第六章 集中趋势与离散程度的测定

1. 算数平均数 $\overline{x}$

2. 调和平均数 H

3. 几何平均数 G

4. 中位数 M_e

5. 众数 M_o

第一节 平均指标的概念与作用

一、概念:

平均指标：指同类社会经济现象总体内各单位某一数量标志在一定时间、地点和条件下数量差异抽象化的代表性水平指标，其数值表现为平均数，具有单位名称，其计量单位与标志值的计量单位一致

二、作用:

1.可以了解总体次数分布的集中趋势；
2.可以对若干同类现象在不同单位、地区间进行比较研究；
3.可以研究某一总体数值的平均水平在时间上的变化，说明总体的发展过程和趋势；
4.可以分析现象之间的依存关系；
5.可以作为某些科学预测、决策和某些推算的依据.

三、种类

1. 按计算方法

- 数值平均数
- 位置平均数

2. 按时间状况不同

- 静态平均数
- 动态平均数

第二节 算数平均数

一、算数平均数的基本形式

算数平均数=$\frac{\text{总体标志总量}}{\text{总体单位总数}}$

二、算数平均数的计算方法：$\overline{x}$

（一）简单算术平均数

- 计算公式：

简单算术平均数=$\frac{\text{各单位标志值之和}}{\text{总体单位总数}}$（应用条件：资料未分组）

$\overline{x}=\frac{x_1+x_2+\cdots+x_n}{n}=\frac{\sum x}{n}$

（二）加权算术平均数

- 计算公式：

加权算术平均数=$\frac{各组标志值与该组单位数乘积的总和}{各组单位数的总和}$ （应用条件：单项式分组，各组次数不同）

$\overline{x}=\frac{x_1f+x_2f+\cdots+x_nf}{\sum f}=\frac{\sum xf}{\sum f}$

几种特殊平均数

1. 是非平均数： 把具有某种特征的用“1”表示，不具有该种特征的用“0”表示。
2. 评分平均数：通常把人们对评价项目的评分组作为x,把事先规定的项目重要程度作权数f。
3. 等级平均数
4. 先进平均数：比算术平均数大的数求平均。
5. 截尾平均数：将数据分成四组，去掉首尾。

（三）算术平均数的性质

1. 各个变量值与其平均数离差之和等于零
   $\sum(x-\overline{x})=0$
   $\sum(x-\overline{x})f=0$
2. 各个变量值与其平均数离差平方之和为最小值
   $\sum(x-\overline{x})^2=最小值$
   $\sum(x-\overline{x})^2f=最小值$

设$x_0\neq x \qquad x_0=\overline{x}-c$

$$\begin{aligned}
\sum(x-x_0)^2&=\sum[x-(\overline{x}-c)]^2\\
&=\sum[(x-\overline{x})+c]^2\\
&=\sum(x-\overline{x})^2+2c\sum(x-\overline{x})+nc^2\\
&=\sum(x-\overline{x})^2+nc^2\\
&\because nc^2\geq 0\\
&\therefore \sum(x-x_0)^2\geq\sum(x-\overline{x})^2\\
&\therefore \sum(x-\overline{x})^2为最小值
\end{aligned}$$

第三节 调和平均数(H)

一、概念：

变量值倒数得算数平均数的倒数，又称倒数平均数.

特点：

用特定的权数($m=xf$)加权,其变量值多为相对数和平均数。

二、简单调和平均数

- 计算公式：
$H=\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}=\frac{n}{\sum\frac{1}{x}}$
- 应用条件：资料未分组，各个变量值次数都是1

三、加权调和平均数

- 计算公式：
$H=\frac{m_1+m_2+\cdots+m_n}{\frac{m_1}{x_1}+\frac{m_2}{x_2}+\cdots+\frac{m_n}{x_n}}=\frac{\sum m}{\sum\frac{m}{x}}$
- 应用条件：资料分组，各组次数不同

第四节 几何平均数(G)

