

第二十三章 数据分析

23.3 方差 (1)

学习新知

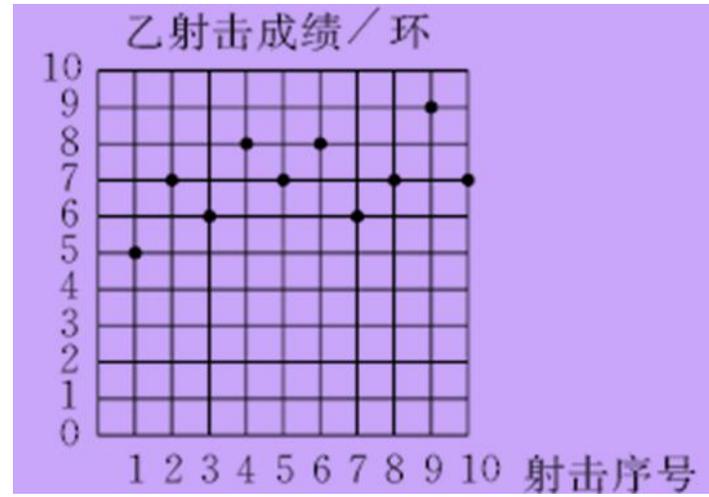
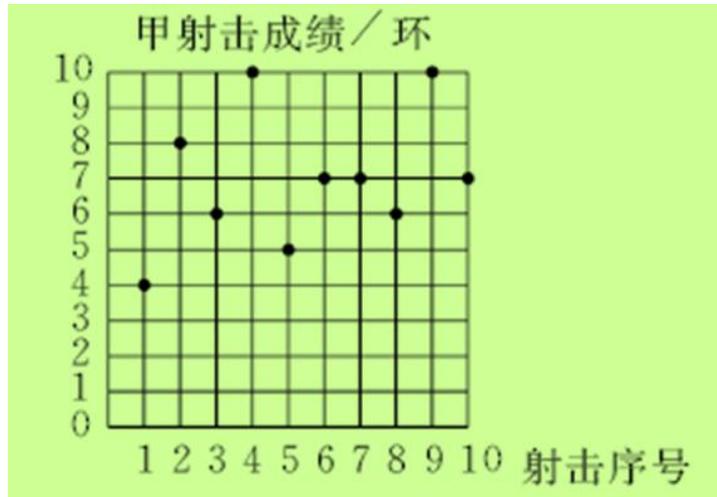
检测反馈

知识复习

- 1.表示一组数据的“集中趋势”的统计量有哪些？
- 2.什么是平均数、中位数和众数？
- 3.如何计算一组数据的平均数、中位数和众数？

观察与思考

甲、乙两名业余射击选手参加了一次射击比赛,每人各射10发子弹,成绩如图所示.



(1)观察上图,甲、乙射击成绩的平均数、中位数各是多少?

(两人射击成绩的平均数和中位数都是7环)

(2)甲、乙射击成绩的平均数是否相同?若相同,他们的射击水平就一样吗?

(两人射击成绩的平均数相同,但并不能说明射击水平一定相同)

(3)哪一组数据相对于其平均数波动较大?波动大小反映了什么?

(甲射击成绩波动较大,波动的大小反映射击的稳定性有差异)

1.如何描述每个数据与平均数的偏差?

$$(x_1 - \bar{x}, x_2 - \bar{x}, \dots, x_n - \bar{x})$$

2.把所有的偏差直接相加能表示所有数据的总偏差吗?

(不能,因为正负偏差会相互抵消,偏差总和为0)

3.如何防止正负偏差相互抵消?

(将各偏差平方后再求和)

4. 如何消除数据个数的影响?

(将各偏差平方后再求平均数)

设 n 个数据 x_1, x_2, \dots, x_n 的平均数为 \bar{x} , 各个数据与平均数偏差的平方分别是

$$(x_1 - \bar{x})^2, (x_2 - \bar{x})^2, \dots, (x_n - \bar{x})^2.$$

偏差平方的平均数叫做这组数据的**方差**, 用 s^2 表示, 即

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

思考

- 1.方差的取值范围是什么?
- 2.如何求一组数据的方差?
- 3.如何用方差的大小衡量离散程度的大小?
- 4.方差为0的一组数据有什么特点?

5.你能通过求方差的方法,说明上述问题中哪个射击选手的成绩比较稳定吗?

$$s_{\text{甲}}^2 = \frac{1}{10} [(4-7)^2 + (5-7)^2 + 2(6-7)^2 + 3(7-7)^2 + (8-7)^2 + 2(10-7)^2] = 3.4,$$

$$s_{\text{乙}}^2 = \frac{1}{10} [(5-7)^2 + 2(6-7)^2 + 4(7-7)^2 + 2(8-7)^2 + (9-7)^2] = 1.2.$$

因为 $s_{\text{甲}}^2 > s_{\text{乙}}^2$, 所以乙的射击成绩比甲的波动小, 乙的成绩更稳定些.

结论

方差的值为非负数;当方差为0时,这组数据为相同的一组数值;当数据分布比较分散时,方差较大;当数据分布比较集中时,方差较小.因此,方差的大小反映了数据波动(或离散程度)的大小.

