



第二十一章

数据特征的测度



第二十一章 数据特征的测度

单选题	多选题	合计
1题1分	2题4分	3题5分



第二十一章 数据特征的测度

一、集中趋势的测度★★★

二、离散程度的测度★★



一、集中趋势的测度



一、集中趋势的测度

知识点：众数★★★

知识点：中位数★★★★

知识点：算术平均数★★

知识点：几何平均数★



一、集中趋势的测度

分布的集中趋势	反映各数据向其中心值靠拢或聚集的程度
分布的离散程度	反映各数据远离其中心值的趋势
分布的偏态和峰度	反映数据分布的形状



一、集中趋势的测度

集中趋势：一组数据向某一中心值靠拢的倾向，是寻找数据一般水平的代表值或中心值。主要包括**位置平均数**和**数值平均数**。

位置平均数：按数据的大小顺序或出现频数的多少确定集中趋势的代表值，有众数、中位数等；

数值平均数：根据全部数据计算出来的平均数，有算术平均数、几何平均数等。



一、集中趋势的测度

知识点：众数★★★

众数是一组数据中**出现频数最多**的那个数值，用 M_0 表示。

【举例】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）



一、集中趋势的测度

众数反映集中趋势，非常直观，**不仅适用于品质数据，也适用于数值型数据。**

众数是一个**位置代表值**，**不受极端值**（明显比其他数据大得多或小得多的数值）**的影响**，抗干扰性强。



一、集中趋势的测度

【例题·单选】下列对数据特征的测度中，既适用于品质数据也适用于数值型数据的是（ ）。

- A. 离散系数
- B. 众数
- C. 算术平均数
- D. 方差



一、集中趋势的测度

答案：B

解析：用众数反映集中趋势非常直观，不仅适用于品质数据，也适用于数值型数据。



一、集中趋势的测度

知识点：中位数★★★

把一级数据按从**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。

中位数将数据分成两部分，其中一半的数据小于中位数，另一半数据大于中位数。



一、集中趋势的测度

中位数位置 = $(n+1) / 2$ ，式中 n 为数据的个数。最后确定中位数的具体数值。

中位数为：

$$M_e = \begin{cases} X_{(\frac{n+1}{2})} \\ \frac{1}{2} [X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}] \end{cases}$$

当 n 为奇数时

当 n 为偶数时



一、集中趋势的测度

【举例】某地级市下辖9个县，每个县的面积如下（单位：平方千米），计算该市下辖县面积的中位数：

1455 2019 912 1016 1352 1031 2128 1075 2000

首先，将上面的数据排序结果如下：

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 $M_e = 1352$ （平方千米）



一、集中趋势的测度

若行政区划调整，临市的一个面积为1000平方千米的县划归该市。行政区划调整后，该市现在下辖10个县，该市下辖县的面积（单位：平方千米）从小到大依次为：

912 1000 1016 1031 1075 1352 1455 2000 2019
2128

计算行政区划调整后该市下辖县面积的中位数：

中位数位置= $(10+1) \div 2 = 5.5$,

$M_e = (1075 + 1352) / 2 = 1213.5$ （平方千米）



一、集中趋势的测度

中位数主要用于顺序数据，也适用于数值型数据，但不适用于分类数据。

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。



一、集中趋势的测度

【真题·2022年单选】2021年市下辖10个县区的参加基本养老保险人数（单位：万人）分别为：100、120、122、123、124、126、133、133、133、133，这组数据的中位数是（ ）万人。

A.126

B.125

C.124

D.123



一、集中趋势的测度

答案：B

解析：中位数位置 = $(10+1) / 2 = 5.5$ ，中位数 = $(124+126) / 2 = 125$ 。



一、集中趋势的测度

知识点：算术平均数★★

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，是进行统计分析和统计推断的基础。

算术平均数**主要适用于数值型数据**，但不适用于品质数据。



一、集中趋势的测度

1.简单算术平均数，主要用于处理未分组的原始数据。

设一组数据为 $X_1, X_2, \dots, X_n$ ，简单算术平均数的计算公式为：

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



一、集中趋势的测度

【举例】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，求该日平均销售额。

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

计算结果表明，元旦5名营业员的平均销售额为570元。



一、集中趋势的测度

2. 加权算术平均数，主要用于处理已分组整理的数据。

设原始数据被分成k组，各组的组中值为 $X_1, X_2, \dots, X_k$ ，各组的频数分别 $f_1, f_2, \dots, f_k$ ，加权算术平均数的计算公式为：

$$\bar{X} = \frac{X_1 f_1 + X_2 f_2 + \dots + X_n f_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum_{i=1}^k X_i f_i}{\sum_{i=1}^k f_i}$$



一、集中趋势的测度

【举例】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 X_i	商业企业数 f_i	$X_i f_i$
100~150	125	4	500
150~200	175	16	2 800
200~250	225	40	9 000
250~300	275	28	7 700
300~350	325	10	3 250
350~400	375	2	750
合计	—	100	24 000

计算年平均销售额。



一、集中趋势的测度

首先确定组中值，计算结果在上表中列出。代入公式得：

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

计算结果表明，100个商业企业的年均销售额为240万元。



一、集中趋势的测度

计算和运用算术平均数需注意：

第一，算术平均数同时受到两个因素的影响，一个是**各组数值**的大小，另一个是**各组分布频数**的多少。

第二，算术平均数**易受极端值**的影响。



一、集中趋势的测度

【例题·多选】影响加权算术平均数大小的因素包括（ ）。

- A. 与各组变量值大小无关
- B. 与各组频率多少次无关
- C. 各组变量值的大小
- D. 各组频率的多少
- E. 极端值



一、集中趋势的测度

答案：CDE

解析：计算和运用算术平均数（包括简单算术平均数和加权算术平均数）需要注意：（1）算术平均数同时受到两个因素的影响，一个是各组数值的大小，另一个是各组分布频数的多少。在数值不变的情况下，哪一组的频数多，该组的数值对平均数的作用就大；反之哪一组的频数少，该组数值对平均数的影响就小。（2）算术平均数易受极端值的影响。



一、集中趋势的测度

知识点：几何平均数★

n 个观察值连乘积的 n 次方根就是几何平均数。

几何平均数也有加权和不加权之分。这里只介绍简单几何平均数的计算。



一、集中趋势的测度

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, Π 为连乘积符号。



一、集中趋势的测度

【举例】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表所示，计算平均合格率。

工序名称	合格率
冲料	98.2
料废	97.5
车工	97.0
加热	96.6
扫槽	95.5
接柄	95.0



一、集中趋势的测度

用几何平均数的方法进行计算，得：

$$\overline{X}_G = \sqrt[n]{\prod_{i=1}^n X_i} = \sqrt[6]{98.2\% \times 97.5\% \cdots \times 95.0\%} \approx 96.63\%$$

各道工序的平均合格率为96.63%。

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。



一、集中趋势的测度

【例题·多选】下列统计指标中，应采用几何平均数方法计算平均数的有（ ）。

- A. 发展速度
- B. 营业收入
- C. 投资完成额
- D. 销售收入
- E. 产品合格率



一、集中趋势的测度

答案：AE

解析：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。



一、集中趋势的测度

众数	适用于品质数据，也适用于数值型数据。 是位置代表值，不受极端值的影响。
中位数	适用于顺序数据，也适用于数值型数据，但不适用于分类数据。 是位置代表值，不受极端值的影响。
算术平均数	适用于数值型数据，不适用于品质数据。 简单算术平均数、加权算术平均数。
几何平均数	对比率、指数等进行平均；计算平均发展速度。