一、简单几何平均

- 计算方法
$G=\sqrt[n]{x_1x_2\cdots x_n}=\sqrt[n]{\prod x}$
- 应用条件：资料未分组，各变量值次数是1

二、加权几何平均

- 计算方法
$G=\sqrt[\sum f]{x_1^{f_1}x_2^{f_2}\cdots x_n^{f_n}}=\sqrt[\sum f]{\prod x^f}$
- 资料经过分组，各组次数不同

第五节 众数和中位数

一、众数(M_o)

1. 由未分组资料确定众数

出现次数最多的数

2. 由单项数列确定众数

按日产量分组（件）	工人数（人）
20	15
21	30
22	20
23	10

$M_o=21(件)$

3. 由组距数列确定众数

$$\begin{aligned}M_0=&L+\frac{f_0-f_{-1}}{(f_0-f_{-1})+(f_0-f_{+1})}\times i\\&=L+\frac{\Delta_1}{\Delta_1+\Delta_2}\times i\,(下限公式)\end{aligned}$$

$$
\begin{aligned}
M_0 =&U-\frac{f_0-f_{+1}}{(f_0-f_{-1})+(f_0-f_{+1})}\times i \\
&=U-\frac{\Delta_2}{\Delta_1+\Delta_2}\times i \quad（限 公式）
\end{aligned}
$$

二、中位数(M_e)

1. 由未分组资料确定中位数：先排序$X_1, X_2, ..., X_n$

$$
\boldsymbol{M}_e=\begin{cases}\frac{X_{\frac{n}{2}}+X_{\left(\frac{n}{2}+1\right)}}{2}, & n\text{为偶数}\\ X_{\frac{n+1}{2}}, & n\text{为奇数}\end{cases}
$$

2. 由单项数列确定中位数：中点位置$\frac{\sum f}{2}$，该位置所在组就是中位数。

3. 由组距数列确定中位数

1. 确定中位数所在组$\frac{\sum f}{2}$
2. 计算公式

$$
\boldsymbol{M}_e=\boldsymbol{L}+\frac{\frac{\sum f}{2}-\boldsymbol{s}_{m-1}}{\boldsymbol{f}_m}\times \boldsymbol{i}(\text{ 下限公式 })
$$

$$
\boldsymbol{M}_e=\boldsymbol{U}-\frac{\frac{\sum f}{2}-\boldsymbol{s}_{m+1}}{\boldsymbol{f}_m}\times \boldsymbol{i}(\text{ 上限公式 })
$$

三、四分位数(Q)

$$
\text{未分组数据：}\begin{cases}\text{下四分位数}(Q_L)\text{位置}=\frac{N+1}{4}\\ \text{上四分位数}(Q_U)\text{位置}=\frac{3(N+1)}{4}\end{cases}
$$

$$
\text{组距分组数据：}\begin{cases}\text{下四分位数}(Q_L)\text{位置}=\frac{N}{4}\\ \text{上四分位数}(Q_U)\text{位置}=\frac{3N}{4}\end{cases}
$$

数值型分组数据的四分位数(计算公式)

- 下四分位数：$Q_L=L_L+\frac{\frac{N}{4}-S_L}{f_L}\times i_L$
- 上四分位数：$Q_U\doteq L_U+\frac{\frac{3N}{4}-S_U}{f_U}\times i_U$

第六节标志变异指标

一、作用

1. 可以衡量平均数代表性的大小；
2. 可以反映社会经济活动过程节奏性和均衡性；
3. 可以反映总体单位标志值的均匀性和稳定性；
4. 是科学地确定必要抽样单位数应考虑的重要因素.

