第十八课

第六章 样本及抽样分布（习题课及单元测试）

内容提要

一、 随机样本

1、 总体、样本、容量

我们把对某一个问题的研究对象的全体称为总体，它是个随机变量。

组成总体的每一个基本单元称为个体。也是随机变量。

从总体中随机地抽取 n 个个体组成的集合就称为容量为 n 的一个样本，其中 n 称为样本的容量。

2、 简单随机样本

设 $X_1,X_2,\ldots,X_n$ 是来自总体 X 的容量为 n 的样本，又 $X_1,X_2,\ldots,X_n$ 相互独立且与总体 X 有相同的概率分布，则 $X_1,X_2,\ldots,X_n$ 就是一个简单随机样本。

3、 样本的数字特征

设 $X_1,X_2,\ldots,X_n$ 是总体 X 的一个样本，那么 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 分别称为

样本均值及样本方差。

4、 统计量

设 $X_1,X_2,\ldots,X_n$ 是总体 X 的一个样本。$X_1,X_2,\ldots,X_n$ 的任一函数 $g(X_1,X_2,\ldots,X_n)$，称为一个统计量。它是一个随机变量。

二、 抽样分布

1、 x^2 分布

2、 t 分布

3、 F 分布

4、 正态总体的样本均值与样本方差的分布

要求

1、 掌握数理统计的基本概念：总体、样本、容量、样本值、统计量、简单随机样本等

2、　了解几个常用的抽样分布，熟悉它们的概率密度函数的图形，会应用查表计算它们的分位点。

典型例题分析

例 1：举例说明总体、个体及样本等概念。

例如，对一批灯泡进行抽测了解使用寿命情况。了解一天生产的灯泡寿命，这一天生产的灯泡就是研究的对象全体，于是它就是总体，记为 X。这天生产的每一个灯泡都是一个个体。随机地从总体（这一天生产的所有灯泡）中抽 100 只，抽到的这 100 只就是一个样本，它的容量是 100，记为 $X_1,X_2,\ldots,X_{100}$。测试结果记录下来，譬如 $x_1=1832,x_2=1958,\ldots,x_{100}=1901$(小时)，这就是一个样本值。

例 2：设 $X_i\sim N(\mu_i,\sigma^2)$ $(i=1,2,\ldots,10)$，μ_i不全等。试问 $X_1,X_2,\ldots,X_{10}$是简单随机样本吗？

分析：据简单随机样本定义：i)它们来自一个总体；ii)它们相互独立；iii)它们同分布。用定义验证本题，第 i)、ii)两条设有交待，而第 iii)条不满足。因为μ_i不全等。故不是同分布。

解：因为 $X_i \sim N(\mu_i, \sigma^2)$ $(i=1,2,\ldots,10)$，μ_i 不全等，所以不是同分布，故 $X_1,X_2,\ldots,X_{10}$ 不是简单随机样本。

进一步：

若 $\mu_1=\mu_2=\ldots=\mu_{10}$，是否是简单随机样本？不一定。因为不能判断是不独立。

若 $X_1,X_2,\ldots,X_{10}$ 相互独立，μ_i 又全等，则 $X_1,X_2,\ldots,X_{10}$ 就是简单随机样本了。

例 3：设总体 X 服从两点分布 B(1,p)，其中 p 是未知参数，$X_1,X_2,\ldots,X_5$ 是从中抽取的一个样本，试指出在 X_1+X_2，$\min\limits_{1\le i\le 5}\{X_i\}$，$X_5+2p$，$(X_5-X_1)^2$，$X_3+1$，$X_4-4$ 之中那些是统计量，那些不是统计量，为什么？

分析：故 $X_1,X_2,\ldots,X_n$ 的一个函数 g(故 $X_1,X_2,\ldots,X_n$)若称为一个统计量，就不能含有未知参数，但可以含有已知常数。

解：由定义验证题中的 6 个式子中，X_5+2p 含有未知量 p，故不是统计量，其余均为统计量。

自测题:请返回在测试题中进行。(注意：后面还有第七章内容)

第七章 参数估计

设总体 X 的分布函数的形式为已知，但它的一个或多个参数为未知，如果有 X 的一组样本观察值 $x_1, x_2, \ldots, x_n$。会想到：用这组数据来估计总体的参数值。

原因：

这些样本观察值来至母体，
就应带有母体的某种信息。

参数估计 → 点估计
参数估计 → 区间估计

7. 1 点估计

定义：

设总体 X，分布函数 F(x)中含有未知参数 θ， θ 与数字特征(期望，方差等)有密切关系。用统计量 $\hat{\theta}(X_1, X_2, \ldots, X_n)$ 来估计 θ，取得观察值 $\hat{\theta}(x_1, x_2, \ldots, x_n)$ 作为 θ 的估计值。这种方法称为参数的点估计。

两种方法：

r 阶原点矩和 r 阶中心矩

总体的矩一般为数学期望和方差

一、 矩估计法

利用样本矩 $A_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$， $B_r = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^r$ 来对总体矩进行估计。

即： $E(X) \leftrightarrow \frac{1}{n}\sum_{i=1}^{n} X_i$

$D(X) \leftrightarrow \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$

实施方法：

在实际应用中一般是用样本的一阶原点矩与总体的数学期望相等来建立关系；用样本的二阶中心矩与总体的方差来建立关系。

例 1：设总体 X 在[a,b]上服从均匀分布

$$f(x,a,b)=\begin{cases}\dfrac{1}{b-a} & x\in[a,b]\\ 0 & x\notin[a,b]\end{cases}$$，试求参数 a,b 的矩法估计量。

解：因为 $E(X)=\dfrac{a+b}{2}$，$D(X)=\dfrac{(b-a)^2}{12}$，而 $\overline{X}=\dfrac{1}{n}\sum_{i=1}^{n}X_i$，$S^2(n)=\dfrac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$

所以可建立方程：$\begin{cases}\overline{X}=\dfrac{\hat{a}+\hat{b}}{2}\\ S^2(n)=\dfrac{(\hat{a}-\hat{b})^2}{12}\end{cases}$

由于这时的 a,b 为估计值，

所以记为 $\hat{a}$ 和 $\hat{b}$。

解得：$\hat{a}=\overline{X}-\sqrt{3}S(n)$，$\hat{b}=\overline{X}+\sqrt{3}S(n)$，这就是参数 a,b 的矩法估计值。

例 2：已知 X~B(n,p)，试求参数 n,p 的矩法估计值。

解：因为 E(X)=np，D(X)=np(1-p)，由样本的一阶原点矩和二阶中心矩及矩估计法知有：

$$\frac{1}{n}\sum_{i=1}^{n}X_i=\hat{n}\hat{p}，\quad \frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=\hat{n}\hat{p}(1-\hat{p})$$

可解得：$\hat{p}=1-\dfrac{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{\frac{1}{n}\sum\limits_{i=1}^{n}X_i}$，$\hat{n}=\dfrac{(\frac{1}{n}\sum\limits_{i=1}^{n}X_i)^2}{\frac{1}{n}\sum\limits_{i=1}^{n}X_i-\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$