



第二节

数值型数据的整理与显示



第二节 数值型数据的整理与显示

【知识点一】数据的分组★★★

品质数据的整理与图示方法，也都适用于对数值型数据的整理与显示。但是数值型数据一些特定的整理和图示方法，并不适用于品质数据。



第二节 数值型数据的整理与显示

（一）数据的分组

1. 数据分组

单变量值分组

- 把每一个变量值作为一组
- 适合：离散变量且变量值较少

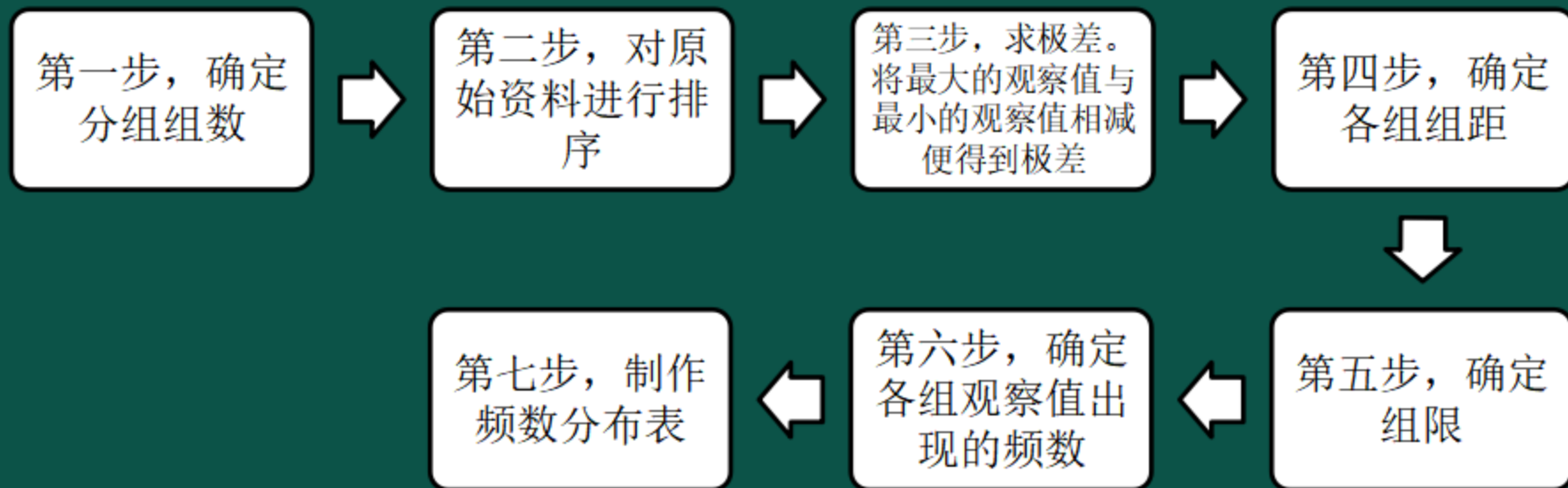
组距分组

- 将全部变量值依次划分为若干个区间，并将这一区间的变量值作为一组
- 适用：连续变量或变量值较多



第二节 数值型数据的整理与显示

2. 组距分组的步骤





第二节 数值型数据的整理与显示

【例题】某高中一年级一班共有55名学生，高一语文考试中成绩分别为：

59 73 87 65 89 85 77 94 69 97

56 80 68 95 96 50 63 88 91 90

96 92 93 79 74 65 74 89 83 51

74 79 94 67 92 92 93 70 87 86

54 87 86 54 62 76 86 73 86 70

100 110 108 102 112



第二节 数值型数据的整理与显示

(1) 确定分组组数

【注意】其要求为：

- ①划分的组数不太多也不太少。
- ②组数的确定要保证组间资料的差异性与组内资料的同质性。
- ③采用的分组办法，要能够充分显示客观现象本身存在的状态。

$$K = 1 + \frac{\log_{10} N}{\log_{10} 2} = 1 + \frac{\log_{10} 55}{\log_{10} 2} = 1 + 5.78 \approx 7$$



