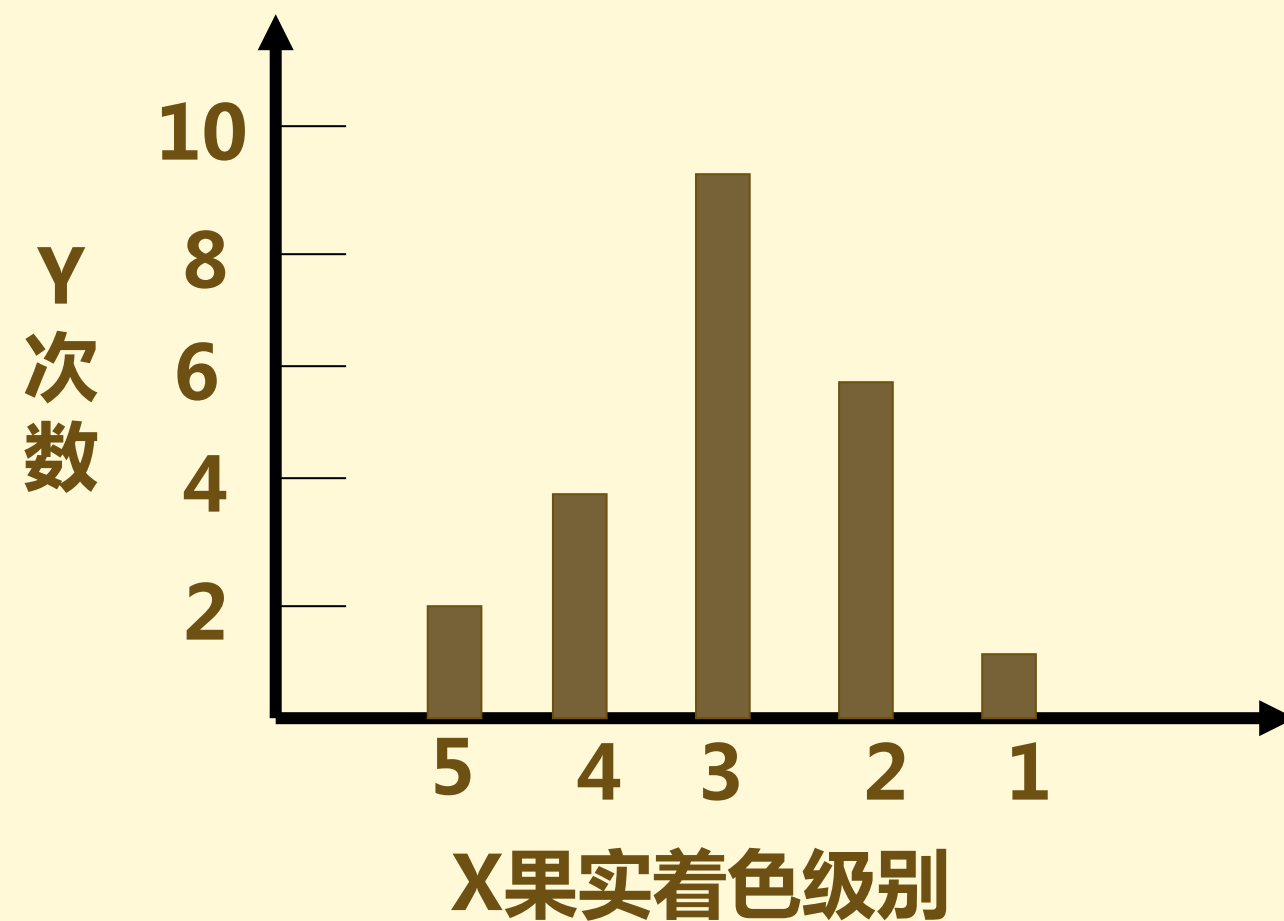




二、次数分布图

3. 条形图

红星苹果着色次数分布条形图



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/397115111151006121>