



主讲老师：徐娜

中级经济师
经济基础知识
课程精讲班



第二十四章

描述统计



目录

第二十四章描述统计

第一节集中趋势的测度

1. 均值
2. 中位数
3. 众数
4. 均值、中位数和众数的比较及适用范围

第二节离散程度的测度

1. 方差
2. 标准差
3. 离散系数

第三节分布形态的测度

1. 偏态系数
2. 标准分数

第四节变量间的相关分析

1. 变量间的相关关系
2. 散点图
3. 相关系数



第一节 集中趋势的测度

一、数据特征

1. 集中趋势：反映各数据向其中心值靠拢或聚集的程度。
2. 离散程度：反映各数据之间的差异程度，也能反映中心值对数据的代表程度。
3. 分布的偏态：反映数据分布的不对称性。



第一节 集中趋势的测度

二、集中趋势的测度

均值

均值也叫作平均数，就是数据组中所有数值的总和除以该组数值的个数。设一组数据为 $X_1, X_2, X_3, \dots, X_n$ ，平均数 \bar{X}

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

均值是集中趋势最主要的测度值，它是一组数据的重心所在，解释了一组数据的平均水平。它主要适用于数值型数据，但不适用于分类和顺序数据，此外，均值容易受到极端值的影响，极端值会使得均值向极大值或极小值方向倾斜，使得均值对数据组的代表性减弱。



第一节 集中趋势的测度

二、集中趋势的测度

中位数

把一组数据按从小到大或从大到小的顺序进行排列，位置居中的数值叫作中位数。中位数是位置代表值，主要用于顺序数据和数值型数据，但不适用于分类数据。中位数的优点是不受极端值的影响，抗干扰性强，尤其适于收入这类偏斜分布的数值型数据。

众数

众数是指一组数据中出现次数（频数）最多的变量值。众数适用于描述分类数据和顺序数据的集中趋势，不适用于描述定量数据。而且有些情况下可能出现双众数、多众数或者没有众数，难以描述数据的集中位置。



第一节 集中趋势的测度

三、集中趋势指标区别

指标	计算	适用数据类型	是否受极端值影响	是否利用数据全部信息	稳定性
均值 (最主要)	算术平均数	数值型数据	受	是	最好
中位数	中间位置	顺序数据/ 数值型数据	不受	否	居中
众数	出现次数最多	分类数据/ 顺序数据	不受	否 (不唯一)	最差



典型真题

【真题·2023单选】下列变量的观测数据中，适合用均值测度集中趋势的是（ ）。

- A. 企业类型
- B. 工人工龄
- C. 地域城乡属性
- D. 产业类别



典型真题

答案：B

解析：均值是集中趋势最主要的测度值，它是一组数据的重心所在，解释了一组数据的平均水平。它主要适用于数值型数据，但不适用于分类数据和顺序数据。工人工龄属于数值型数据，因此适合用均值测度集中趋势。



典型真题

【真题·2023单选】某省下辖6个地级市60岁以上人口数（单位：万人）分别：32、38、48、52、73、93，中位数是（ ）万人

A. 50

B. 48

C. 52

D. 56



典型真题

答案：A

解析：把一组数据按从小到大或从大到小的顺序进行排列，位置居中的数值叫作中位数，用M表示。中位数将数据分成两部分，其中一半数据小于中位数，另一半数据大于中位数。

中位数 = $(48+52) / 2 = 50$ 。



第二节 离散程度的测度

一、方差

方差是数据组中各数值与其均值离差平方的平均数。方差越小，说明数据值与均值的平均距离越小，均值的代表性越好。方差的单位是原数据的平方。

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n - 1}$$



第二节 离散程度的测度

二、标准差

标准差即方差的平方根。

1. 不仅能度量数值与均值的平均距离，还与原始数值具有相同的计量单位

2. 标准差越小，说明数据值与均值的平均距离越小，均值的代表性越好

3. 标准差的大小不仅与数据的计量单位有关，也与观测值的均值大小有关，不能直接用标准差比较不同变量的离散程

度

$$s = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n - 1}}$$



第二节 离散程度的测度

三、离散系数

也称为变异系数或标准差系数，即标准差与均值的比值，主要用于不同类别数据离散程度的比较，记为CV。

离散系数消除了测度单位和观测值水平不同的影响，可以直接用来比较变量的离散程度。

$$CV = \frac{S}{\bar{X}}$$



典型真题

【真题·2023单选】一组数据的方差为100、均值为40，则这组数据的离散系数为（ ）。

- A. 4
- B. 0.25
- C. 2.5
- D. 0.4



典型真题

答案：B

解析：离散系数也称为变异系数或标准差系，即标准差与均值的比值。 $CV=10/40=0.25$ 。