二、绝对数形式

1. 全距(R)

- 公式：$R =$ 最大值 $-$ 最小值
  或 最高组的上限 $-$ 最低组的下限

2. 四分位差

- 公式：$Q_D = \frac{(Q_U - Q_L)}{2}$

3.平均差(A.D)

- 公式：$\mathbf{A.D} = \frac{\sum|\mathbf{x}-\overline{\mathbf{x}}|}{\mathbf{n}}$

4. 加权平均差

- 公式：$\mathbf{A.D} = \frac{\sum|\mathbf{x}-\overline{\mathbf{x}}|\mathbf{f}}{\sum\mathbf{f}}$

5. 标准差 σ

- 公式：

$$\sigma = \sqrt{\frac{\sum(\mathrm{x}-\overline{\mathrm{x}})^2}{\mathrm{n}}} = \sqrt{\frac{\sum\mathrm{x}^2}{\mathrm{n}} - \left(\frac{\sum\mathrm{x}}{\mathrm{n}}\right)^2}$$

6. 加权标准差

- 公式：

$$\sigma = \sqrt{\frac{\sum(\mathrm{x}-\overline{\mathrm{x}})^2\mathrm{f}}{\sum\mathrm{f}}} = \sqrt{\frac{\sum\mathrm{x}^2\mathrm{f}}{\sum\mathrm{f}} - \left(\frac{\sum\mathrm{xf}}{\sum\mathrm{f}}\right)^2}$$

