第5章 ——数理统计的基础知识

- 本课程内容：研究随机现象的统计规律。
- 概率论是在分布已经知道情况下对随机变量进行研究。
- 数理统计的基本研究内容：推断随机变量的分布，即参数估计与假设检验。

作者：lfb

第5章 ——数理统计的基础知识

- ◆数理统计的基本概念（*****）
- ◆三大统计分布（*****）
- ◆正态总体的抽样分布（*****）

第5. 1节 ——数理统计的基本概念

◆总体，个体，样本（*****）

◆常用的统计量（*****）

◆ 总体，个体，样本

问题：检查一箱子螺钉的重量是否合格，怎么操作？

方法：随机选取部分螺钉进行检查，由检查结果推断整箱螺钉重量是否合格.

◆ 总体，个体，抽样，样本

总体：研究对象的全体构成的集合，每个成员称为**个体**；总体用表征其某指标的随机变量(或向量)X表示。

抽样：从总体中抽取部分个体进行观察这个过程即抽样。

样本：对总体进行抽样所得的部分个体构成的集合，其包含的个体数目称为样本容量，样本用表征其某指标的随机变量序列X1, X2, ---, Xn表示。

◆ 总体，个体，抽样，样本

总体： X

个体： X_i

样本： $X_1,X_2,\cdots,X_n$

样本容量： n

总体： $X \xrightarrow{\text{抽样}}$ 样本：$X_1,X_2,\cdots,X_n$

◆ 简单随机样本

$X_1, X_2, \cdots, X_n$独立同分布且与总体X同分布。

特征：（1）$X_1, X_2, \cdots, X_n$相互独立.

（2）$X_1, X_2, \cdots, X_n$同分布.

（3）个体X_i与总体X同分布.

◆ 总体分布与样本分布（略）

例：已知总体X的分布律如下，X_1,X_2,X_3为一样本，求$p\{X_1=1,X_2=2,X_3=1\}$.

X	1	2	3
P	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

样本特征：（1）$X_1,X_2,\cdots,X_n$相互独立.

（2）$X_1,X_2,\cdots,X_n$同分布.

（3）个体X_i与总体X同分布.

例：已知总体X的分布律如下，X_1, X_2, X_3为一样本，求$p\{X_1=1, X_2=2, X_3=1\}$.

X	1	2	3
P	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$

解：$p\{X_1=1, X_2=2, X_3=1\}=p\{X_1=1\}p\{X_2=2\}p\{X_3=1\}$

$$=p\{X=1\}p\{X=2\}p\{X=1\}$$

$$=\theta^2\cdot 2\theta(1-\theta)\theta^2$$

例：已知总体X的概率密度为$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & \text{else}\end{cases}$，$X_1,X_2,X_3$为一样本，

求$f_{X_1X_2X_3}(x_1,x_2,x_3)$.

样本特征：（1）$X_1,X_2,\cdots,X_n$相互独立.

（2）$X_1,X_2,\cdots,X_n$同分布.

（3）个体X_i与总体X同分布.

例：已知总体X的概率密度为$f_X(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & \text{else}\end{cases}$，$X_1,X_2,X_3$为一样本，

求$f_{X_1X_2X_3}(x_1,x_2,x_3)$.

解：$f_{X_1X_2X_3}(x_1,x_2,x_3)=f_{X_1}(x_1)f_{X_2}(x_2)f_{X_3}(x_3)$

$$=f_X(x_1)f_X(x_2)f_X(x_3)$$

$$=\begin{cases}\lambda e^{-\lambda x_1}\cdot\lambda e^{-\lambda x_2}\cdot\lambda e^{-\lambda x_3}, & x_1>0,x_2>0,x_3>0\\ 0, & \text{else}\end{cases}$$

$$=\begin{cases}\lambda^3 e^{-\lambda(x_1+x_2+x_3)}, & x_1>0,x_2>0,x_3>0\\ 0, & \text{else}\end{cases}$$

◆ **统计量：不包含总体分布中的未知参数的样本函数.**

◆ **常用统计量**

1. 样本均值： $\bar{X}=\dfrac{1}{n}\sum\limits_{i=1}^{n}X_i$

2. 样本方差： $S^2=\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2=\dfrac{1}{n-1}\left(\sum\limits_{i=1}^{n}X_i^{\ 2}-n\bar{X}^2\right)$

3. 样本标准差： $S=\sqrt{\dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i-\bar{X})^2}$

