

第二十章 统计数据的整理与显示

第二十章 统计数据的整理与显示		
考点 1	品质数据的整理和显示	★★★★
考点 2	数值型数据的整理和显示	★★★★
考点 3	统计表	★★★

考点 1 品质数据的整理和显示★★★★

一、分类数据的整理与显示	
1. 频数与频数分布 分类数据整理时常用的指标	
比例	总体中各个部分数量占总体数量的比重，反映总体的构成或结构。各部分的比例之和等于 1
百分比	比例乘以 100% 就是百分比或百分数，用 % 表示。是一个更为标准化的数值，很多相对数都用百分比表示
比率	各不同类别数量的比值，可以是一个总体中各不同部分的数量对比。如经济学中积累与消费之比、国内生产总值中第一、二、三产业产值之比。比率的值可能大于 1

2. 分类数据		
	条形图	圆形图
作用	在表示定类数据的分布时，用条形图的高度来表示各类别数据的频数或频率	主要用于表示总体中各组成部分所占的比例，对于研究结构性问题十分有用

二、顺序数据的整理与显示方法	
1. 累积频数和累积频率	
累积频数	向上累积 类别顺序开始一方——最后一方累加频数 数值型数据从变量值小一方——大的一方累加频数)
	向下累积 类别顺序最后一方——开始一方累加频数 数值型数据从变量值大一方——小的一方累加频数)
累积频率	就是将各类别的百分比逐级累加起来，也有向上累积和向下累积两种方法

二、顺序数据的整理与显示方法						
工资等级	工人数	百分比	向上累积		向下累积	
			工人数 (人)	百分比 (%)	工人数 (人)	百分比 (%)
一级	100	18.9	100	18.9	530	100.0
二级	125	23.6	225	42.5	430	81.1
三级	105	19.8	330	62.3	305	57.5
四级	80	15.1	410	77.4	200	37.7
五级	50	9.4	460	86.8	120	22.6
六级	35	6.6	495	93.4	70	13.2

七级	25	4.7	520	98.1	35	6.6
八级	10	1.9	530	100.0	10	1.9

二、顺序数据的整理与显示方法

2. 顺序数据

【单选题】【2014年】2013年某企业共有500人，这些工人工资水平从低到高分为一到五级。根据该企业工人工资状况整理累积频数分布如下：则2013年该企业工资水平为三级的工人有（ ）人。

- A.80
- B.190
- C.150
- D.400

工资水平	向上累积百分比
一级	12
二级	50
三级	80
四级	90
五级	100

【答案】C

【解析】累积频数（Cumulative Frequencies）。就是将各类别的频数逐级累加起来。其方法有两种：一是从类别顺序的开始一方向类别顺序的最后一方累加频数（数值型数据则是从变量值小的一方向变量值大的一方累加频数），称为向上累积；反之，则是向下累积。该企业工资等级为三级的工人占全体工人的百分比为 $80\% - 50\% = 30\%$ ，则人数为 $500 \times 30\% = 150$ 人。

考点2 数值型数据的整理和显示★★★★

一、数据的分组		
1. 数据分组概述		
方法	单变量值分组	把每一个变量值作为一组。 只适合于离散变量且变量值较少的情况。
	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。 适用于连续变量或变量值较多的情况。

一、数据的分组

2. 组距分组的步骤

第一步，确定分组组数。

第二步，对原始资料进行排序。

第三步，求极差。

将最大的观察值与最小的观察值相减便得到极差。

极差值=最大观察值-最小观察值

【例如】某组数据 50、60、67、75、80、91、105、112 中，极差值=最大观察值-最小观察值=112-50=62

第四步，确定各组组距。

(1) 组距=极差/组数

(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=某组上限值-该组下限值

第五步，确定组限。

【单选题】2019年某公司7名销售经理完成的销售收入分别为：142万元，148万元，155万元，177万元，187万元，211万元，355万元，则这组数据的极差为（ ）万元。

- A.142。

B.177

C.355

D.213

【答案】D

考点 2 数值型数据的整理和显示★★★★

一、数据的分组	
下限	一个组的最小值
上限	一个组的最大值
组距	上限与下限的差值
组中值	上限值与下限值的平均值

一、数据的分组			
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

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。

第六步，确定各组观察值出现的频数。

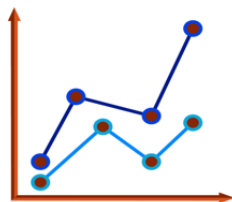
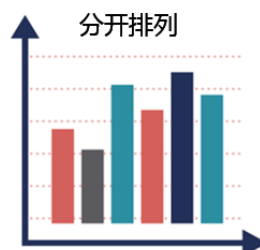
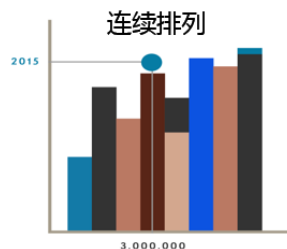
采用组距分组时，需要遵循“不重不漏”的原则。

【例如】“60~70”、“70~80”

70 这个数字应划到“70~80”组里。

第七步，制作频数分布表，并填上相关的内容，以及其他需要说明的事项

二、数值型数据的图示
直方图
条形图
折线图



二、数值型数据的图示		
【总结】数据的整理与显示		
类型	整理方法	显示方法
分类数据	频数、比例、比率、百分比	条形图、圆形图
顺序数据	频数、比例、比率、百分比、累积频数、累积频率	条形图、圆形图、 累积分布图
数值型数据	频数、比例、比率、百分比、累积频数、累积频率、 分组（单变量分组、组距分组）	条形图、圆形图、 累积分布图、直方图、折线图

考点3 统计表★★★★

一、统计表的构成

- (1) 表头：放在表的上方，说明统计表的主要内容
- (2) 行标题：统计表的第一列
- (3) 列标题：统计表的第一行
- (4) 数字资料：表的其余部分是具体的数字资料

二、统计表的设计

设计统计表应注意的问题：

第一，要合理安排统计表的结构。

第二，表头一般应包括表号、总标题和表中数据的单位等内容。

第三，表中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。

- (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。
- (3) 表中数据一般是右对齐，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。

第四，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。