§6.1 总体和参数

- 总体
- 个体
- 样本（简单随机样本）

- 联合分布函数　　联合密度函数

总体

- 什么是总体?
- 什么是数理统计中的总体?

相关概念

（容量　有限总体　无限总体）

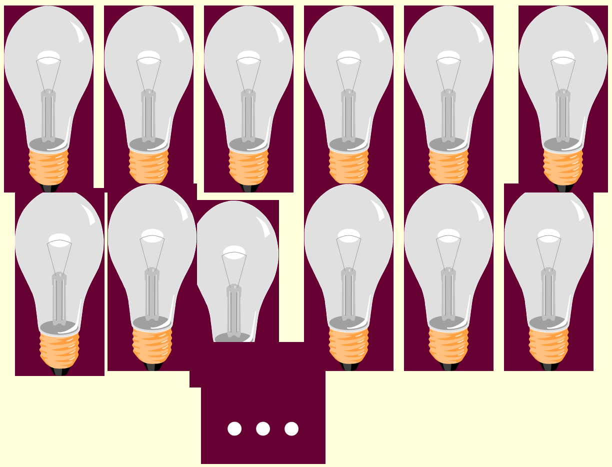
总 体

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.

例如



研究某批灯泡的质量

总 体

又例如：研究北京理工大学学生的学习情况

总体：北京理工大学全体学生

注：　总体随研究范围而定

总体中所包含的个体的个数称为**总体的容量**

容量有限的称为**有限总体**

容量无限的称为**无限总体**

例

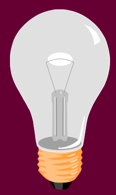
测量一湖泊任一地点深度

观察并记录某一地点每天最高温度

注　有限总体**容量很大**，可看作无限总体

然而在统计研究中，人们关心总体仅仅是关心其每个个体的一项(或几项)**数量指标**和**该数量指标在总体中的分布情况**.

某批
灯泡的寿命



该批灯泡寿命的
全体就是总体

国产轿车每公里耗油
量的全体就是总体

总体

每个个体具有的数量指标的全体

X

任取一个个体的数量指标

这样，总体就可以用一个**随机变量及其分布**来描述. 如说总体X或总体$F(x)$.

例如 检查自生产线出来的零件是正品还是次品

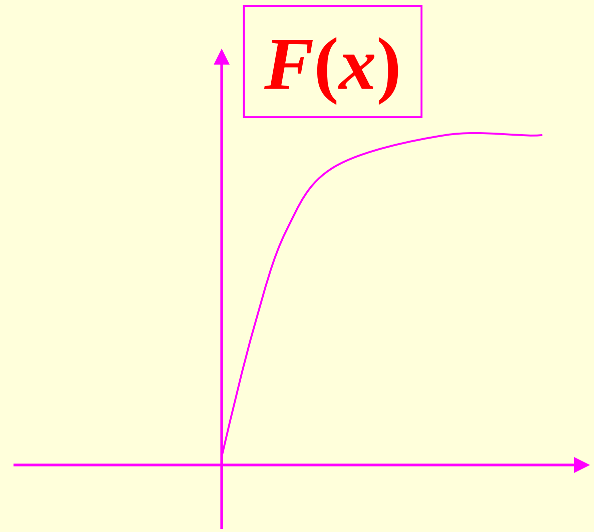
$$X=\begin{cases}1 & \text{正品}\\0 & \text{次品}\end{cases}$$

总体 X 对应的分布 （0 - 1分布）

$$P(X=x)=p^{x}(1-p)^{1-x},x=0,1.$$

例如:研究某批灯泡的寿命时，关心的数量指标就是寿命，那么，此总体就可以用随机变量X表示，或用其分布函数$F(x)$表示.

寿命 X 可用一概率分布来刻划



类似地，在研究某地区中学生的营养状况时，若关心的数量指标是身高和体重，我们用X和Y分别表示身高和体重，那么此总体就可用二维随机变量(X,Y)或其联合分布函数$F(x,y)$来表示.

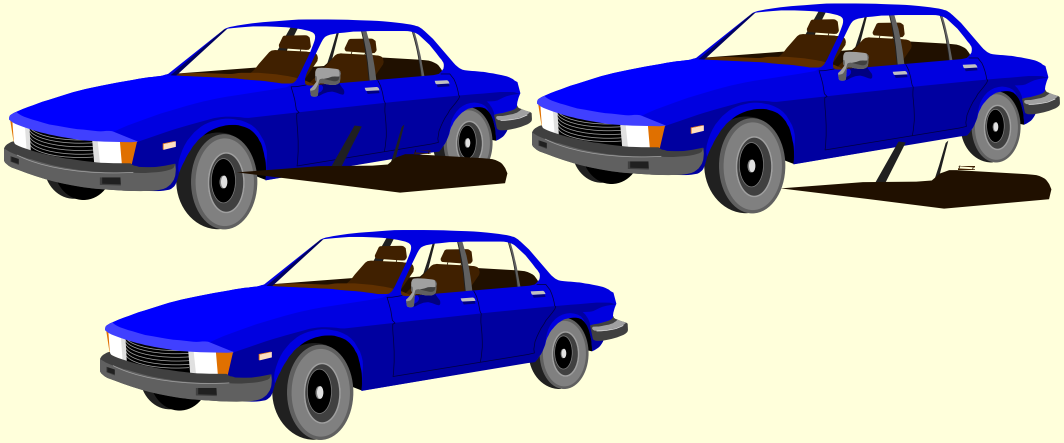
统计中，总体这个概念的要旨是：总体就是一个概率分布.