◆ 常用统计量

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

2. 样本方差：　$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^{\ 2}-n\bar{X}^2\right)$

3. 样本标准差：　$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2}$

4. 样本K阶原点矩：　$A_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k$

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

6. 顺序统计量：　$X_{(1)},X_{(2)},\cdots,X_{(n)}$

例：已知总体$X\sim N\left(\mu,\sigma^2\right)$，$X_1, X_2, \cdots X_n$为样本，　求$E\left(\bar{X}\right), D\left(\bar{X}\right)$

解：$E\left(\bar{X}\right)=E\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n}E\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E\left(X_i\right)=\frac{1}{n}\sum_{i=1}^{n}E\left(X\right)=E\left(X\right)=\mu$

$$D\left(\bar{X}\right)=D\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}D\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D\left(X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D\left(X\right)=\frac{1}{n}D\left(X\right)=\frac{1}{n}\sigma^2$$

结论：无论总体X服从什么分布，均有

$$E\left(\bar{X}\right)=E\left(X\right),\quad D\left(\bar{X}\right)=\frac{1}{n}D\left(X\right)$$

结论：无论总体X服从什么分布，均有

$$E(\bar{X})=E(X),\quad D(\bar{X})=\frac{1}{n}D(X)$$

例：已知总体$X\sim B(n,p)$，$X_1,X_2,\cdots X_n$为样本，　求$E(\bar{X}),D(\bar{X})$

解：$E(\bar{X})=E(X)=np,\quad D(\bar{X})=\dfrac{1}{n}D(X)=\dfrac{1}{n}np\cdot(1-p)=p\cdot(1-p)$

◆ 作业

习题5-1（Page138）： 8

8. 已知总体$X \sim e(\lambda)$，X_1, X_2为容量为2的样本，　求$X_{(1)}, X_{(2)}$的概率密度.

2：　$M=\max\{X,Y\}, N=\min\{X,Y\}$的分布$(X,Y$独立$)$

$$F_M(z)=F_X(z)F_Y(z)$$

$$F_N(z)=1-[1-F_X(z)][1-F_Y(z)]$$

推广：当X,Y独立同分布时：

$$F_M(z)=[F(z)]^2,\quad F_N(z)=1-[1-F(z)]^2$$

推广：设$X_1, X_2, \cdots, X_n$独立同分布，$M=\max\{X_1, X_2, \cdots, X_n\}, N=\min\{X_1, X_2, \cdots, X_n\}$，则

$$F_M(z)=[F(z)]^n,\quad F_N(z)=1-[1-F(z)]^n$$

8. 已知总体$X\sim e(\lambda)$，X_1,X_2为容量为2的样本，　求$X_{(1)},X_{(2)}$的概率密度.

解：（1）由题目知：$F_X(x)=\begin{cases}1-e^{-\lambda x}, x>0\\ 0, \quad \text{else}\end{cases}$，　$X_{(1)}=\min\{X_1,X_2\}$

$$F_{X_{(1)}}(t)=1-[1-F_{X_1}(t)]\cdot[1-F_{X_2}(t)]=1-[1-F_X(t)]\cdot[1-F_X(t)]=1-[1-F_X(t)]^2$$

$$=\begin{cases}1-e^{-2\lambda t}, & t>0\\ 0, & \text{else}\end{cases} \Rightarrow f_{X_{(1)}}(t)=F'_{X_{(1)}}(t)=\begin{cases}2\lambda e^{-2\lambda t}, & t>0\\ 0, & \text{else}\end{cases}$$

注：$X_{(1)}\sim e(2\lambda)$

8. 已知总体$X \sim e(\lambda)$，X_1, X_2为容量为2的样本，求$X_{(1)}, X_{(2)}$的概率密度.

解：(2) 由题目知：$F_X(x)=\begin{cases}1-e^{-\lambda x}, x>0\\ 0, \quad \text{else}\end{cases}$， $X_{(2)}=\max\{X_1, X_2\}$

$$F_{X_{(2)}}(t)=F_{X_1}(t)F_{X_2}(t)=F_X(t)^2=\begin{cases}\left(1-e^{-\lambda t}\right)^2, t>0\\ 0, \quad \text{else}\end{cases}$$

$$\Rightarrow f_{X_{(2)}}(t)=F'_{X_{(2)}}(t)=\begin{cases}2\lambda e^{-\lambda t}\left(1-e^{-\lambda t}\right), \ t>0\\ 0, \quad \text{else}\end{cases}$$