

第二十章 数据的分析

20.2 数据的波动程度

第1课时 方差

学习目标

1. 理解方差的概念及统计学意义；
2. 会计算一组数据的方差；（重点）
3. 能够运用方差判断数据的波动程度，并解决简单的实际问题。（难点）

导入新课

刘教练到我班选拔一名篮球队员，刘教练对陈方楷和李霖东两名学生进行5次投篮测试，每人每次投10个球，下图记录的是这两名同学5次投篮中所投中的个数。

队 员	第 1次	第2次	第3次	第4次	第5次
李霖东	7	8	8	8	9
陈方楷	10	6	10	6	8

- (1) 请求出以上两组数据的平均数、中位数、众数；
- (2) 用复式折线统计图表示上述数据；
- (3) 若要选一个投篮稳定的队员，选谁更好？

方差的意义

问题1 农科院计划为某地选择合适的甜玉米种子。选择种子时，甜玉米的**产量**和**产量的稳定性**是农科院所关心的问题。为了解甲、乙两种甜玉米种子的相关情况，农科院各用10块自然条件相同的试验田进行试验，得到各试验田每公顷的产量（单位：t）如下表：



甲	7.65	7.50	7.62	7.59	7.65	7.64	7.50	7.40	7.41	7.41
乙	7.55	7.56	7.53	7.44	7.49	7.52	7.58	7.46	7.53	7.49

根据这些数据估计，农科院应该选择哪种甜玉米种子呢？



甲	7.65	7.50	7.62	7.59	7.65	7.64	7.50	7.40	7.41	7.41
乙	7.55	7.56	7.53	7.44	7.49	7.52	7.58	7.46	7.53	7.49

