



浙江财经大学
Zhejiang University of Finance & Economics

算术平均数



授课教师： 洪兴建

浙江财经大学数据科学学院



算术平均数

引例

今后国家统计局不再计算和发布全国70个大中城市房价涨幅的平均数，重点发布各个城市不同对比基期的分类指数和总指数。

<http://finance.people.com.cn/GB/13936560.html>





算术平均数

引例

人民网北京2月16日电（记者 朱剑红）国家统计局16日公布，《住宅销售价格统计调查方案》（以下简称《新方案》）已于今年1月起开始实施，今后，国家统计局不再计算和发布全国70个大中城市房价涨幅的平均数，重点发布各个城市不同对比基期的分类指数和总指数。依据《新方案》统计的1月份房价数据将在2月18日向社会公布。据国家



- ◆ 怎么计算70个大中城市房价平均涨幅的？
- ◆ 为什么不再计算了？



算术平均数

计算方法

公式

算术平均数 = $\frac{\text{总体标志值 (数据) 总和}}{\text{总体单位的总数}}$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{\underbrace{1 + 1 + \cdots + 1}_{n \text{ 个 } 1}}$$

作用 消除个体标志值之间的差异，体现出总体的一般水平。



人均GDP是算术平均数吗？ ➤ 分子分母一一对应



算术平均数

加权算术平均数

■ 问题

按师生分组	身高 x (厘米)	人数 f (人)
老师	226	1
学生	170	49
合 计	—	50

问：师生的平均身高为多少？

$$\bar{x} = \frac{\sum x_i}{n} = \frac{226 + 170}{2} = 198 \quad \times$$

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{226 + 170 \times 49}{50} = 171.12 \quad \checkmark$$



算术平均数

加权算术平均数计算公式

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_n f_n}{f_1 + f_2 + \cdots + f_n} = \frac{\sum_{i=1}^n x_i f_i}{\sum_{i=1}^n f_i}$$

分组数据中， x 表示各组水平值， f 代表各组变量值出现的**频数**（frequency）。



算术平均数

例1

某厂职工按日产量分组后所得组距数列如下表所示，求该厂职工的平均日产量。

按日产量分组(千克)	工人数(人)	组中值 x_i	$x_i f_i$
60 以下	10	55	550
60 – 70	19	65	1235
70 – 80	50	75	3750
80 – 90	36	85	3060
90 – 100	27	95	2565
100 – 110	14	105	1470
110 以上	8	115	920
合 计	164	-	13550

$$\bar{x} = \frac{\sum x_i \cdot f_i}{\sum f_i}$$

$$= \frac{13550}{164} = 82.62(\text{千克})$$



算术平均数

注意点

1

算术平均数取决于**变量值(x)**和**频数 (f)**或**频率 (f / Σf)**。

2

权重一般是相对数

$$\frac{f_i}{\Sigma f_i}$$

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \sum \left(x_i \cdot \frac{f_i}{\sum f_i} \right)$$



例2



问题：某市2016年市中心和郊区的价格都比2015年上涨了，请问2016年的平均价格是否一定上涨？

按地区 分组	成交价(元/平方米)		成交面积(万平方米)	
	2015年	2016年	2015年	2016年
市区	15000	16000	1000	500
郊区	5000	6000	500	1000
合 计	—	—	1500	1500



算术平均数

例2计算

$$\bar{X}_{2015} = \frac{\sum x_i f_i}{\sum f_i} = \frac{15000 \times 10000000 + 5000 \times 5000000}{15000000}$$

$$= 11667(\text{元} / \text{平方米})$$

$$\bar{X}_{2016} = \frac{\sum x_i f_i}{\sum f_i} = \frac{16000 \times 5000000 + 6000 \times 10000000}{15000000}$$

$$= 9333(\text{元} / \text{平方米})$$

- 结论：市区、郊区价格都上涨了，平均价格未必上涨。
- 问题：结构优化的含义是什么？

增大有利部分的比重，降低不利部分的比重。



算术平均数

性质

(1) 离差之和等于零 $\sum_i (x_i - \bar{x}) = 0$

推导
$$\sum_i (x_i - \bar{x}) = \sum_i x_i - n\bar{x} = \sum_i x_i - n \frac{\sum_i x_i}{n} = 0$$

(2) 离差平方和最小 $\sum (x_i - \bar{x})^2$ 的值最小

设c为任意实数
$$\sum (x_i - \bar{x})^2 \leq \sum (x_i - c)^2$$

推导
$$\begin{aligned} \sum (x_i - \bar{x})^2 &= \sum ((x_i - c) - (\bar{x} - c))^2 = \sum ((x_i - c)^2 - 2(x_i - c)(\bar{x} - c) + (\bar{x} - c)^2) \\ &= \sum (x_i - c)^2 - 2(\bar{x} - c) \sum (x_i - c) + \sum (\bar{x} - c)^2 = \sum (x_i - c)^2 - n(\bar{x} - c)^2 \leq \sum (x_i - c)^2 \end{aligned}$$



算术平均数

优缺点

■ 优点

- 推算总体标志总量
- 进行代数运算
- 抽样中具有良好的稳定性和可靠性

■ 缺点

- 受极值影响较大
张庄有个张千万，两个邻居穷光蛋
加在一起求平均，个个收入超百万