第二节 数值型数据的整理与显示

(2) 成绩按照从低到高排序结果如下:

50 51 54 54 56 59 62 63 65 65

67 68 69 70 70 73 73 74 74 74

76 77 79 79 80 83 85 86 86 86

86 87 87 87 88 89 89 90 91 92

92 92 93 93 94 94 95 96 96 97

100 102 108 110 112



第二节 数值型数据的整理与显示

(3) 求极差

极差值 = 最大观察值 - 最小观察值

极差 = (最大的观察值 - 最小的观察值) = $112 - 50 = 62$

(4) 确定各组组距

【方法】

(1) 组距 = 某组上限值 - 该组下限值

(2) 等距分组的情况下，最好把组距取成接近于能被5除尽的一个数。

(3) 关系：组距与组数成反比关系。

(4) 定义：组距是每组观察值的最大差。

组距 = $62 \div 7 \approx 8.9$ ，组距可取10。



第二节 数值型数据的整理与显示

(5) 确定组限

【注意】确定组限时应注意的问题：

第一，第一组的下限值应比最小的观察值小一点，最后一组的上限值应比最大的观察值大一点。

第二，特别需要或不得已的情况除外，最好不要使用开口组。

第三，组限应取得美观些，按数字偏好，组限值应能被5除尽，且一般要用整数表示。

①下限和上限：成绩50-60组的下限是50（最小值），上限是60（最大值）。

②成绩50-60组的组距 $=60-50=10$ 。组距是每组观察值的最大差。



第二节 数值型数据的整理与显示

(6) 确定各组观察值出现的频数。

① “不重不漏”的原则。

不重：一项观察值只能分在其中的某一组，不能在其他组重复出现

不漏：在所分的全部组别中每项数据都能分在其中的某一组，不能遗漏

② “上组限不在内”。即当相邻两组的上下限重叠时，恰好等于某一组上限的观察值不算在本组内，而计算在下一组内。



第二节 数值型数据的整理与显示

(7) 制作频数分布表。如下表，并填上相关的内容，以及其他需要说明的事项。

观察值区间	组中值	频数	频率 (%)
50-60	55	6	10.9
60-70	65	7	12.7
70-80	75	11	20.0
80-90	85	13	23.6
90-100	95	13	23.6
100-110	105	3	5.5
110-120	115	2	3.7
合计	—	55	100.0



真题回顾

【真题·单选】一组数据分为“0~1000”，“1000~2000”，“2000~3000”，“3000~4000”，“4000~5000”，根据统计分组习惯规定，“2000”应计入（ ）组内。

A. 1000~2000

B. 2000~3000

C. 3000~4000

D. 4000~5000



真题回顾

答案：B

解析：本题考查组距式分组。根据“上组限不在内”的原则，2000应该计入在2000~3000这一组中。



真题回顾

【真题·单选】关于数据分组的说法，正确的是（ ）。

- A. 组距分组首先要确定组限
- B. 单变量值分组不适合离散变量并且变量值较少的情况
- C. 组距分组不需考虑组间资料的差距性
- D. 数据分组适用于数据型数据