(1) 甜玉米的产量可用什么量来描述？请计算后说明。

$$\bar{x}_{\text{甲}} \approx 7.54, \bar{x}_{\text{乙}} \approx 7.52$$

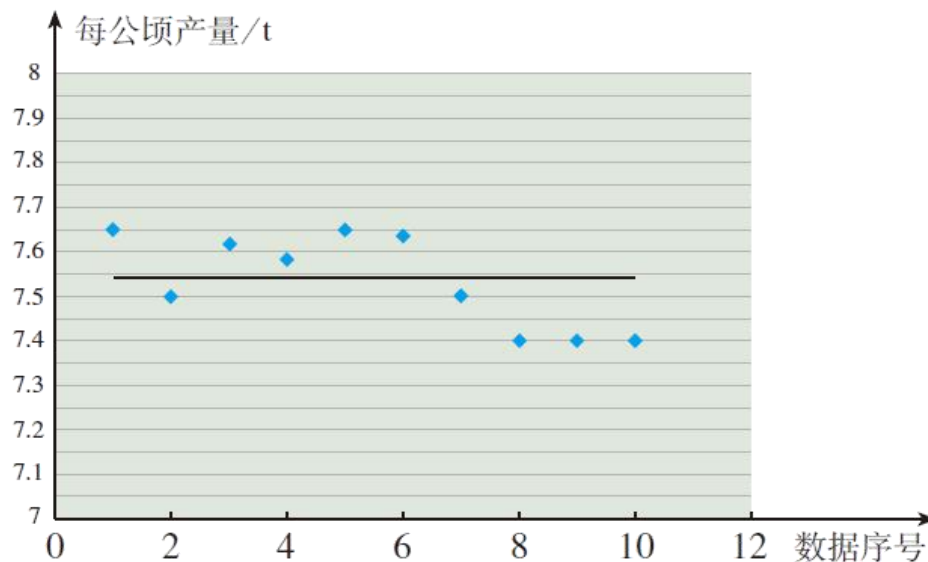
说明在试验田中，甲、乙两种甜玉米的**平均产量**相差不大。

可估计这个地区种植这两种甜玉米的**平均产量**相差不大。

(2) 如何考察一种甜玉米产量的**稳定性**呢？

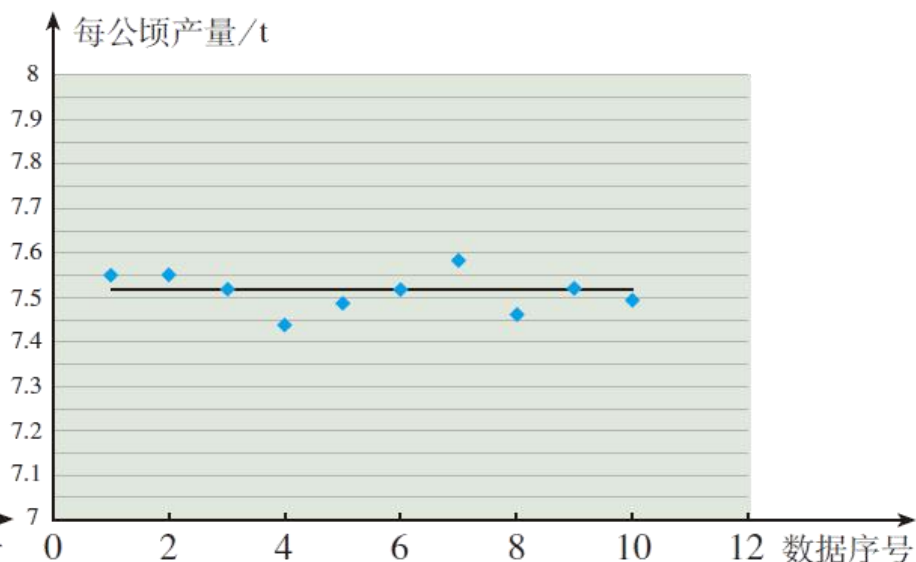
①请设计统计图直观地反映出甜玉米产量的分布情况。

甲种甜玉米的产量



产量波动较大

乙种甜玉米的产量



产量波动较小

探究新知



统计学中常采用下面的做法来量化这组数据的波动大小：

设有 n 个数据 x_1, x_2, \dots, x_n ，各数据与它们的平均数 \bar{x} 的差的平方分别是 $(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \dots, (x_n - \bar{x})^2$ ，我们用这些值的平均数，即用

$$s^2 = \frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]$$

来衡量这组数据的波动大小，称它为这组数据的**方差**。

方差**越大**，数据的波动**越大**；

方差**越小**，数据的波动**越小**。

②请利用方差公式分析甲、乙两种甜玉米的波动程度.

两组数据的方差分别是:

$$s_{\text{甲}}^2 = \frac{(7.65-7.54)^2 + (7.50-7.54)^2 + \dots + (7.41-7.54)^2}{10} \approx 0.01$$

$$s_{\text{乙}}^2 = \frac{(7.55-7.52)^2 + (7.56-7.52)^2 + \dots + (7.49-7.52)^2}{10} \approx 0.002$$

显然 $s_{\text{甲}}^2 > s_{\text{乙}}^2$, 即说明甲种甜玉米的波动较大,

这与我们从产量分布图看到的结果一致.

据样本估计总体的统计思想, 种乙种甜玉米产量较稳定.

应用新知



例 在一次芭蕾舞比赛中，甲、乙两个芭蕾舞团都表演了舞剧《天鹅湖》，参加表演的女演员的身高（单位：cm）分别是：

甲团	163	164	164	165	165	166	166	167
乙团	163	165	165	166	166	167	168	168

哪个芭蕾舞团女演员的身高更整齐？



练一练

1.用条形图表示下列各组数据，计算并比较它们的平均数和方差，体会方差是怎样刻画数据的波动程度的.

(1) 6 6 6 6 6 6;

