

平均数、方差与标准差

A bouquet of white and purple flowers in a glass vase, serving as a background for the text. The flowers are arranged in a vase, with white flowers at the top and purple flowers at the bottom. The background is a soft, light blue color with a subtle pattern of small white dots.

总体特征数：

在数学中，通常把能反映**总体**某种特征的量称为**总体特征数**

如何反映**总体**的特征数？

用样本的特征数估计总体的特征数！

情境引入一：

在利用**单摆检验重力加速度的实验**中，全班同学在相同的条件下**进行测试**，得到下列数据（单位： m/s^2 ）

9.62 9.54 9.78 9.94 10.01 9.66 9.88 9.68
10.32
9.76 9.45 9.99 9.81 9.56 9.78 9.72 9.93
9.94
9.65 9.79 9.42 9.68 9.70 9.84 9.90

怎样利用这些数据对重力加速度进行估计？

平 均 数

为什么呢
?

问题转化为:

实验结果测得一组数据为

用 **算术平均数** 作为重力加速度“最理想的”近似值，依据是什么呢？

处理实验数据的原则是使这个近似值与实验数据之间的离差（偏差）最小、设近似值为 x ，则它与 n 个实验值 a_i ($i=1, 2, 3, \dots, n$) 的离差分别为 $x-a_1, x-a_2, \dots, x-a_n$

称为这 n 个数的平均数或者均值

读作: a 平均

$$= \frac{a_1 + a_2 + \dots + a_n}{n} = \frac{1}{n} \sum_{i=1}^n a_i$$

平均数最能代表一个样本数据的集中趋势，也就是说它与样本数据的离差最小。

例1 某校高一年级的甲乙两个班级（均为50人）的数学成绩如下（总分150），试确定这次考试中，哪个班的数学成绩更好一些。

甲班

112 86 106 84 100 105 98 102 94 107
 87 112 94 94 99 90 120 98 95 119
 108 100 96 115 111 104 95 108 111 105
 104 107 119 107 93 102 98 112 112 99

乙班

92 102 93 84 94 94 100 90 84 114
 116 95 109 96 106 98 108 99 110 103
 94 98 105 101 115 104 112 101 113 96
 108 100 110 98 107 87 108 106 103 97
 107 106 111 121 97 107 114 122 101 107
 107 111 114 106 104 104 95 111 111 110

甲班均分

乙班均分

思考

某公司有**经理**1人，另有6名**管理人员**，5名**高级技工**，10名**工人**和10名**学徒**，**现**需要增加一名新工人。小张前来**应聘**，**经理说**：“我公司**报酬不错**，**平均工资**每月1695元。”小张工作几天后找到**经理说**：“你**欺骗**了我，我**问过**其他工人，每月一个人的**工资超过**1500元，平均月**工资**怎么能是1695元呢？”**经理**拿出如下表所示的**工资表说**：“你看，平均**周工资**就是1695元。”

某公司内部**结构**以及**工资**分布：

| 人员 | 经理 | 管理人员 | 技工 | 工人 | 学徒 | 合计 |
|-----|-------|------|------|------|-----|----|
| 月工资 | 11000 | 1250 | 1100 | 1000 | 500 | |
| 人数 | 1 | 6 | 5 | 10 | 10 | 23 |

在这个问题中，总体月平均数能客观地反映工人的月工资水平吗？为什么？

在这个问题中，总体月平均数能客观地反映工人的月工资水平吗？为什么？



总体月平均数不能反映工人的月工资水平，因为公司中少数人的月工资额与大多数的月工资额差别较大，这样导致平均中位数的偏差较大，所以月平均数不能反映这个公司工人的月工资水平，而应用中位数或众数来反映工人的月工资水平



加权平均值（用频率计算平均值）

一般地，若取值为 x_1, x_2, \dots, x_n ，出现的次数分别为 f_1, f_2, \dots, f_n ，设频率为 p_1, p_2, \dots, p_n

则其加权平均数为

$$x_1 p_1 + x_2 p_2 + \dots + x_n p_n$$

其中 $(p_1 + p_2 + \dots + p_n = 1)$



例3：由下表估计学生日平均睡眠时间

| 睡眠时间 | 人数 | 频率 |
|------------|-----|------|
| $[6, 6.5)$ | 5 | 0.05 |
| $[6.5, 7)$ | 17 | 0.17 |
| $[7, 7.5)$ | 33 | 0.33 |
| $[7.5, 8)$ | 37 | 0.37 |
| $[8, 8.5)$ | 6 | 0.06 |
| $[8.5, 9]$ | 2 | 0.02 |
| 合 计 | 100 | 1 |



例4：由某单位年收入表试估计该单位职工的平均年输入

| 收入范围 | | | 所占百分比 |
|-------|---|-------|-------|
| 10000 | 至 | 15000 | 10% |
| 15000 | 至 | 20000 | 15% |
| 20000 | 至 | 25000 | 20% |
| 25000 | 至 | 30000 | 25% |
| 30000 | 至 | 35000 | 15% |
| 35000 | 至 | 40000 | 10% |
| 40000 | 至 | 50000 | 5% |



平均数的计算方法:

(1)定义法: 已知 $x_1, x_2, x_3, \dots, x_n$ 为某样本的 n

个数据, 则这 n 个数据的平均数为: $\bar{x} =$

$$\frac{x_1 + x_2 + x_3 + \dots + x_n}{n}.$$

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/357156016002006136>