7.是非标志的标准差

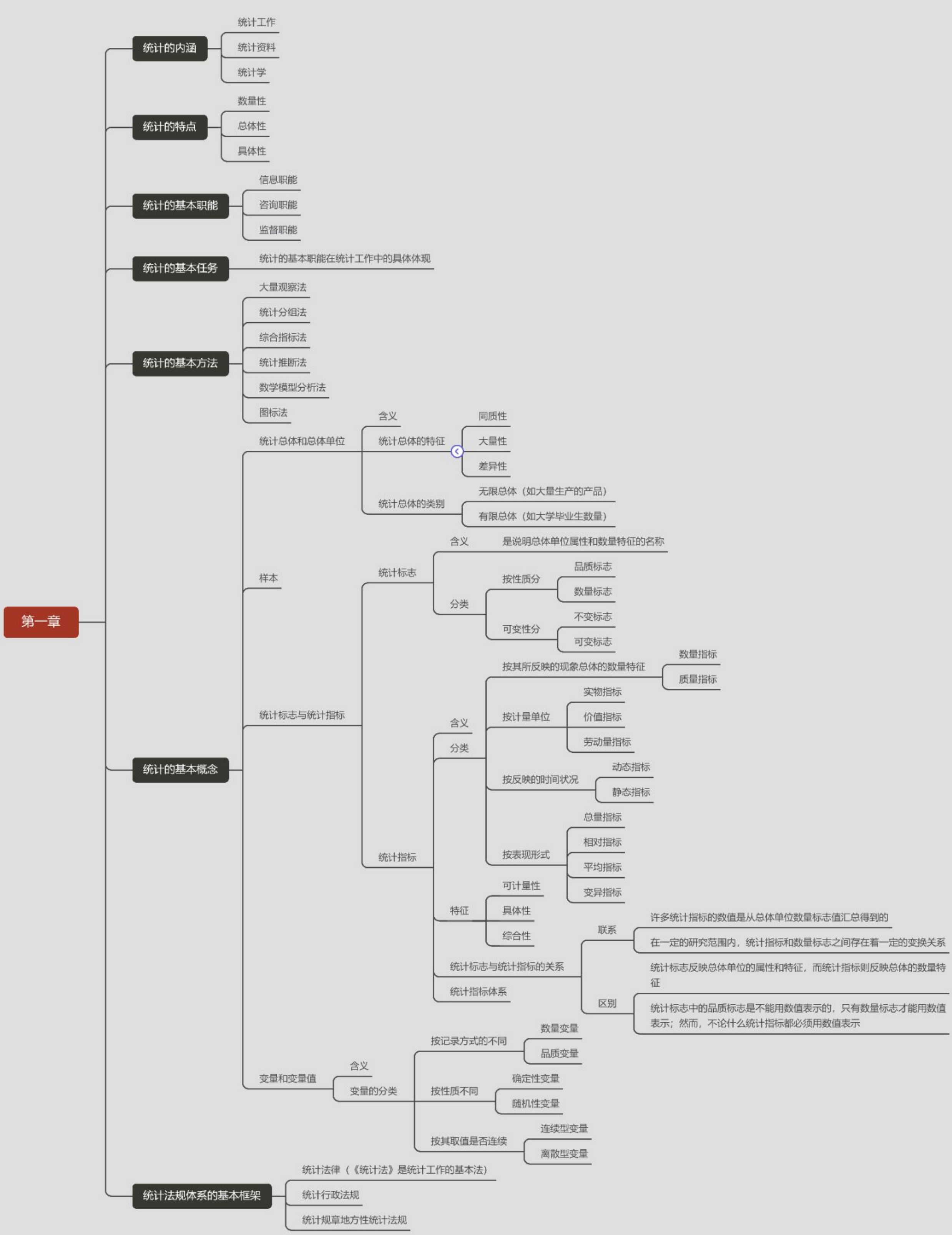
- 公式：

$$\begin{aligned}\sigma_{是} &= \sqrt{\frac{(1-\mathrm{P})^2\,\mathrm{N}_1 + \mathrm{P}^2\,\mathrm{N}_0}{\mathrm{N}}} \\ &= \sqrt{\mathrm{P}(1-\mathrm{P})}\end{aligned}$$

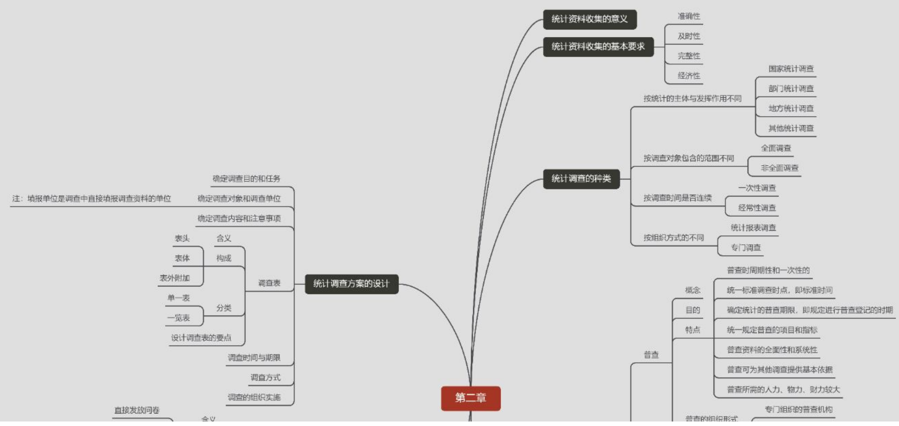
标志值x	单位数f	$(x-\overline{x})^2f$
1	N_1	$(1-P)^2N_1$
0	N_0	$(0-p)^2N_0$
合计	$\boldsymbol{N}$	$(1-P)^2N_1 + P^2N_0$