例1 利用计算器计算下列数据的平均数和方差.
(结果精确到0.01)

66 78 81 75 86 82

解:(1)进入统计状态,选择一元统计.

(2)输入数据.

(3)显示结果.

按 $\boxed{\text{Rcl}}$ \bar{x} 键,显示结果为 78.

按 $\boxed{\text{Rcl}}$ s_x^2 键,显示结果为 40.33333.

所以 $\bar{x}=78, s^2 \approx 40.33$.

[知识拓展]

1. 方差是用来衡量一组数据波动大小的重要量.
2. 方差反映的是数据在它的平均数附近波动的情况.
3. 对于同类问题的两组数据, 方差较大的波动较大, 方差较小的波动较小.
4. 一组数据的每一个数据加上(或减去)同一个常数, 所得的一组新数据的方差不变.
5. 一组数据的每一个数据都变为原来的 k 倍, 则所得的一组新数据的方差将变为原数据方差的 k^2 倍.

1.(2015·上海中考)下列统计量中,能表示一组数据波动程度的是 (C)

- A.平均数 B.众数
C.方差 D.频率

解析:表示一组数据波动大小的统计量是方差.故选C.

2.甲、乙两班举行电脑汉字输入比赛,参赛学生每分钟输入汉字的个数统计结果如下表:

班级	参赛人数	中位数	方差	平均数
甲	55	149	191	135
乙	55	151	110	135

某同学分析上表后得出如下结论:

- ①甲、乙两班学生成绩平均水平相同;
- ②乙班优秀的人数多于甲班优秀的人数(每分钟输入汉字 ≥ 150 个为优秀);
- ③甲班成绩的波动比乙班大.

上述结论正确的是 (A)

- A.①②③ B.①② C.①③ D.②③

解析： $\because \overline{x_{甲}} = \overline{x_{乙}}$ ， \therefore ①正确； \because 乙的中位数为151，甲的中位数为149， \therefore 乙班优秀的人数多于甲班优秀的人数，②正确； $\because s_{甲}^2 > s_{乙}^2$ 甲班成绩的波动比乙班大，③正确. 故选A.

3. 一名学生军训时连续射靶10次,命中的环数分别为4,7,8,6,8,5,9,10,7,6.则这名学生射击环数的方差是3.

解析:数据4,7,8,6,8,5,9,10,7,6的平均数=

$$\frac{4+7+8+6+8+5+9+10+7+6}{10} = 7, \text{方差} =$$

$$\frac{1}{10} (9+0+1+1+1+4+4+9+0+1) = 3. \text{故填} 3.$$

4.某次射击练习,甲、乙二人各射靶5次,命中的环数如下表:

甲命中环数	7	8	6	8	6
乙命中环数	9	5	6	7	8

那么射击比较稳定的是甲.

解析:根据题意得 $\bar{x}_{甲} = (7+8+6+8+6) \div 5 = 7$,

$$\bar{x}_{乙} = (9+5+6+7+8) \div 5 = 7,$$

$$s_{甲}^2 = \frac{1}{5} [(7-7)^2 + (8-7)^2 + (6-7)^2 + (8-7)^2 + (6-7)^2] = 0.8,$$

$$s_{乙}^2 = \frac{1}{5} [(9-7)^2 + (5-7)^2 + (6-7)^2 + (7-7)^2 + (8-7)^2] = 2,$$

$\because s_{甲}^2 < s_{乙}^2$, \therefore 射击成绩比较稳定的是甲.故填甲.

5. 已知两组数据:

甲: 9.9 10.3 9.8 10.1 10.4 10 9.8 9.7

乙: 10.2 10 9.5 10.3 10.5 9.6 9.8 10.1

分别计算这两组数据的方差, 并说明数据波动的大小.

$$\text{解: } \bar{x}_{\text{甲}} = \frac{1}{8} \times (9.9 + 10.3 + 9.8 + 10.1 + 10.4 + 10 + 9.8 + 9.7) = 10.$$

$$\bar{x}_{\text{乙}} = \frac{1}{8} \times (10.2 + 10 + 9.5 + 10.3 + 10.5 + 9.6 + 9.8 + 10.1) = 10.$$

$$\begin{aligned} s_{\text{甲}}^2 &= \frac{1}{8} \times [(9.9-10)^2 + (10.3-10)^2 + \dots + (9.7-10)^2] \\ &= \frac{1}{8} \times (0.01 + 0.09 + \dots + 0.09) \\ &= \frac{1}{8} \times 0.44 = 0.055, \end{aligned}$$

$$\begin{aligned} s_{\text{乙}}^2 &= \frac{1}{8} \times [(10.2-10)^2 + (10-10)^2 + \dots + (10.1-10)^2] \\ &= \frac{1}{8} \times (0.04 + 0 + \dots + 0.01) \\ &= \frac{1}{8} \times 0.84 = 0.105. \end{aligned}$$

因为 $s_{\text{甲}}^2 < s_{\text{乙}}^2$,所以乙组数据比甲组数据波动大. 