Chapter 5

数理统计的基本知识

第五章作业（2019.05.27交）：习题五5.1, 5.9, 5.10(1), 5.11, 5.14, 5.18

5.1 已知样本观测值为

15.8	24.2	14.5	17.4	13.2	20.8
17.9	19.1	21.0	18.5	16.4	22.6

计算样本均值、样本方差与样本二阶中心矩的观测值。

解： 样本均值$\overline{X}$的观测值

$$\bar{x} = \frac{1}{12}\sum_{i=1}^{12} x_i = 18.45,$$

样本方差S^2的观测值

$$s^2 = \frac{1}{11}\sum_{i=1}^{12}(x_i - \bar{x})^2 \approx 10.7755,$$

样本二阶中心矩U_2的观测值

$$\widetilde{\sigma}^2 = \frac{1}{12}\sum_{i=1}^{12}(x_i - \bar{x})^2 = 9.8775.$$

□

5.9 设总体$X \sim N(40, 5^2)$，

(1) 抽取容量为36的样本，求样本均值$\overline{X}$在38与43之间的概率；

(2) 抽取容量为64的样本，求$|\overline{X}-40|<1$ 的概率；

(3) 抽取样本容量n多大时，才能使概率$P\{|\overline{X}-40|<1\}$达到0.95？

解： 已知总体$X \sim N(40, 5^2)$，则统计量

$$u = \frac{\overline{X}-40}{5/\sqrt{n}} \sim N(0,1).$$

(1) 因$n=36$，所以有

$$u = \frac{\overline{X}-40}{5/6} \sim N(0,1).$$

由此得所求概率

$$P\{38<\overline{X}<43\} = P\left\{-2.4 < \frac{\overline{X}-40}{5/6} < 3.6\right\} = \Phi(3.6)-\Phi(-2.4)$$
$$= 0.99984-(1-0.9918) = 0.99164.$$

(2) 因$n=64$，所以有

$$u = \frac{\overline{X}-40}{5/8} \sim N(0,1).$$

由此得所求概率

$$P\{|\overline{X}-40|<1\} = P\left\{-1.6 < \frac{\overline{X}-40}{5/8} < 1.6\right\}$$
$$= 2\Phi(1.6)-1 = 2\times 0.9452 = 0.8904.$$

(3)

$$P\{|\overline{X}-40|<1\} = P\left\{-\frac{\sqrt{n}}{5} < \frac{\overline{X}-40}{5/\sqrt{n}} < \frac{\sqrt{n}}{5}\right\} = 2\Phi\left(\frac{\sqrt{n}}{5}\right)-1.$$

依题意有，

$$2\Phi\left(\frac{\sqrt{n}}{5}\right)-1 = 0.95, \quad \text{即} \quad \Phi\left(\frac{\sqrt{n}}{5}\right) = 0.975.$$

查表得$\Phi(1.96) = 0.975$，从而有

$$\frac{\sqrt{n}}{5} = 1.96,$$

解得$n \approx 96$。

□

5.10 设总体$X \sim N(\mu, \sigma^2)$，从总体中抽取容量为16的样本，

(1) 已知$\sigma = 2$，求概率$P\{|\overline{X} - \mu| < 0.5\}$；

解：

(1) 已知总体$X \sim N(\mu, 2^2)$，样本容量$n = 16$，得

$$\frac{\overline{X} - \mu}{2/4} \sim N(0, 1).$$

有

$$P\{|\overline{X} - \mu| < 0.5\} = P\left\{\left|\frac{\overline{X} - \mu}{2/4}\right| < 1\right\}$$
$$= 2\Phi(1) - 1 = 2 \times 0.8413 - 1 = 0.6826.$$

□

5.11 设总体$X \sim N(\mu, 2^2)$，收取容量为20的样本$X_1, X_2, \cdots, X_{20}$，

(1) 已知μ，求概率$P\left\{43.6 \leq \sum_{i=1}^{20}(X_i - \mu)^2 \leq 150.4\right\}$；

(2) 未知μ，求概率$P\left\{46.8 \leq \sum_{i=1}^{20}(X_i - \overline{X})^2 \leq 154.4\right\}$。

解：

(1) 已知总体$X \sim N(\mu, 2^2)$，其中μ已知，样本容量$n=20$，则统计量

$$\chi^2 = \frac{1}{2^2}\sum_{i=1}^{20}(X_i-\mu)^2 \sim \chi^2(20),$$

所以有

$$\begin{aligned} P\left\{43.6 \le \sum_{i=1}^{20}(X_i-\mu)^2 \le 150.4\right\} &= P\left\{10.9 \le \frac{1}{2^2}\sum_{i=1}^{20}(X_i-\mu)^2 \le 37.6\right\} \\ &= P\{10.9 \le \chi^2 \le 37.6\} \\ &= P\{\chi^2 \ge 10.9\} - P\{\chi^2 > 37.6\}. \end{aligned}$$