8. 标准差系数

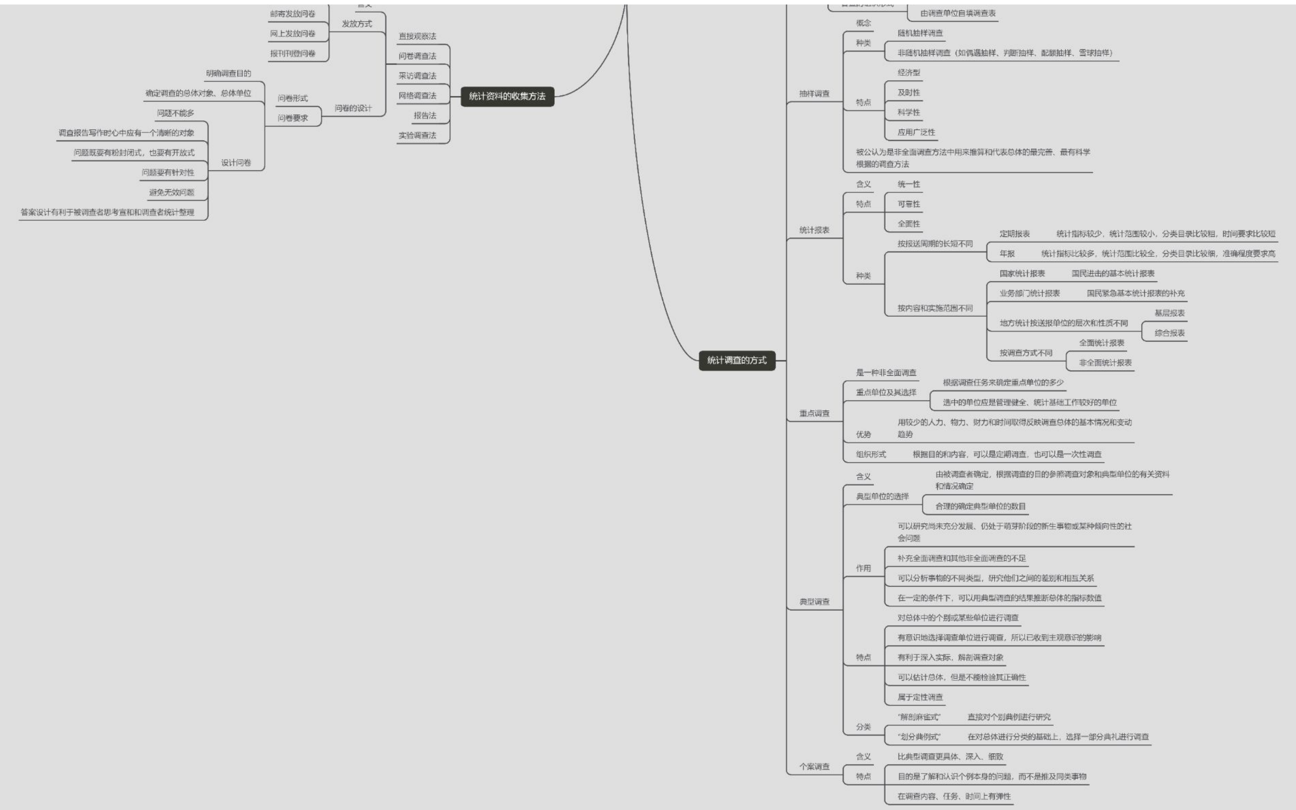
- 公式:$V_\sigma = \frac{\sigma}{x}$

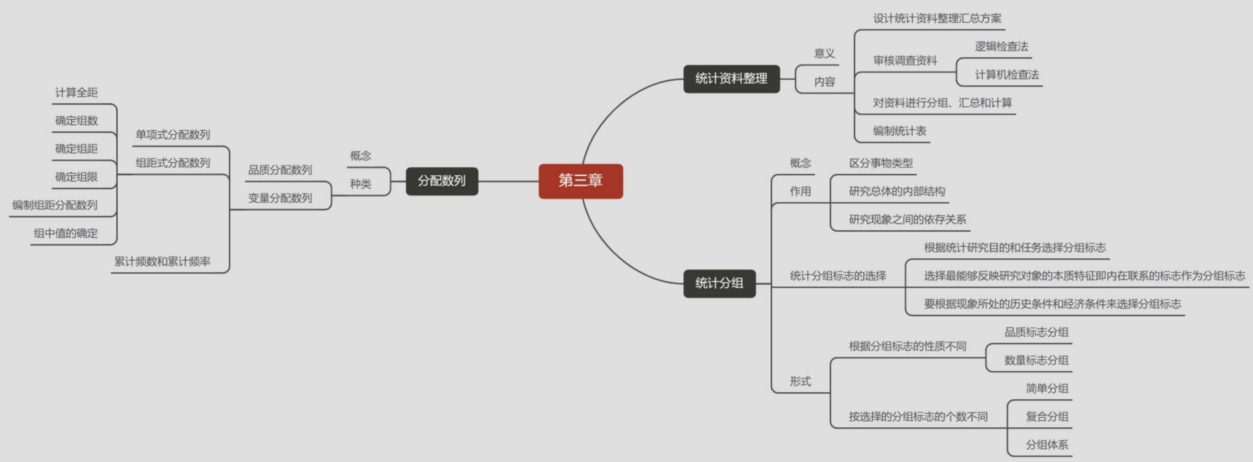


第一章
统计的内涵
统计工作
统计资料
统计学
统计的特点
数量性
总体性
具体性
统计的基本职能
信息职能
咨询职能
监督职能
统计的基本任务
统计的基本职能在统计工作中的具体体现
统计的基本方法
大量观察法
统计分组法
综合指标法
统计推断法
数学模型分析法
图标法
统计的基本概念
统计总体和总体单位
含义
统计总体的特征
同质性
大量性
差异性
统计总体的类别
无限总体（如大量生产的产品）
有限总体（如大学毕业生数量）
样本
统计标志与统计指标
统计标志
含义
是说明总体单位属性和数量特征的名称
分类
按性质分
品质标志
数量标志
可变性分
不变标志
可变标志
统计指标
含义
分类
按其所反映的现象总体的数量特征
数量指标
质量指标
按计量单位
实物指标
价值指标
劳动量指标
按反映的时间状况
动态指标
静态指标
按表现形式
总量指标
相对指标
平均指标
变异指标
特征
可计量性
具体性
综合性
统计标志与统计指标的关系
联系
许多统计指标的数值是从总体单位数量标志值汇总得到的
在一定的研究范围内，统计指标和数量标志之间存在着一定的变换关系
区别
统计标志反映总体单位的属性和特征，而统计指标则反映总体的数量特征
