

## 第六章 样本及抽样分布

### § 6.1 随机样本

总体，个体，样本，样本容量，经验分布，  
样本矩

## 一、总体与总体分布

一个统计问题总有它明确的研究对象.

研究对象的全体称为**总体(母体)**,

总体中每个成员称为**个体**.



研究某批灯泡的质量



考察国产 轿车的质量

然而在统计研究中，人们往往关心每个个体的一项(或几项)数量指标和该数量指标的分布情况。这时，就把数量指标叫作**总体**。



灯泡的寿命




灯泡的寿命就是  
总体



国产轿车每公里的  
耗油量

国产轿车的耗油量就是  
总体



因此，总体就是我们要研究的随机变量（或其分布）。为此常用随机变量的符号或分布的符号来表示总体。

通常，我们用随机变量  $X, Y$  等表示总体。当我们说到总体时，就是指一个具有固定但未知概率分布的随机变量。而总体的分布是指该随机变量的分布。

10月24日

设总体 $X$ 代表从一群人中任取一人的身高，其分布函数为 $F$ 。

为了研究 $X$ 的分布，我们的办法就是取  $n$  个人并测量他们的身高。

记 $X_1, X_2, \dots, X_n$ 分别为这些人的身高，则

- (1)  $X_1, X_2, \dots, X_n$  都有相同的分布函数  $F$ ;
- (2)  $X_1, X_2, \dots, X_n$  相互独立。

我们称  $X_1, X_2, \dots, X_n$  为样本。把这两个特点归纳为一般情形就得到如下定义：

定义. 设总体 $X$ 的分布函数为 $F$ 。若

(1)  $X_1, X_2, \dots, X_n$  都有相同的分布函数 $F$ ;

(2)  $X_1, X_2, \dots, X_n$  相互独立。

则称 $X_1, X_2, \dots, X_n$  为从总体 $X$ (或总体 $F$ 、或分布函数 $F$ )得到的容量为 $n$ 的简单随机样本, 简称样本。

它们的观察值 $x_1, x_2, \dots, x_n$ 称为样本值, 又称为 $X$ 的 $n$ 个独立的观察值.

也可以将样本看成是一个随机向量,  
写成  $(X_1, X_2, \dots, X_n)$ , 此时样本值相  
应地写成  $(x_1, x_2, \dots, x_n)$ .

$(X_1, X_2, \dots, X_n)$  的分布函数为

$$F^*(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F(x_i)$$

又若  $X$  具有概率密度  $f$ , 则  $(X_1, X_2, \dots, X_n)$  的概率密度为

$$f^*(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i)$$

设离散型随机变量（总体） $X$  的分布律

$$P\{X = x\} = p(x)$$

则 $(X_1, \dots, X_n)$ 的联合分布律为：

$$P\{X_1 = x_1, \dots, X_n = x_n\} = \prod_{i=1}^n p(x_i)$$

其中， $(x_1, x_2, \dots, x_n)$  代表  $(X_1, X_2, \dots, X_n)$  的可能值。



## 注意：

- 1、在统计中，上述的  $x_1, x_2, \dots, x_n$  代表的是样本观测值。因此，它们可能重复，也可能只包括一部分的可能值。但是，如果是写密度函数或者概率分布，则与第二、三章一致，是代表自变量、可能值。
- 2、在统计中，总体可能指我们研究的随机变量；或者一批车、灯泡等等；还有可能指该随机变量的分布。
- 3、当我们考虑统计方法时，把样本作为随机变量（大写）。当我们具体计算数值时则用样本值（小写）。

1、当我们考虑样本时， $x_1, x_2, \dots, x_n$ 代表样本值， $X_1, X_2, \dots, X_n$ 代表样本。

2、当我们考虑样本的联合密度（连续型总体）或者样本的分布函数时， $x_1, x_2, \dots, x_n$ 代表函数的自变量。

3、当我们考虑样本的联合分布（离散型总体）时， $(x_1, x_2, \dots, x_n)$ 代表 $(X_1, X_2, \dots, X_n)$ 的可能值。