第一节　随机样本

一、总体与个体

二、随机样本的定义

三、小结

一、总体与个体

1. 总体

试验的全部可能的观察值称为总体.

2. 个体

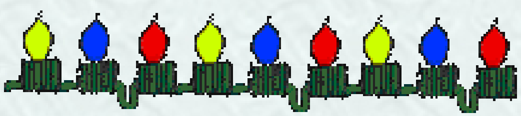
总体中的每个可能观察值称为个体.

实例1 在研究2000名学生的年龄时, 这些学生的年龄的全体就构成一个总体, 每个学生的年龄就是个体.

3. 有限总体和无限总体

实例2 某工厂10月份生产的灯泡寿命所组成的总体中, 个体的总数就是10月份生产的灯泡数, 这是一个有限总体; 而该工厂生产的所有灯泡寿命所组成的总体是一个无限总体, 它包括以往生产和今后生产的灯泡寿命.

当有限总体包含的个体的总数很大时, 可近似地将它看成是无限总体.



4. 总体分布

实例3 在2000名大学一年级学生的年龄中, 年龄指标值为“15”，“16”，“17”，“18”，“19”，“20”的依次有9，21，132，1207，588，43 名, 它们在总体中所占比率依次为

$$\frac{9}{2000},\quad \frac{21}{2000},\quad \frac{132}{2000},\quad \frac{1207}{2000},\quad \frac{588}{2000},\quad \frac{43}{2000},$$

即学生年龄的取值有一定的分布.

一般地, 我们所研究的总体, 即研究对象的某项数量指标 X, 其取值在客观上有一定的分布, X 是一个随机变量.

总体分布的定义

我们把数量指标取不同数值的比率叫做总体分布.

如**实例**3中, 总体就是数集 $\{15, 16, 17, 18, 19, 20\}$.

总体分布为

年龄	15	16	17	18	19	20
比率	$\frac{9}{2000}$	$\frac{21}{2000}$	$\frac{132}{2000}$	$\frac{1207}{2000}$	$\frac{588}{2000}$	$\frac{43}{2000}$

二、随机样本的定义

1. 样本的定义

设 X 是具有分布函数 F 的随机变量，若 X_1, $X_2,\cdots,X_n$ 是具有同一分布函数 F、相互独立的随机变量，则称 $X_1,X_2,\cdots,X_n$ 为从分布函数 F (或总体 F、或总体 X) 得到的容量为 n 的简单随机样本，简称样本.

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

2. 简单随机抽样的定义

获得简单随机样本的抽样方法称为简单随机抽样.

根据定义得：若 $X_1, X_2, \cdots, X_n$ 为 F 的一个样本，

则 $X_1, X_2, \cdots, X_n$ 的联合分布函数为

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

又若 X 具有概率密度 f,

则 $X_1, X_2, \cdots, X_n$ 的联合概率密度为

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

例4 设总体 X 服从参数为 $\lambda\,(\lambda > 0)$ 的指数分布, $(X_1, X_2, \cdots, X_n)$ 是来自总体的样本, 求样本 $(X_1, X_2, \cdots, X_n)$ 的概率密度.

解 总体 X 的概率密度为 $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0, \\ 0, & x \le 0, \end{cases}$

因为 $X_1, X_2, \cdots, X_n$ 相互独立, 且与 X 有相同的分布,

所以 $(X_1, X_2, \cdots, X_n)$ 的概率密度为

$$f_n(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} f(x_i) = \begin{cases} \lambda^n e^{-\lambda \sum\limits_{i=1}^{n} x_i}, & x_i > 0, \\ 0, & \text{其他}. \end{cases}$$

例5 设总体 X 服从两点分布 $B(1,p)$, 其中$0<p<1$, $(X_1,X_2,\cdots,X_n)$是来自总体的样本, 求样本 $(X_1,X_2,\cdots,X_n)$ 的分布律.

解 总体 X 的分布律为

$$P\{X=i\}=p^i(1-p)^{1-i} \qquad (i=0,1)$$

因为 $X_1,X_2,\cdots,X_n$相互独立,

且与 X 有相同的分布,

所以 $(X_1,X_2,\cdots,X_n)$ 的分布律为

$$P\{X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n\}$$

$$= P\{X_1 = x_1\}P\{X_2 = x_2\}\cdots P\{X_n = x_n\}$$

$$= p^{\sum_{i=1}^{n} x_i}(1-p)^{n-\sum_{i=1}^{n} x_i}$$

其中 $x_1, x_2, \cdots, x_n$ 在集合 $\{0,1\}$ 中取值.

三、小结

基本概念：个体　总体 $\begin{cases} \text{有限总体} \\ \text{无限总体} \end{cases}$　随机样本

说明1　一个总体对应一个随机变量 X, 我们将不区分总体和相应的随机变量, 统称为总体 X.

说明2　在实际中遇到的总体往往是有限总体, 它对应一个离散型随机变量; 当总体中包含的个体的个数很大时, 在理论上可认为它是一个无限总体.