Chapter 5

数理统计的基本知识

5.9 设$X_1, X_2, \cdots, X_{15}$是来自正态总体$N(0, 2^2)$的简单随机样本，求统计量

$$Y = \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{2(X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2)}$$

所服从的概率分布。

解： 因为$X_i \sim N(0, 2^2)$，所以$\dfrac{X_i}{2} \sim N(0,1)$，$(\dfrac{X_i}{2})^2 \sim \chi^2(1)$，

$$\xi = \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{4} \sim \chi^2(10),$$
$$\eta = \frac{X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2}{4} \sim \chi^2(5).$$

故

$$Y = \frac{X_1^2 + X_2^2 + \cdots + X_{10}^2}{2(X_{11}^2 + X_{12}^2 + \cdots + X_{15}^2)} = \frac{\xi/10}{\eta/5} \sim F(10, 5).$$

□

5.10 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

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

(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

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

(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，

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

查表得$\Phi(1.96) = 0.975$，从而有

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

解得$n \approx 96$。

□

5.12 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$P\left\{43.6 \le \sum_{i=1}^{20}(X_i-\mu)^2 \le 150.4\right\}$；

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

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$X\sim N(50,6^2)$，总体$Y\sim N(46,4^2)$，从总体X中抽取容量为10的样本，从总体Y中抽取容量为8的样本，求下列概率：

(1) $P\{0<\overline{X}-\overline{Y}<8\}$；

(2) $P\left\{\dfrac{S_1^2}{S_2^2}<8.28\right\}$。

解：

(1) 因为总体$X\sim N(50,6^2), Y\sim N(46,4^2), n_1=10, n_2=8$，可以得到统计量

$$U=\frac{(\overline{X}-\overline{Y})-(50-46)}{\sqrt{\dfrac{6^2}{10}+\dfrac{4^2}{8}}}=\frac{\overline{X}-\overline{Y}-4}{\sqrt{5.6}}\sim N(0,1).$$

所以

$$\begin{aligned}P\{0<\overline{X}-\overline{Y}<8\}&=P\left\{-\frac{4}{\sqrt{5.6}}<\frac{\overline{X}-\overline{Y}-4}{\sqrt{5.6}}<\frac{4}{\sqrt{5.6}}\right\}\\&=2\Phi(\frac{4}{\sqrt{5.6}})-1\approx 2\Phi(1.69)-1\approx 2\times 0.9545-1=0.909.\end{aligned}$$

(2) 已知总体$X\sim N(50,6^2), Y\sim N(46,4^2), n_1=10, n_2=8$，可以得到统计量

$$F=\frac{S_1^2/6^2}{S_2^2/4^2}=\frac{4S_1^2}{9S_2^2}\sim F(9,7).$$

则有

$$P\left\{\frac{S_1^2}{S_2^2}<8.28\right\}=P\left\{\frac{4S_1^2}{9S_2^2}<3.68\right\}=P(F<3.68)=1-P(F\ge 3.68).$$

记$\alpha=P(F\ge 3.68)$，也就是说上分位点$F_\alpha(9,7)=3.68$，查表得$\alpha=0.05$。由此得所求概率为

$$P\left\{\frac{S_1^2}{S_2^2}<8.28\right\}=1-0.05=0.95.$$

□

5.15 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

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

证明： 已知总体$X \sim N(\mu, \sigma^2)$，则$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$，$X_{n+1} \sim N(\mu, \sigma^2)$。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

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

由此得到标准化的统计量

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

又由于$\overline{X}$、X_{n+1}分别与S^2相互独立，所以统计量U与S^2也是相互独立的。统计量$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$，则按照$t$分布的性质可知，统计量

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

□

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

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

□