



浙江财经大学
Zhejiang University of Finance & Economics

位置平均数



授课教师： 洪兴建

浙江财经大学数据科学学院



位置平均数

引例

3-2 位置平均数 讲课.pptx

http://news.sohu.com/20120122/n332877554.shtml

2012年01月22日08:48 中国广播网 来源: 中国广播网

正文

2012年01月22日08:48 中国广播网 来源: 中国广播网

中广网北京1月22日消息(记者张棉棉)据中国之声《新闻纵横》报道,国家统计局昨天首次公布了城乡居民收入的中位数。国家统计局报告显示,2011年农村居民人均纯收入中位数为6194元,2011年城镇居民人均可支配收入中位数为19118元,两者均较上年有所增长。

在国家统计局最新公布的《2011年城乡居民收入增长情况》报告中,首次提出了“中位数”的概念,那么,究竟什么是中位数?它相比过去公布的平均数又有哪些差别?

在过去,居民收入的数据一公布,很多人就会说自己又“被增长”了。其实,数字并没有水分,“被增长”原因之一是“平均数”,就现实国情而言,国内收入结构并非理想的中产占多数的

女清洁工拍写真被赞像闫妮 老公... 2162627

战斗民族的客机! 冲开特大浓雾降... 2056517

12岁女孩扶起跌倒大妈 被恐吓... 1580331

希拉里计划 美国做好准备迎接... 1580331

比例较高的中低收入者自然不会关

22:29 2017/8/19



位置平均数

定义

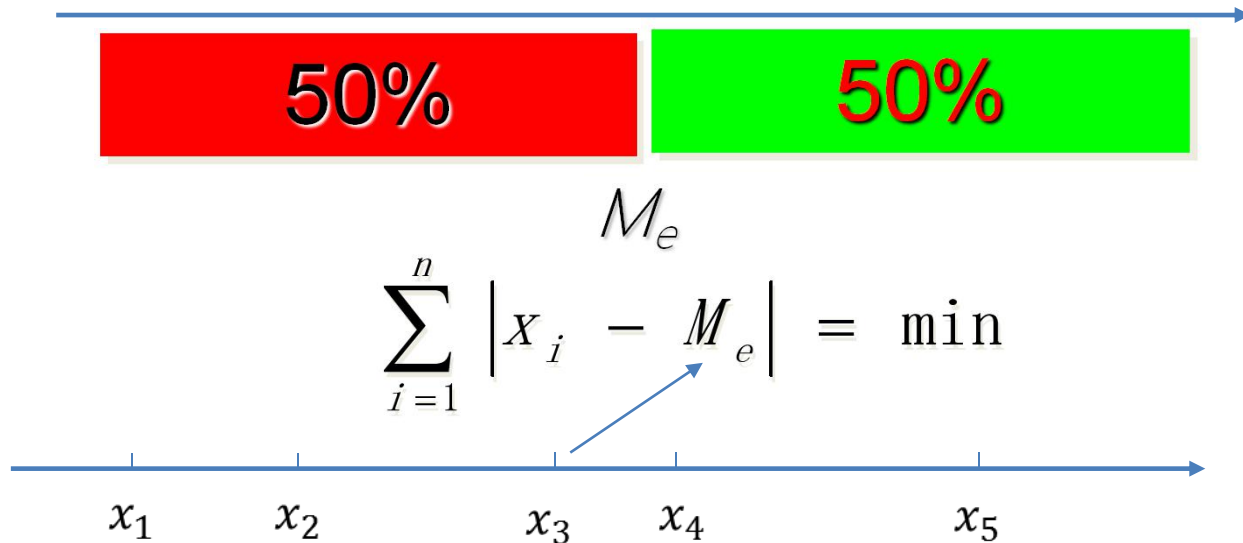
- 特殊位置上的数据作为代表值。
- 常用的位置平均数有中位数、众数。



中位数与分位数

中位数 (Median) 的含义

- 将数据按大小顺序排列，居于**中间位置**的数据。





中位数与分位数

中位数的计算：原始数据

- 对于未分组的原始资料：

$$x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_n$$

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



中位数与分位数

中位数的计算: 单项数列

例1: 某车间50名工人日加工零件数分组表

零件数 (个)	频数 (人)	零件数 (个)	频数 (人)	零件数 (个)	频数 (人)
107	1	119	1	128	2
108	2	120	2	129	1
110	1	121	1	130	1
112	2	122	4	131	1
113	1	123	4	133	2
114	1	124	3	134	2
115	1	125	2	135	1
117	3	126	2	137	1
118	3	127	3	139	2



