

第十九章 统计调查

统计调查方式	普查	为某一特定目的而专门组织的一次性全面调查，如人口普查、农业普查和经济普查				
	抽样调查 (最广泛)	定义	从总体中抽取单位作为样本推断总体数量特征的非全面调查			
		特点	经济性、时效性强、适应面广、准确性高			
		方法	<table border="1" style="width: 100%;"> <tr> <td>概率抽样</td> <td>如简单随机抽样、分层抽样、整群抽样、等距(系统)等抽样形式</td> </tr> <tr> <td>非概率抽样</td> <td>如偶遇抽样、判断抽样、配额抽样、滚雪球抽样等。</td> </tr> </table>	概率抽样	如简单随机抽样、分层抽样、整群抽样、等距(系统)等抽样形式	非概率抽样
	概率抽样	如简单随机抽样、分层抽样、整群抽样、等距(系统)等抽样形式				
非概率抽样	如偶遇抽样、判断抽样、配额抽样、滚雪球抽样等。					
重点调查	非全面调查，在总体中选择重点单位调查					
典型调查	非全面调查，选择若干具有典型意义的或有代表性的单位进行的调查					

统计数据	第一手统计数据	①农产量用 直接观察法 ；②现行的统计制度是 报告法 ；③生死用 登记方法 ；④西方流行 电话访问法 ；⑤新产品、新工艺用 试验设计法
	第二手统计数据	方法(主要来源)：公开的出版物、未公开的内部调查等
		①评估数据的可用价值②指标口径可比性③弥补缺失数据和进行质量检查④引用统计数据要注明数据来源
大数据	特征：容量大、类型多、存取速度快、应用价值高	
	内容：行政记录数据(个人、单位、自然资源和其他管理记录数据)、商业记录数据、互联网数据、点设备感应数据以及其他传统数据	

误差及其来源	登记性	如填报错误、抄录错误、汇总错误等。有意虚报或瞒报调查数据，从理论上讲可以消除
	代表性	如没有遵循随机原则；样本结构与总体结构存在差异；样本容量不足等，通常无法消除，但事先可以进行控制或计算
质量评价标准	①精度(抽样误差最低)；②准确性(非抽样误差最小)；③关联性(满足需要)；④及时性(时间最短)；⑤一致性(可比性)；⑥最低成本	

第二十章 统计数据的整理与显示

品质数据	分类数据	整理	频数和频数分组、比例(部分对部分)、比率(部分对总体)、百分比
		图示	条形图，纵置称为条形图；横置称为柱形图
			圆形图也称饼图，用于表示总体中各组成部分所占的比例
	顺序数据	整理	累积频数/累积频率，分为向上累积、向下累积
图示		累积频数或频率分布图	

数值型数据数据的整理	数据分组	单变量值	每一个变量值作为一组，适合于离散变量且变量值较少的情况。
		组距分组步骤	①确定分组组数；②排序；③极差(最大的观察值-最小的观察值)；④定组距；⑤定组限；⑥定频数；⑦作频数分布表。【注意】不重不漏、上组限不在内
	图示	直方图	矩形宽度表示组距；高度表示每组的组数或百分比；(条形图只是高度或长度有意义，宽度无意义)
		折线图	以直方图为基础，连接直方图顶部的中点(即组中值)，再把直方图抹掉就是折线图

第二十一章 数据特征的测度

集中趋势	位置平均数	众数：一组数据中出现频数最多的那个数值。
		中位数：从小到大的顺序进行排列，位置居中的数值
	数值平均数	算术平均数：简单平均数和加权平均数

	(最主要)	几何平均数: n 个观察值连乘积的 n 次方根
离散程度	极差:	总体或分布中最大的标志值与最小的标志值之差
	方差:	总体所有单位标志值与其平均数离差之平方, 平方根是标准差
	离散系数:	标准除均, 为消除变量值水平高低和计量单位不同