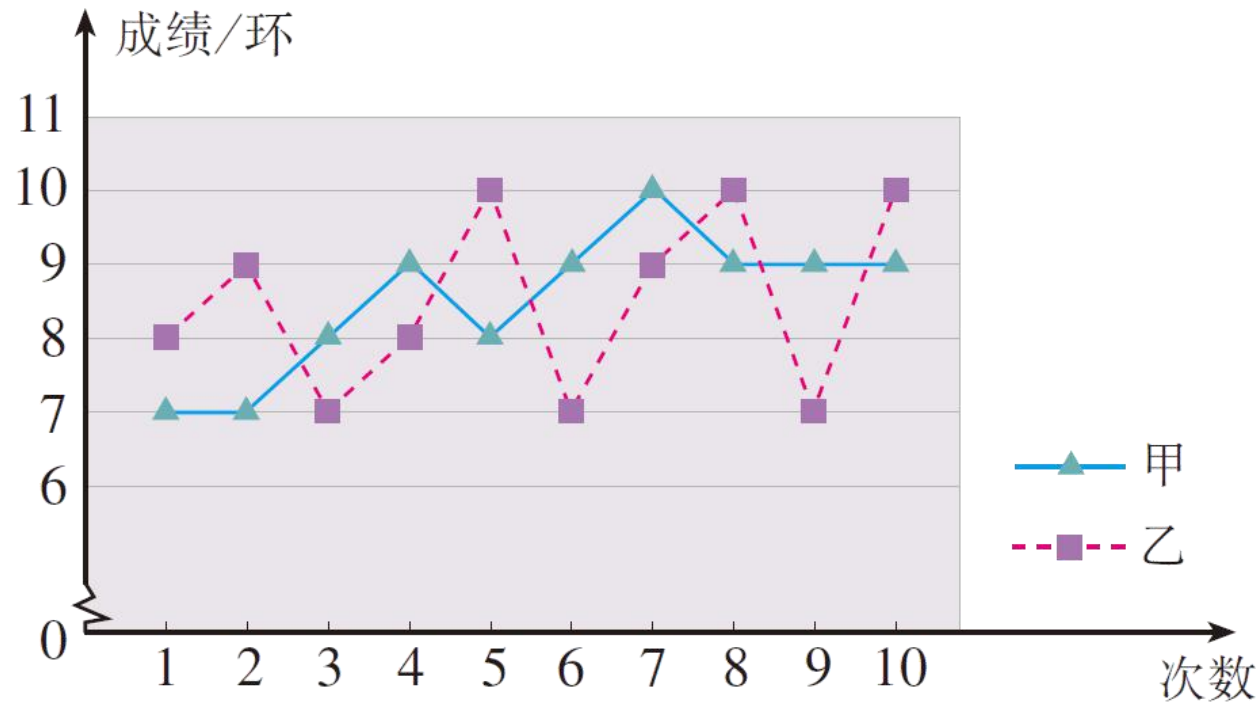
(2) 5 5 6 6 6 7 7;

(3) 3 3 4 6 8 9 9;

(4) 3 3 3 6 9 9 9.

【答】 (1)平均数：6，方差：0； (2)平均数：6；方差： $\frac{4}{7}$
(3)平均数：6，方差： $\frac{44}{7}$ ； (4)平均数：6，方差： $\frac{54}{7}$.

2.如图是甲、乙两射击运动员的10次射击训练成绩的折线统计图. 观察图形, 甲、乙这10次射击成绩的方差哪个大?



【答】乙的射击成绩波动大, 所以乙的方差大.

求一组较大数据的方差，有如下简便计算方法：

方法拓展

1 任取一个基准数 a

2 将原数据减去 a ，得到一组新数据

3 求新数据的方差

知识拓展

若数据 x_1, x_2, \dots, x_n 平均数为 \bar{x} , 方差为 s^2 , 则

① 数据 $x_1 - 3, x_2 - 3, \dots, x_n - 3$

平均数为 $\bar{x} - 3$; 方差为 s^2 .

② 数据 $x_1 + 3, x_2 + 3, \dots, x_n + 3$

③ 数据 $3x_1, 3x_2, \dots, 3x_n$

平均数为 $3\bar{x}$, 方差为 $9s^2$.

④ 数据 $2x_1 - 3, 2x_2 - 3, \dots, 2x_n - 3$

平均数为 $2\bar{x} - 3$, 方差为 $4s^2$.

1.样本方差的作用是（ D ）

A.表示总体的平均水平

B.表示样本的平均水平

C.准确表示总体的波动大小

D.表示样本的波动大小,从而估计总体的波动大小

2.人数相同的八年级（1）、（2）两班学生在同一次数学单元测试中，班级平均分和方差下：

$\bar{x}_{\text{甲}} = \bar{x}_{\text{乙}} = 80$ ， $s_{\text{甲}}^2 = 24$ ， $s_{\text{乙}}^2 = 18$ ，则成绩较为稳定的班级是（**B**）