中位数与分位数

中位数的计算: 单项数列

零件数	频数	向上累计数	零件数	频数	向上累计数	零件数	频数	向上累计数
107	1	1	119	1	16	128	2	39
108	2	3	120	2	18	129	1	40
110	1	4	121	1	19	130	1	41
112	2	6	122	4	23	131	1	42
113	1	7	123	4	27	133	2	44
114	1	8	124	3	30	134	2	46
115	1	9	125	2	32	135	1	47
117	3	12	126	2	34	137	1	48
118	3	15	127	3	37	139	2	50



中位数与分位数

中位数的计算

- 单项数列，直接按 $\frac{\sum f}{2}$ 求出中位数所在组的位置，组值即是中位数。



中位数与分位数

中位数的计算: 组距数列

例2:某企业50名工人加工零件如下表, 计算50名工人日加工零件数的中位数

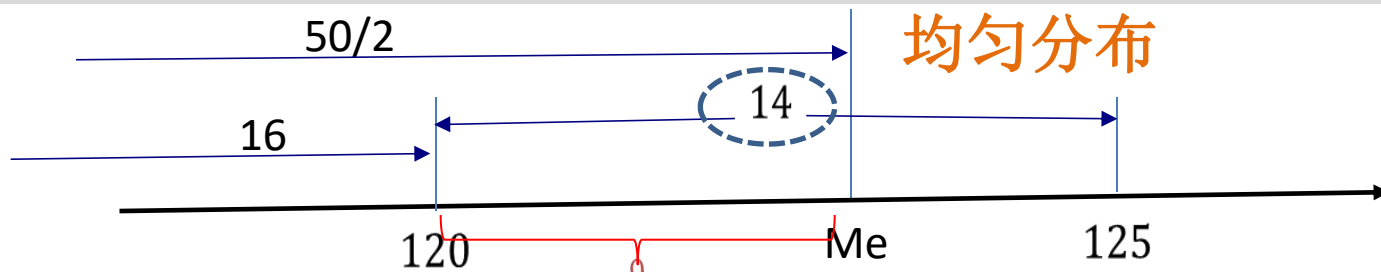
按零件数分组(个)	频数(人)	向上累计(人)	向下累计(人)
105~110	3	3	50
110~115	5	8	47
115~120	8	16	42
120~125	14	30	34
125~130	10	40	20
130~135	6	46	10
135~140	4	50	4





中位数与分位数

中位数



$$\frac{\frac{50}{2} - 16}{14} \times 5$$

$$M_e = 120 + \frac{\frac{50}{2} - 16}{14} \times 5 = 123.21 (\text{件})$$



中位数与分位数

中位数的计算: 组距数列

- 由组距数列确定中位数，先按 $\frac{\sum f}{2}$ 求出中位数所在组的位置，然后再按下限公式或上限公式确定中位数。

$$M_e = L + \frac{\frac{\sum f}{2} - S_{m-1}}{f_m} \times i$$

$$M_e = U - \frac{\frac{\sum f}{2} - S_{m+1}}{f_m} \times i$$

L、U分别为中位数所在组的下限和上限
Sm-1为中位数所在组以下的累计次数。
Sm+1为中位数所在组以上的累计次数。
fm为中位数所在组的次数



中位数

中位数特点

- 不受极值影响
- 缺乏敏感性

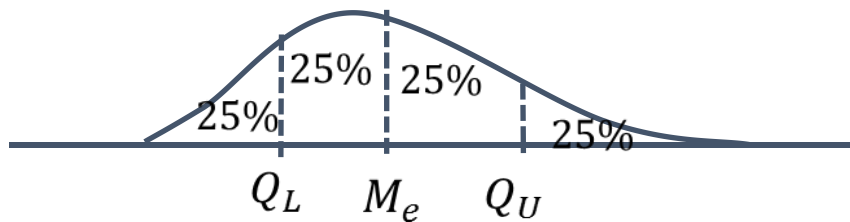


中位数与分位数

分位数

- 处于等分点位置的数值
- 常用的有四分位数(Quartile)、十分位数和百分位数

- 四分位数有几个?
- 四分位数的公式:



