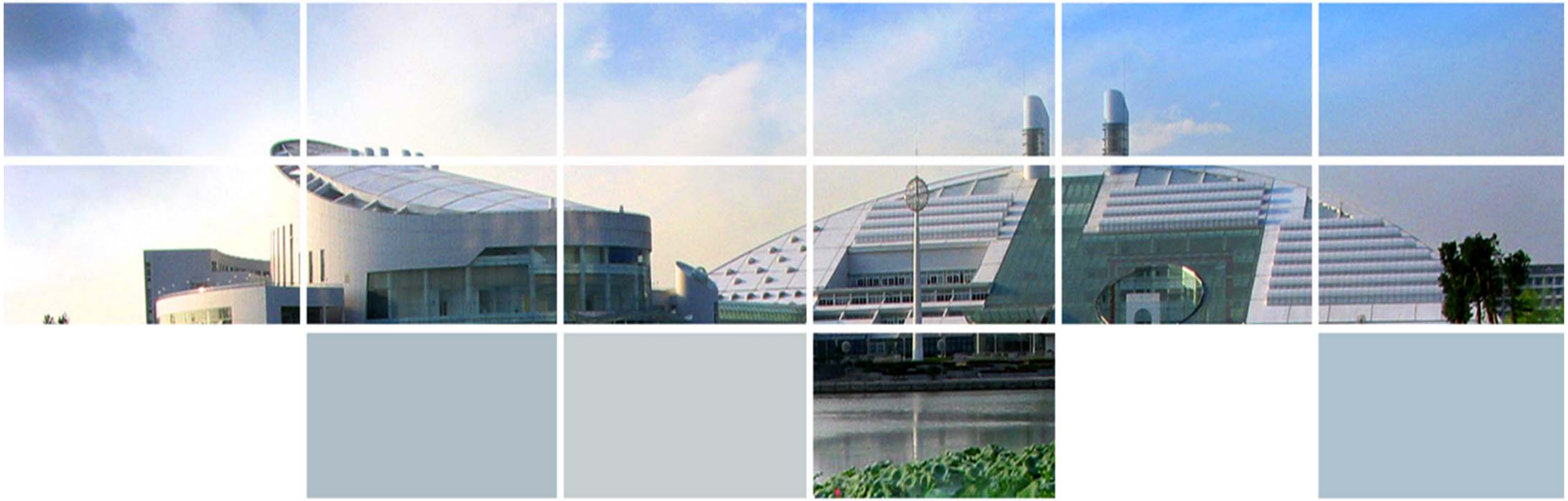

第39讲 统计量与常用统计量

在上一讲例3中，为了估计指数分布的参数λ,进行抽样观测，得到样本$X_1,\dots,X_{10}$和样本值

6394，1105， 4717, 1399, 7952,

17424, 3275, 21639, 2360, 2896.

样本中包含了许多信息。

对于推断总体的参数或分布而言，有些是有用的，重要的信息，有些则并不重要。

上例的样本至少提供了两种信息：

1）10个灯泡的平均寿命;—有用且重要的信息

2）灯泡寿命的序号(如6394是第1个).—不重要信息

从样本中提取有用的信息来研究总体的分布及各种特征数.——构造统计量.

- 统计量：样本的不含任何未知参数的函数。

设$(X_1, X_2, \ldots, X_n)$为样本, 若$g(X_1, X_2, \ldots, X_n)$不含任何未知参数, 则称$g(X_1, X_2, \ldots, X_n)$为统计量.

一旦有了样本观察值$x_1, x_2, \ldots, x_n$,就可以算出统计量的具体值$g(x_1, x_2, \ldots, x_n)$.

比如10个灯泡的平均寿命$(X_1 + X_2 + \ldots + X_{10})/10$是统计量.

平均寿命的观测值是6916.1小时.

常用统计量：

1\. 样本均值 $\bar{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i,$

2\. 样本方差 $S^2=\dfrac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2,$

样本标准差 $S=\sqrt{S^2}$

常用统计量：

3. 样本矩　　k阶矩：　$A_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$

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

$k = 1, 2, \ldots$

根据样本数据，用$Excel$计算$\bar{X}, S^2, B_2$,见实验11.

例1： 设X为总体,$E(X)=\mu$存在，$X_1,\ldots,X_n$是总体X的简单随机样本,则$\overline{X}=\mu$，对吗？

答：不对.

$E(X)=\mu$是一个数,可能已知,可能未知；

$\overline{X}$是随机变量,依赖于样本值,

对于不同的样本值,$\overline{X}$的取值可能不一样.

例2 接上一讲例2，总体为88，75，70，63，总体均值为74，总体方差为83.5.计算全部16个样本的样本均值，样本方差和样本二阶中心矩.

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

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

$$B_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

样本编号	样本	样本均值	样本方差	样本中心矩
1	(88,88)	88	0	0
2	(88,75)	81.5	84.5	42.25
3	(88,70)	79	162	81
4	(88,63)	75.5	312.5	156.25
5	(75,88)	81.5	84.5	42.25
6	(75,75)	75	0	0

样本编号	样本	样本均值	样本方差	样本中心矩
7	(75,70)	72.5	12.5	6.25
8	(75,63)	69	72	36
9	(70,88)	79	162	81
10	(70,75)	72.5	12.5	6.25
11	(70,70)	70	0	0
12	(70,63)	66.5	24.5	12.25
13	(63,88)	75.5	312.5	156.25
14	(63,75)	69	72	36
15	(63,70)	66.5	24.5	12.25
16	(63,63)	63	0	0
平均		74	83.5	41.75

与总体均值74相同

与总体方差83.5相同

比总体方差小

当总体数字特征未知时

- 用样本均值$\overline{X}$估计总体均值$\mu = E(X)$
- 用样本方差S^2估计总体方差$\sigma^2 = E(X-\mu)^2$
- 用样本原点矩A_k估计总体原点矩$\mu_k = E(X^k)$
- 用样本中心矩B_k估计总体中心矩$\nu_k = E(X-\mu)^k$

这些非常直观的想法，有什么理论依据吗？
这部分内容我们会在第44讲介绍。

- 统计量的分布被称为**抽样分布**.
- 当总体X服从一般分布（如指数分布、均匀分布等），要得出统计量的分布是很困难的.
- 当总体X服从正态分布时，统计量$\overline{X}, S^2$是可以计算的，那么服从什么分布呢？
- 下两讲我们将介绍数理统计中三个重要的抽样分布——χ^2分布，t分布，F分布.