查表知

$$\chi^2_{0.95}(20) = 10.9, \quad \chi^2_{0.01}(20) = 37.6.$$

所以，

$$P\left\{43.6 \le \sum_{i=1}^{20}(X_i-\mu)^2 \le 150.4\right\} = 0.95 - 0.01 = 0.94.$$

(2) 已知总体$X \sim N(\mu, 2^2)$，其中μ未知，样本容量$n=20$，则统计量

$$\chi^2 = \frac{(20-1)S^2}{2^2} = \frac{1}{2^2}\sum_{i=1}^{20}(X_i-\overline{X})^2 \sim \chi^2(19),$$

所以有

$$\begin{aligned} P\left\{46.8 \le \sum_{i=1}^{20}(X_i-\overline{X})^2 \le 154.4\right\} &= P\left(11.7 \le \frac{1}{2^2}\sum_{i=1}^{20}(X_i-\overline{X})^2 \le 38.6\right) \\ &= P\{11.7 \le \chi^2 \le 38.6\} \\ &= P\{\chi^2 \ge 11.7\} - P\{\chi^2 > 38.6\}. \end{aligned}$$

查表知

$$\chi^2_{0.90}(19) = 11.7, \quad \chi^2_{0.005}(19) = 38.6$$

所以，

$$P\left\{46.8 \le \sum_{i=1}^{20}(X_i-\overline{X})^2 \le 154.4\right\} = 0.90 - 0.005 = 0.895.$$

□

5.14 设总体$X \sim (\mu, \sigma^2)$, 抽取样本$X_1, X_2, \cdots, X_n$,样本均值为$\overline{X}$,样本方差为S^2. 如果再抽取一个样本X_{n+1}, 证明：统计量

$$\sqrt{\frac{n}{n+1}}\frac{X_{n+1}-\overline{X}}{S} \sim t(n-1).$$

解： 已知总体$X \sim N(\mu, \sigma^2)$,则$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$.

若再抽取一个样本X_{n+1},则$X_{n+1} \sim N(\mu, \sigma^2)$.

由于所有样本$X_1, X_2, \cdots, X_n, X_{n+1}$是相互独立的，

所以$\overline{X}$与X_{n+1}也相互独立的，则

$$X_{n+1} \sim \overline{X} \sim N(0, \frac{n+1}{n}\sigma^2).$$

由此得到标准化的统计量

$$U = \frac{X_{n+1}-\overline{X}}{\sqrt{\frac{n+1}{n}}\sigma} \sim N(0,1).$$

又由于统计量$\chi^2 = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

因为$\overline{X}$与S^2是独立的，所以统计量U与X^2也是相互独立的.则按照t分布的定义可知，统计量

$$\frac{U}{\sqrt{\frac{\chi^2}{(n-1)}}} = \sqrt{\frac{n}{n+1}}\frac{X_{n+1}-\overline{X}}{S} \sim t(n-1).$$

□

5.18 设总体的分布函数为$F(x)$，概率密度为$f(x)$，抽取样本$X_1, X_2, \cdots, X_n$，求：

(1) 样本最大值$\max(X_1, X_2, \cdots, X_n)$的概率密度；

(2) 样本最小值$\min(X_1, X_2, \cdots, X_n)$的概率密度。

解：

(1) 因为样本与总体服从相同的分布，所以X_i的分布函数及概率密度分别是

$$F_i(x) = F(x), \quad f_i(x) = f(x), \quad i = 1, 2, \cdots, n.$$

又因为样本X_i $(i = 1, 2, \cdots, n)$相互独立，则样本最大值的分布函数

$$F_{\max}(x) = [F(x)]^n,$$

对x求导得概率密度

$$f_{\max}(x) = n[F(x)]^{n-1} f(x).$$

(2) 样本最小值$\min(X_1, X_2, \cdots, X_n)$的分布函数

$$F_{\min}(x) = 1 - [1 - F(x)]^n,$$

对x求导得概率密度

$$f_{\min}(x) = n[1 - F(x)]^{n-1} f(x).$$

□

作业情况：

1 本次作业大家完成得挺好，但仍有11位同学未交作业，希望同学们下次补上；

2 5.9(2)(3)计算错误的较多；有部分同学5.11、5.18没有写出解题过程，须注意。