序号

$$Q_L = \frac{n+1}{4}$$

$$Q_m = M_e = \frac{2(n+1)}{4}$$

$$Q_U = \frac{3(n+1)}{4}$$



定义

问题：鞋店老板进货时，35-45码的进货数量相等吗？

□ Mode：总体中出现次数最多的标志值



众数

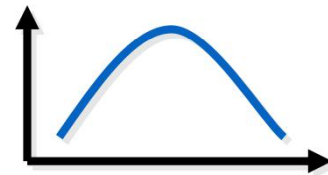
众数的计算

离散型数据的众数

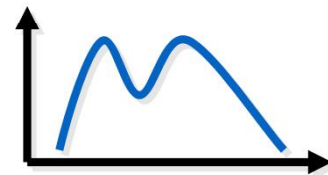
无众数: 10 5 9 12 6 8



一个众数: 6 5 9 8 5 5



多个众数: 25 28 28 36 42 42



众数的计算

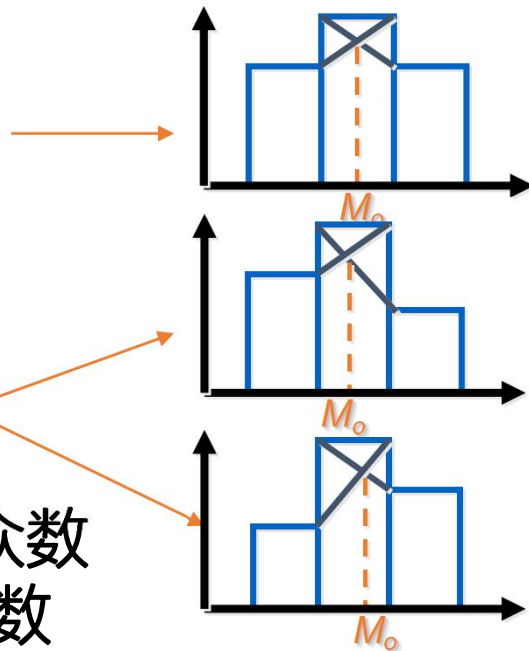
数值型分组数据的众数

➤ 相邻两组的频数相等时，组中值

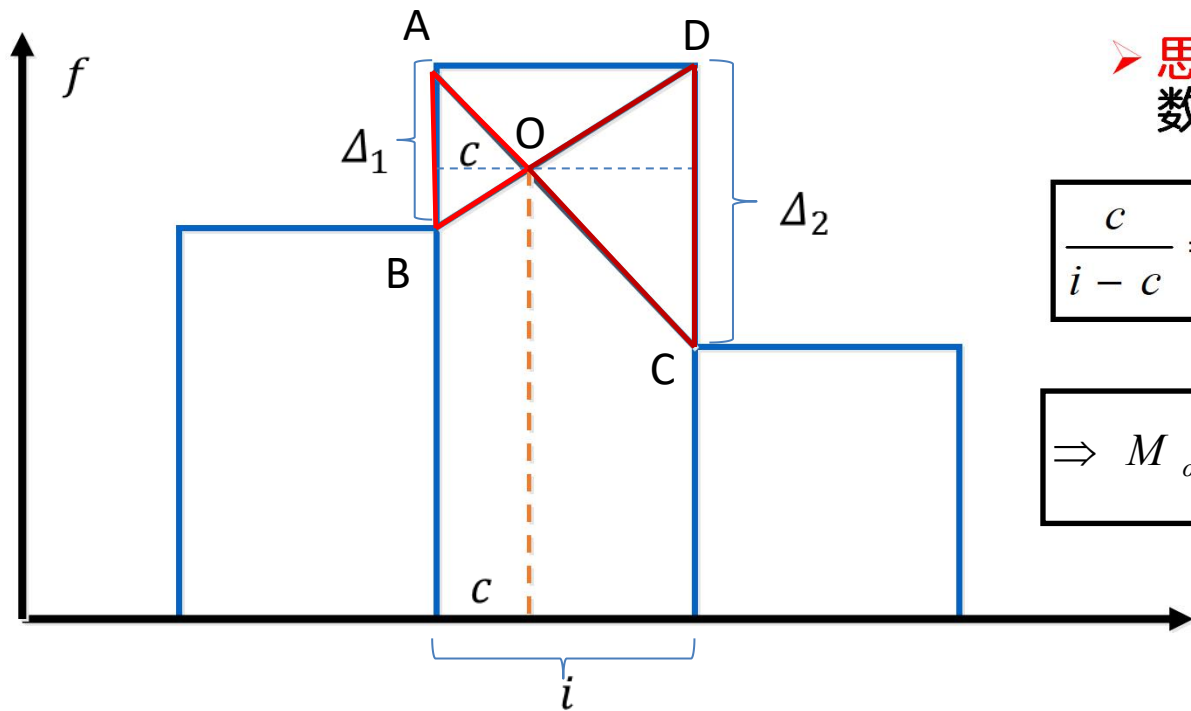
➤ 相邻两组的频数不等时，近似有：

$$M_o = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = U - \frac{\Delta_2}{\Delta_1 + \Delta_2} \times i$$