真题回顾

答案：D

解析：品质数据的整理与图示方法，也都适用于对数值型数据的整理与显示。但是数值型数据一些特定的整理和图示方法，并不适用于品质数据，如数据分组。

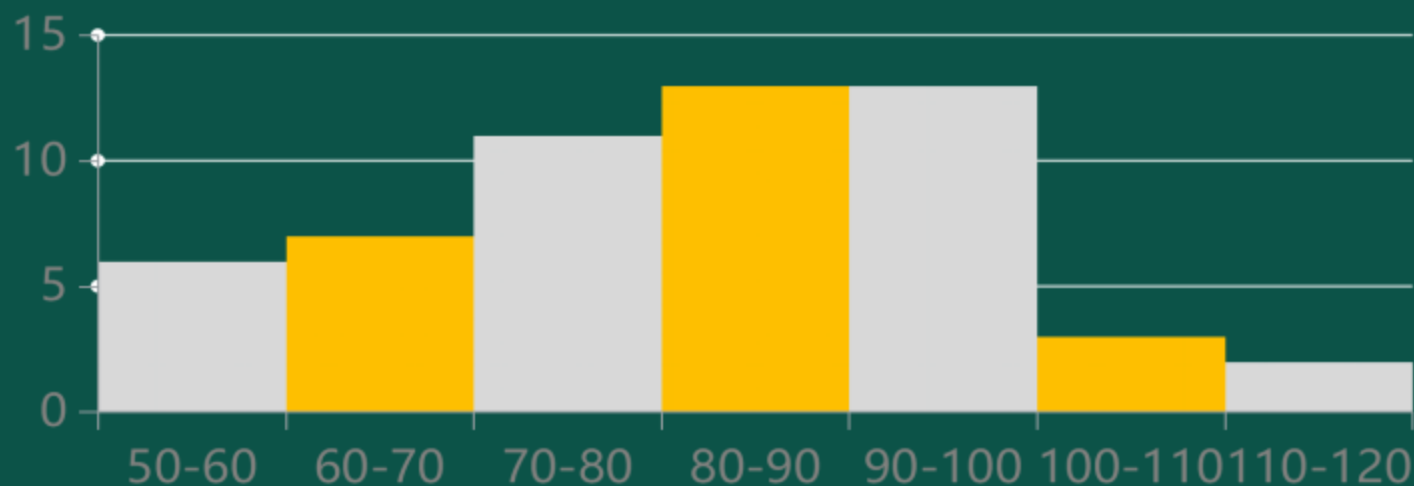


第二节 数值型数据的整理与显示

【知识点二】数值型数据的图示★★

1. 直方图

在直方图中，用矩形的面积表示各组的频数分布。如果取矩形的宽度为一个单位，高度表示比例（即频率），直方图下的总面积等于1。





第二节 数值型数据的整理与显示

【区分】直方图VS条形图

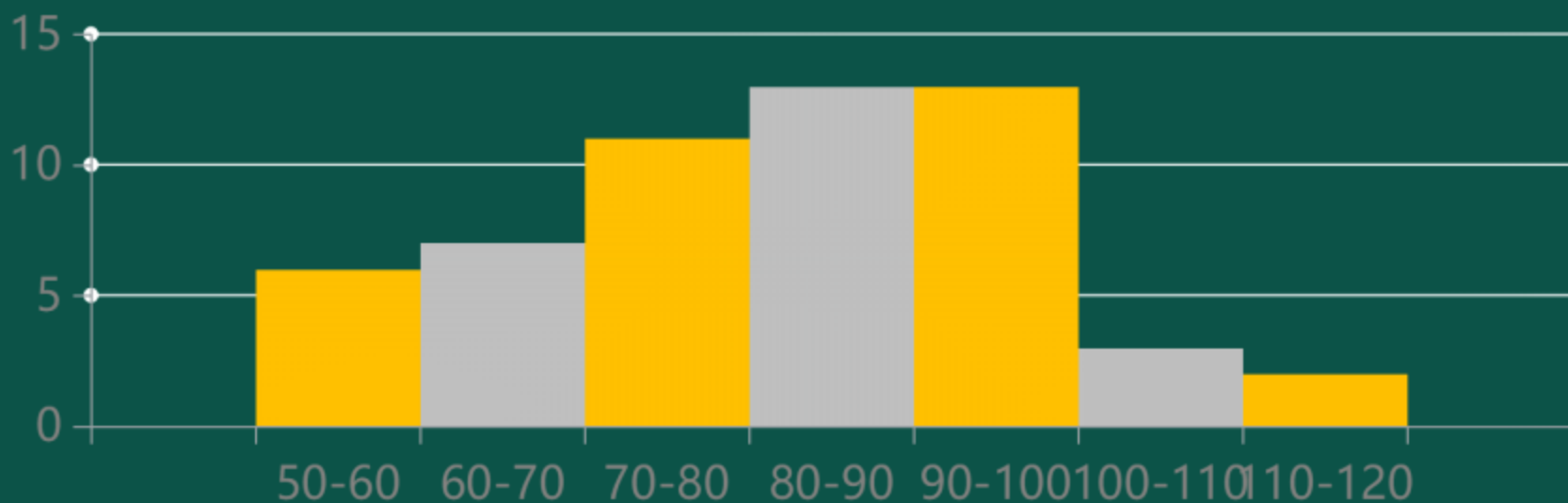
直方图	条形图
<p><u>面积</u>表示<u>各组频数的多少</u></p> <p><u>高度</u>表示<u>每一组的频数或百分比</u></p> <p><u>宽度</u>表示<u>各组的组距</u></p> <p>矩形通常<u>连续排列</u></p>	<p>—</p> <p>长度：表示各类别频数多少</p> <p>宽度：表示类别</p> <p>宽度则是固定的，且矩形<u>分开排列</u></p>