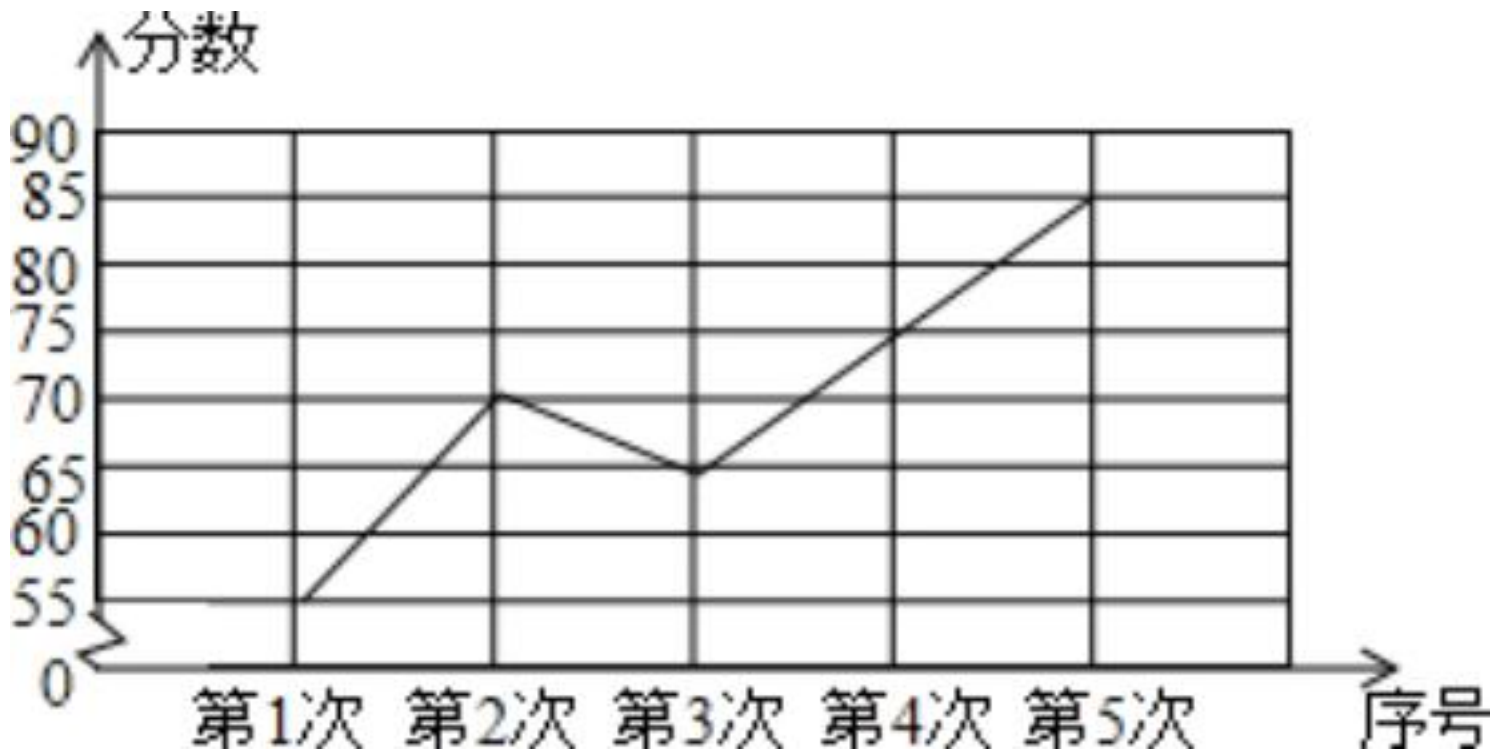
A.甲班

B.乙班

C.两班成绩一样稳定

D.无法确定

3.小凯同学参加数学竞赛训练，近期的五次测试成绩得分情况如图所示，则他这五次成绩的方差为 100.



4.在样本方差的计算公式 $s^2 = \frac{1}{10}[(x_1 - 20)^2 + (x_2 - 20)^2 \dots + (x_n - 20)^2]$

中，数字10表示 样本容量，数字20表示 平均数。

5.五个数1, 3, a , 5, 8的平均数是4, 则 $a =$ 3, 这五个数的方差 5.6。

6.为了从甲、乙两名学生中选择一人去参加电脑知识竞赛，在相同条件下对他们的电脑知识进行10次测验，成绩（单位：分）如下：

甲的成绩	76	84	90	84	81	87	88	81	85	84
乙的成绩	82	86	87	90	79	81	93	90	74	78

(1) 填写下表：

同学	平均成绩	中位数	众数	方差	85分以上的频率
甲	84	84	84	14.4	0.3
乙	84	84	90	34	0.5

(2) 利用以上信息，请从不同的角度对甲、乙两名同学的成绩进行评价.

解：从众数看，甲成绩的众数为84分，乙成绩的众数是90分，乙的成绩比甲好；

从方差看， $s^2_{\text{甲}}=14.4$ ， $s^2_{\text{乙}}=34$ ，甲的成绩比乙相对稳定；从甲、乙的中位数、平均数看，中位数、平均数都是84分，两人成绩一样好；

从频率看，甲85分以上的次数比乙少，乙的成绩比甲好.

方差

公式：
$$s^2 = \frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2]$$

方差的统计学意义（判断数据的波动程度）：
方差越**大**（**小**），数据的波动越**大**（**小**）。