样本

为推断总体分布及各种特征，按一定规则从总体中抽取若干个体进行观察试验，以获得有关总体的信息，这一抽取过程称为“抽样”

所抽取的部分个体称为样本.

样本中所包含的个体数目称为样本容量.



从国产轿车中抽5辆
进行耗油量试验

样本容量为5

简单随机样本

1. 代表性： $X_1, X_2,\ldots, X_n$中每一个与所考察的总体X有相同的分布.

2. 独立性： $X_1, X_2,\ldots, X_{\mathrm{n}}$是相互独立的随机变量.

简单随机样本是应用中最常见的情形，今后，当说到“$X_1, X_2, \ldots, X_n$ 是来自某总体的样本”时，若不特别说明，就指简单随机样本.

上页 下页 返回

样本是随机变量

容量为n的样本可以看作n维随机向量$(X_1,X_2,\ldots,X_n)$.

但是，一旦取定一组样本，得到的是 n 个具体的数 $(x_1, x_2,\ldots, x_n)$，称为样本的一次观察值，简称样本值 .

总体、样本、样本值的关系

如我们从某班大学生中欲抽取10人测量身高，用 $X_1, X_2,\ldots,X_{10}$表示10个人的身高。进行试验后，得到10个数，$x_1, x_2,\ldots,x_{10}$，它们是样本取到的值．我们只能观察到随机变量取的值而见不到随机变量.

总体（理论分布）

?

样本 样本值

统计是从手中已有的资料--样本值，去推断总体的情况---总体分布$F(x)$的性质.

样本　是联系二者的桥梁

样本的联合分布

1　若总体X的分布函数为*F(x)*，

$X_1, X_2, \ldots, X_n$为来自总体的样本，

$(X_1, X_2, \ldots, X_n)$的 联合分布函数为

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

2　若总体X的密度函数为*f(x)*，

$(X_1, X_2, \ldots, X_n)$的 联合密度函数为

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

3 若总体*X*为离散型随机变量，其分布律为$P(X=y_K)=p_k,\ k=1,2,\ \cdots$

$(X_1, X_2,\ldots, X_n)$的 联合分布律为

$$f(x_1, x_2, \cdots x_n) = P(X_1 = x_1, X_2 = x_2, \cdots X_n = x_n)$$

$$= \prod_{i=1}^{n} P(X_i = x_i)$$

$$x_i \in \{y_1, y_2, \cdots\}, i = 1,2,\cdots$$

例1:总体X～$N(\mu,\sigma^2)$，样本（$X_1,X_2\ldots X_n$） 联合密度。

例2： 总体X～$B(1, p)$， 0<p<1写出其样本（$X_1,X_2,\ldots X_n$）的联合概率函数。

统计量

不含任何未知参数的样本的函数$g(X_1, X_2, \ldots, X_n)$称为统计量. 它是完全由 样本决定的量.

练习

$X \sim B(1, p)$, p未知，$X_1, \cdots X_n$为来自总体X的样本,下面那些是统计量

$$\sum_{i=1}^{n} X_i, \quad \max\{X_i\}-10, \quad X_n + pX_1, \quad \sum_{i=1}^{n}(X_i - \overline{X})^2$$

几个常见统计量

它反映了总体均值的信息

样本均值

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

它反映了总体方差的信息

样本方差

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

它反映了总体*k* 阶矩的信息

样本*k*阶原点矩 $$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

样本*k*阶中心矩 $$B_k = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^k$$

它反映了总体*k* 阶中心矩的信息

k=1,2,…

例如

$$A_1 = \overline{X}$$

$$B_2 = \frac{n-1}{n}S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \overset{\triangle}{=} S_n^2$$

性质：设总体X不论服从什么分布，只要其二阶矩存在，即$E(X)=\mu$、$D(X)=\sigma^2$都存在，则：

(1) $E(\overline{X})=E(X)=\mu$　　$D(\overline{X})=\dfrac{\sigma^2}{n}$

(2) $E(S^2)=D(X)=\sigma^2$

证明：

$$(2)S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

$$=\frac{1}{n-1}\sum_{i=1}^{n}(X_i^2-2X_i\bar{X}+\bar{X}^2)$$

$$=\frac{1}{n-1}[\sum_{i=1}^{n}X_i^2-n\overline{X}^2]$$

$$E(S^2)=\frac{1}{n-1}[\sum_{i=1}^{n}E(X_i^2)\ -nE(\overline{X}^2)]$$

$$=\frac{1}{n-1}[\sum_{i=1}^{n}(\mu^2+\sigma^2)-n(\mu^2+\frac{\sigma^2}{n})]=\sigma^2$$

顺序统计量

设$X_1, X_2, \ldots, X_n$是来自总体X的样本，将$X_1, X_2, \ldots, X_n$按照从小到大排列为

$$X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$$

称$(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$为$X_1, X_2, \ldots, X_n$的顺序统计量。

$X_{(k)}$, $1 \le k \le n$ 为第k个顺序统计量；

$X_{(1)} = \min\limits_{1 \le i \le n} X_i = \min(X_1, X_2, \cdots, X_n)$ 为最小顺序统计量；

$X_{(n)} = \max\limits_{1 \le i \le n} X_i = \max(X_1, X_2, \cdots, X_n)$ 为最大顺序统计量；

$R = X_{(n)} - X_{(1)}$ 为样本极差，反映了样本观察值的最大波动程度；

中位数

$$\tilde{X} = \begin{cases} X_{(\frac{n+1}{2})}, & n = 2m+1, \\ \frac{1}{2}\left(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}\right), & n = 2m. \end{cases}$$

称为样本中位数，它是把样本分为两部分，而样本中位数恰好是分界线。

第六章

作业1：1,2,3