Δ_1 为众数组次数减去前一组次数， Δ_2 为众数组次数减去后一组次数， i 为众数组次数



众数的计算



➤ **思想：**渐进变化，众数靠近次数多的组

$$\frac{c}{i - c} = \frac{\Delta_1}{\Delta_2} \Rightarrow c = \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

$$\Rightarrow M_o = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$



众数的计算

例3:

某车间50名工人日加工零件数分组表		
按零件数分组	频数（人）	累积频数
105~110	3	3
110~115	5	8
115~120	8	16
120~125	14	30
125~130	10	40
130~135	6	46
135~140	4	50
合计	50	—

$$M_0 = 120 + \frac{14 - 8}{(14 - 8) + (14 - 10)} \times 5 = 123(\text{个})$$



众数

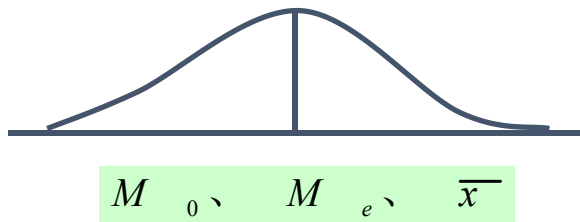
众数的特点

- 不受极值影响
- 均匀分布无众数
- 众数偏向次数较多的组
- 缺乏敏感性



中位数、众数和算术平均数的关系

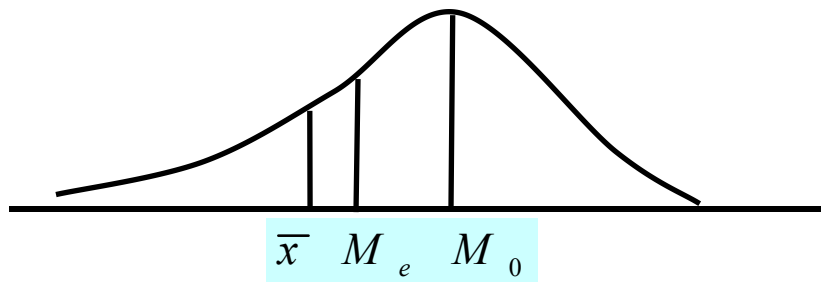
1. 对称分布（正态分布）时



对称分布

$$\bar{x} = M_e = M_o$$

2. 左偏分布时



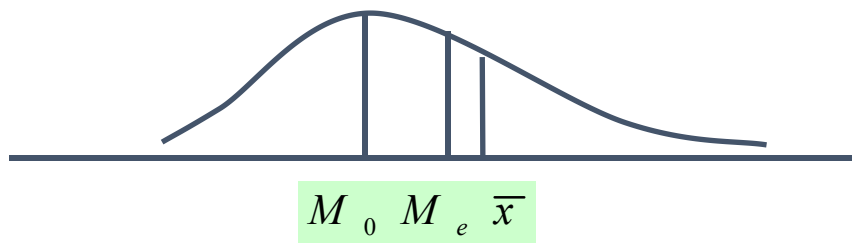
左偏分布

$$\bar{x} < M_e < M_o$$



中位数、众数和算术平均数的关系

3. 右偏分布时



右偏分布

$$M_o > M_e > \bar{x}$$

4. 适度偏态时

$$M_o - \bar{x} = 3(M_e - \bar{x})$$

皮尔逊经验：众数与算术平均数的距离约为中位数与算术平均数距离的3倍



中位数、众数和算术平均数的关系

例4：一组技术人员月薪的众数为7000元，算术平均数为10000元，**适度偏斜**时中位数近似值是多少？

$$M_o - \bar{x} = 3(M_e - \bar{x})$$

$$M_e = \frac{1}{3}(M_0 + 2\bar{X}) = \frac{1}{3}(7000 + 2 \times 10000) = 9000(\text{元})$$

$$\because \bar{X} > M_e > M_0 \quad \therefore \text{分布右偏。}$$