统计标志中的品质标志是不能用数值表示的，只有数量标志才能用数值表示；然而，不论什么统计指标都必须用数值表示
统计指标体系
变量和变量值
含义
变量的分类
按记录方式的不同
数量变量
品质变量
按性质不同
确定性变量
随机性变量
按其取值是否连续
连续型变量
离散型变量
统计法规体系的基本框架
统计法律（《统计法》是统计工作的基本法）
统计行政法规
统计规章地方性统计法规



第二章
统计资料收集的意义
统计资料收集的基本要求
准确性
及时性
完整性
经济性
统计调查的种类
按统计的主体与发挥作用不同
国家统计调查
部门统计调查
地方统计调查
其他统计调查
按调查对象包含的范围不同
全面调查
非全面调查
按调查时间是否连续
一次性调查
经常性调查
按组织方式的不同
统计报表调查
专门调查
普查
概念
目的
特点
普查时周期性和一次性的
统一标准调查时点，即标准时间
确定统计的普查期限，即规定进行普查登记的时期
统一规定普查的项目和指标
普查资料的全面性和系统性
普查可为其他调查提供基本依据
普查所需的人力、物力、财力较大
专门组织的普查机构
统计调查方案的设计
确定调查目的和任务
确定调查对象和调查单位
注：填报单位是调查中直接填报调查资料的单位
确定调查内容和注意事项
调查表
含义
构成
表头
表体
表外附加
分类
单一表
一览表
设计调查表的要点
调查时间与期限
调查方式
调查的组织实施
直接发放问卷