第二节 数值型数据的整理与显示

2. 折线图

折线图也称频数多边形图，它是在直方图的基础上，把直方图顶部的中点（即组中值）用直线连接起来，再把原来的直方图抹掉就是折线图。





真题回顾

【真题·多选】直方图与条形图的区别在于（ ）。

- A. 直方图的各矩形通常是分开排列
- B. 直方图的面积表示各组频数的多少
- C. 直方图的各矩形通常是连续排列
- D. 直方图的矩形高度与宽度均有意义
- E. 直方图的矩形高度有意义而宽度无意义



真题回顾

答案：BCD

解析：直方图的各矩形是连续排列的。在直方图中，用矩形的面积表示各组的频数分布。如果取矩形的宽度为一个单位，高度表示比例（即频率），直方图下的总面积等于1。

第三节 统计表



第三节 统计表

表头	行标题	列标题	数字资料
<ul style="list-style-type: none">• 放在表的上方，说明统计表的主要内容	<ul style="list-style-type: none">• 统计表的第一列	<ul style="list-style-type: none">• 统计表的第一行	<ul style="list-style-type: none">• 表的其余部分是具体的数字资料

【知识点一】统计表的构成★

【补充】表外附加：放在统计表的下方，包括资料来源、指标的注释和必要的说明等。



第三节 统计表

【知识点二】统计表的设计★★

1. 要合理安排统计表的结构。
2. 表头一般应包括表号、总标题和表中数据的单位等内容。
3. 表中的上下两条横线一般用粗线，中间的其他线要用细线，看起来清楚、醒目。

【注意】

1. 通常情况下，统计表的左右两边不封口。
2. 列标题之间一般用竖线隔开，而行标题之间通常不必用横线隔开。



第三节 统计表

3. 表中的数据一般是右对齐，有小数点时应以小数点对齐，而且小数点的位数应统一。

4. 对于没有数字的表格单元，一般用“—”表示，一张填好的统计表不应出现空白单元格。

使用统计表时，必要时可在表的下方加上注释，特别要注意注明资料来源。



真题回顾

- 【真题·多选】关于统计表设计注意事项的说法，正确的有（ ）。
- A. 表头一般应包括表号、总标题和表中数据的单位等内容
 - B. 资料来源不必注明
 - C. 通常情况下，统计表的左右两边不封口
 - D. 表中的数据一般是右对齐，有小数点时应以小数点对齐，而且小数点的位数应统一
 - E. 行标题、列标题、数字资料的位置应安排合理



真题回顾

答案：ACDE

解析：本题考查统计表的设计。在使用统计表时，必要时可在表的下方加上注释，特别要注意注明资料来源，以表示对他人劳动成果和知识产权的尊重，方便读者查阅使用。B的说法有误。



本章内容总结

分类数据	计算出每一类别的频数、频率或比例、比率
	条形图和圆形图
顺序数据	累积频数和累积频率
	累积频数分布图
数值型数据	单变量分组和组距分组
	直方图和折线图
统计表	四大组成：表头、行标题、列标题、数字资料