



第二十一章

数据特征的测度



本章内容介绍

第二十一章 数据特征的 测度

第一节集中趋势的测度

众数、中位数

算术平均数、几何平均数

第二节离散程度的测度

极差、标准差和方差

离散系数



本章内容介绍

原始数据经过整理与显示，达到了一定程度的概括。对统计数据特征的测度，主要从三个方面进行：

集中趋势

- 反映各数据向其中心值靠拢或聚集的程度

离散程度

- 各数据远离其中心值的趋势

偏态和峰度

- 数据分布的形状



第一节

集中趋势的测度



第一节 集中趋势的测度

【知识点一】众数、中位数★★★

1. 众数：一组数据中出现频数最多的那个数值，用 M_0 表示。

【例如】一家连锁超市的10个分店某月的销售额（单位：万元）分别为：61 65 73 78 80 80 80 80 96 97。

这10个分店月销售额的众数为 $M_0=80$ （万元）

【总结】众数既适用于品质数据，也适用于数值型数据。

众数是一个位置代表值，不受极端值的影响，抗干扰性强。



第一节 集中趋势的测度

2. 中位数：把一组数据按从小到大的顺序进行排列，位置居中的数值叫作中位数，用 M_e 表示。

$$M_e = \begin{cases} X_{(\frac{n+1}{2})} & \text{当 } n \text{ 为奇数时} \\ \frac{1}{2}(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}) & \text{当 } n \text{ 为偶数时} \end{cases}$$



第一节 集中趋势的测度

【例如】某地级市下辖9个县，每个县的面积如下（单位：平方公里），计算该市下辖县面积的中位数：

1455 2019 912 1016 1352 1031 2128 1075 2000

首先，将上面的数据排序：912 1016 1031 1075 1352

1455 2000 2019 2128

中位数位置 = $(9+1) \div 2 = 5$ ，中位数为1352，即 $M_e = 1352$ （平方公里）

【总结】中位数主要适用于顺序数据，也适用于数值型数据。中位数也是一个位置代表值，不受极端值的影响，抗干扰性强。



真题回顾

【真题·单选】2020年某快餐品牌旗下10个加盟店的年营业收入（单位：万元）分别为5600、5600、5600、5600、5600、5800、7200、7800、7900、8900，这组数据的中位数是（ ）。

- A. 6000
- B. 5700
- C. 5800
- D. 5600



真题回顾

答案：B

解析：中位数位置 = $(10+1) \div 2 = 5.5$ ，中位数 = $(5600+5800) \div 2 = 5700$ 。



第一节 集中趋势的测度

【知识点二】算术平均数★★

1. 算术平均数：是全部数据的算术平均，又称均值，用 \bar{X} 表示。算术平均数是集中趋势最主要的测度值，主要适用于数值型数据，但不适用于品质数据。

简单算术平均数

- 用于处理未分组的原始数据

- 计算公式： $\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$

加权算术平均数

- 用于处理经分组整理的数据

- 计算公式： $\bar{X} = \frac{X_1 f_1 + X_2 f_2 + \dots + X_k f_k}{f_1 + f_2 + \dots + f_k} = \frac{\sum_{i=1}^k X_i f_i}{\sum_{i=1}^k f_i}$



第一节 集中趋势的测度

【例如·简单算术平均数】某售货小组有5名营业员，元旦一天的销售额分别为520元、600元、480元、750元和500元，求该日5名营业员的平均销售额。

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$

$$\bar{X} = \frac{520 + 600 + 480 + 750 + 500}{5} = 570 \text{ (元)}$$



第一节 集中趋势的测度

【例如·加权算术平均数】某市商业企业协会根据100个会员样本，整理出一年销售额分布资料：

销售额（万元）	组中值	商业企业数 f_i	$x_i f_i$
100-150	125	4	500
150-200	175	16	2800
200-250	225	40	9000
250-300	275	28	7700
300-350	325	10	3250
350-400	375	2	750
合计	—	100	24000



第一节 集中趋势的测度

计算年平均销售额。

$$\bar{X} = \frac{\sum_{i=1}^k X_i f_i}{\sum_{i=1}^k f_i} = \frac{24000}{100} = 240 \text{ (万元)}$$

3. 注意事项:

- (1) 算术平均数的影响因素: 各组数值的大小和各组分布频数的多少。
- (2) 算术平均数易受极端值的影响。



第一节 集中趋势的测度

【知识点三】几何平均数★

1. 几何平均数：n个观察值连乘积的n次方根
2. 简单几何平均数的计算。

设一组数据为 X_1, X_2, \dots, X_n ，且大于0， \bar{X}_G 表示几何平均数，则：

$$\bar{X}_G = \sqrt[n]{X_1 \times X_2 \times \dots \times X_n} = \sqrt[n]{\prod_{i=1}^n X_i}$$

式中， \prod 为连乘积符号。



第一节 集中趋势的测度

【例如】某型号钻头的生产，需经过6道不同的加工工序，各道工序的合格率如下表所示，计算平均合格率。

用几何平均数的方法进行计算，得：

$$\bar{X}_G = \sqrt[n]{\prod_{i=1}^n X_i} = \sqrt[6]{98.2\% \times 97.5\% \times \dots \times 95.5\%} = 96.63\%$$



第一节 集中趋势的测度

计算几何平均数要求各观察值之间存在连乘积关系，它的主要用途是：

- (1) 对比率、指数等进行平均。(2) 计算平均发展速度。

【总结】集中趋势指标特点总结

	指标	极端值	品质数据		数值型数据
			分类数据	顺序数据	
位置平均数	众数	不受影响	√	√	√
	中位数			√	√
数值平均数	算术平均数	受影响			√
	几何平均数				连乘积关系的适用



真题回顾

【真题·多选】对分组数据计算加权算术平均数时，其平均数值会受到（ ）等因素的影响。

- A. 组内极差
- B. 极端值
- C. 组内标准差
- D. 各组数值大小
- E. 各组频数多少



真题回顾

答案：BDE

解析：本题考查算术平均数的相关知识。计算和运用算术平均数须注意：（1）算术平均数同时受到两个因素的影响：各组数值的大小；各组分布频数的多少。（2）算术平均数易受极端值的影响。