统计学是收集 IMG_256 、分析、表述和解释数据

的科学。

研究 数据收集 -> 整理 -> 描述 的统计学分支

只能归于某一类的非数字型数据

只能归于某一有序类别的非数字型数据

按照数值尺度测量的观察值（样本直接给你了具体的数字）

通过 调查/观测 得到的数据

在实验中 控制实验对象 得到的数据

用来描述**总体特征**的概括性数字度量

用来描述**样本特征**的概括性数字度量

比如样本平均数，样本标准差s，样本概率p

 = 1 + 2 + 3 + ... + n





 = 

s = 

 = (  +  + ... + ) ÷ n =  = 

P(A) = 

从总体中**随机**抽取一部分单位作为样本调查，

并根据调查结果推断总体特征

的数据收集方法

根据**已知的概率**来抽取样本单位

比如我们都知道硬币正反面概率都是50%，

让一个机械手去抛硬币，出现任何结果都是随机的

**有意识地**选取样本单位，样本抽取单位**不随机**

含有N个元素的总体中，

抽取n个元素作为样本，

使得每一个容量为n的样本都有**相同的机会（概率）**被抽中

从N个元素的总体里，

抽1个样本，

放回后，

再抽，总共抽n次

从N个元素的总体里，

抽1个样本，

**不放回，从剩下的总体里**

再抽，总共抽n次

在**没有调查员**协助的情况下

被调查者自己填写，完成调查问卷

现场调查中，

调查员与被调查者**面对面**

调查员提问，被调查者回答这种方式

调查人员通过**打电话**的方式

向被调查者实施调查

对照组和实验组的产生不仅应该**是随机的**，

而且应该**背景材料情况类似**

由抽样的**随机性**引起的样本结果与总体真值之间的误差

除了抽样误差之外，

由**其他原因**引起的

样本观察结果与总体真值之间的差异

检查调查单位或个体是否有遗漏

调查项目是否填写齐全

数据是否有错误

纠正

/

/ 是

是否存在异常值，异常值是否记录错误

\ 否

\

保留

x0.75





1.应当精心设计，有助于**洞察问题的实质**

2.使复杂观点得到**简明、确切、高效**的阐述

3.应当在**最短的时间**内以最少的笔墨给读者提供最大量的信息

5.应当表述数据的**真实**情况

**第**

4.应当**多维**

5

****



 = 

**o o o o o**

**↑**

**2 3 4 5 6 7**

**↑ ↑**

**（4+5）/ 2 = 4.5**

.使复杂

合理安排统计表的结构

****

第个数 和第+1个数的中间值

**1，2，3，4，5，6，7，8，9，10**

**（1+2+3+4+5+6+7+8+9+10）÷ 10**

****

****

|  |  |  |
| --- | --- | --- |
| **组中值 =（上限+下限）÷2** | **频数** |  |
| **=（1+5）÷2= 3** | **5** | **15** |
| **=（6+10）÷2=8** | **5** | **40** |

****

****

****

****

**o o o o o o o o**

**↑ ↑ ↑**

的中间值

 = (15 + 40) ÷ 10 = 5.5



μ



G = =

 = **- **

 = 1 - 



**异众比率越大**，非众数组的频率占总频数的比重越大，**众数的代表性越差**

R













 = 

σ

至少有1-的数据落在±k个标准差的范围内（k>1）

超过3σ范围之外的点，为**离群点**

= 

SK

SK=

偏态峰值

SK=