统计资料的收集方法
直接观察法
问卷调查法
采访调查法
网络调查法
报告法
实验调查法
发放方式
邮寄发放问卷
网上发放问卷
报刊刊登问卷
问卷的设计
问卷形式
问卷要求
明确调查目的
确定调查的总体对象、总体单位
设计问卷
问题不能多
调查报告写作时心中应有一个清晰的对象
问题既要有粉封闭式，也要有开放式
问题要有针对性
避免无效问题
答案设计有利于被调查者思考宣和和调查者统计整理
统计调查的方式
由调查单位自填调查表
抽样调查
概念
种类
随机抽样调查
非随机抽样调查（如偶遇抽样、判断抽样、配额抽样、雪球抽样）
特点
经济型
及时性
科学性
应用广泛性
被公认为是非全面调查方法中用来推算和代表总体的最完善、最有科学根据的调查方法
统计报表
含义
特点
统一性
可靠性
全面性
种类
按报送周期的长短不同
定期报表
统计指标较少，统计范围较小，分类目录比较粗，时间要求比较短
年报
统计指标比较多，统计范围比较全，分类目录比较细，准确程度要求高
按内容和实施范围不同
国家统计报表
国民进击的基本统计报表
业务部门统计报表
国民紧急基本统计报表的补充
地方统计按送报单位的层次和性质不同
基层报表
综合报表
按调查方式不同
全面统计报表
非全面统计报表
重点调查
是一种非全面调查
重点单位及其选择
根据调查任务来确定重点单位的多少
选中的单位应是管理健全、统计基础工作较好的单位
优势
用较少的人力、物力、财力和时间取得反映调查总体的基本情况和变动趋势
组织形式
根据目的和内容，可以是定期调查，也可以是一次性调查
典型调查
含义
典型单位的选择
由被调查者确定，根据调查的目的参照调查对象和典型单位的有关资料和情况确定
合理的确定典型单位的数目
作用
可以研究尚未充分发展、仍处于萌芽阶段的新生事物或某种倾向性的社会问题
补充全面调查和其他非全面调查的不足
可以分析事物的不同类型，研究他们之间的差别和相互关系
在一定的条件下，可以用典型调查的结果推断总体的指标数值
特点
对总体中的个别或某些单位进行调查
有意识地选择调查单位进行调查，所以已收到主观意识的影响
有利于深入实际，解剖调查对象
可以估计总体，但是不能检验其正确性
属于定性调查
分类
"解剖麻雀式"
直接对个别典例进行研究
"划分典例式"
在对总体进行分类的基础上，选择一部分典礼进行调查
个案调查
含义
比典型调查更具体、深入、细致
特点
目的是了解和认识个例本身的问题，而不是推及同类事物
在调查内容、任务、时间上有弹性


第三章
统计资料整理
意义
内容
设计统计资料整理汇总方案
审核调查资料
逻辑检查法
计算机检查法
对资料进行分组、汇总和计算
编制统计表
分配数列
概念
种类
品质分配数列
变量分配数列
单项式分配数列
组距式分配数列
计算全距
确定组数
确定组距
确定组限
编制组距分配数列
组中值的确定
累计频数和累计频率
统计分组
概念
区分事物类型
作用
研究总体的内部结构
研究现象之间的依存关系
统计分组标志的选择
根据统计研究目的和任务选择分组标志
选择最能够反映研究对象的本质特征即内在联系的标志作为分组标志
要根据现象所处的历史条件和经济条件来选择分组标志
形式
根据分组标志的性质不同
品质标志分组
数量标志分组
按选择的分组标志的个数不同
简单分组
复合分组
分组体系


第四章
统计表
含义
条理清晰地反映大量统计数字资料
作用
简洁醒目体现大量数据中各项目间的关系，便于检查资料中数字地完整性、正确性，便于进行各种计算
科学合理地组织统计资料，便于阅读，对照比较和分析
构成
形式构成
总标题
横行标题
纵栏标题
指标数值
表外附加（必要时）
内容构成
主语
宾词（指标名称和指标数值）
种类
按用途分类不同
统计调查表
统计汇总表
统计分析表
按照主词分组情况不同
简单表
分组表
符合表
注：不宜超过三个分组标志分组
编制统计表的规则
要遵循科学、实用、简洁、醒目、美观
具体规则
各种标题要简明确切地概括所有表明的内容
表中各行各列的排列应注意反映资料的逻辑关系
表中数值资料必须标明计量单位
统计表中的数字要清晰，对准位数
统计表的表示一般是开口式
表内引用的资料要在表下注明其资料来源
统计图
含义
现象间的对比关系
用途
揭露总体结构
检查计划的执行情况
揭示现象间的依存关系，反映总体单位的分配情况
说明现象在空间上的分布情况
种类
条形图
含义
特点
一眼看出各个数据的大小
易于比较数据之间的差别
直方图
制作统计图
折线图
能够显示数据的变化趋势，反映事物的变化情况
圆形（饼图）
茎叶图