



第四部分 统计



第二十三章

统计与数据科学



第一节 统计学

(一) 定义

统计学是关于收集、整理、分析数据和从数据中得出结论的科学。

统计学有两个分支：描述统计和推断统计。



第一节 统计学

(二) 分类

(1) 描述统计

描述统计是研究数据收集、整理和描述的统计学方法。其内容包括：

- ①如何取得所需要的数据；
- ②如何用图表或数学方法对数据进行整理和展示；
- ③如何描述数据的一般性特征。



第一节 统计学

(2) 推断统计

推断统计是研究如何利用样本数据来推断总体特征的统计学方法，其内容包括参数估计和假设检验两大类。

参数估计：利用样本信息推断总体特征。

假设检验：利用样本信息判断对总体的假设是否成立。



第一节 统计学

【例-多选题】下列统计分析中，采用的是推断统计方法的有（ ）。

- A. 利用样本信息估计总体特征
- B. 利用图表对数据进行展示
- C. 描述一组数据的集中趋势
- D. 利用样本信息检验对总体的假设是否成立
- E. 描述一组数据的离散趋势

答案：AD

解析：推断统计是研究如何利用样本数据来推断总体特征的统计学方法，其内容包括参数估计和假设检验两大类。



第二节 变量和数据

(一) 变量

(1) 变量是研究对象的属性或特征，可以有两个或更多个可能的取值；而常数只有一个固定取值。



第二节 变量和数据

(2) 分类

类别		含义	举例
定量变量或数量变量		变量的取值是数量	企业销售额、注册员工数量等
定性变量	分类变量	变量的取值表现为类别	企业所属行业
	顺序变量	变量的取值表现为类别且具有一定顺序	员工受教育水平



第二节 变量和数据

(二) 数据

(1) 数据是对变量进行测量、观测的结果。数据根据需要可以是数值、文字或者图像等形式。



第二节 变量和数据

(2) 分类

类别	含义	举例
分类数据	分类变量的观测结果， 表现为类别 ，一般用文字来表述，也可用数值代码表示。（对分类数据可以计算出各类别的频率，但对其进行加、减、乘或除等数学运算是没有意义的。）	比如用1 表示“男性”，用2 表示“女性”。
顺序数据	对顺序变量的观测结果， 也表现为类别 ，一般用文字表述，也可用数值代码表示。	比如用1 表示“硕士及以上”，用2表示“本科”，用3表示“大专及以下”。
数值型数据	对定量变量的观测结果， 其取值表现为具体的表示大小或多少数值 。对不同类型的数据，可采用不同的统计方法来处理和分析。（数值型数据可以进行数学运算，比如计算均值和方差等统计量。	



第二节 变量和数据

【例-多选题】下列变量中，通常属于数值型变量的有（ ）。

- A. 商品销售额
- B. 上班出行方式
- C. 家庭收入
- D. 居住地区
- E. 年龄

答案：ACE

解析：数值型数据是对定量变量的观测结果，其取值表现为具体的表示大小或多少数值。很明显除了ACE项，BD项均无法用数值体现，故错误。



第二节 变量和数据

【例-多选题】下列变量中，属于顺序变量的有（ ）。

- A. 行业类别
- B. 可支配收入
- C. 受教育水平
- D. 球队排名
- E. 性别

答案：CD

解析：顺序变量的取值表现为类别且具有一定顺序。AB项均属于分类变量。