5.4 三大抽样分布

2. 设 $x_1, x_2, \cdots, x_n$ 是来自 $N(\mu, 16)$ 的样本，问 n 多大时才能使得 $P(|\bar{x}-\mu|<1) \geqslant 0.95$ 成立?

解：

由于 $\bar{x} \sim N(\mu, \frac{16}{n})$，根据切比雪夫不等式，

$$P(|\bar{x}-E\bar{x}|<\varepsilon) \geqslant 1-\frac{\operatorname{Var}\bar{x}}{\varepsilon}$$

$$P(|\bar{x}-\mu|<1) \geqslant 1-\frac{16}{n}$$

$$\frac{16}{n}=0.95 \Rightarrow n=\frac{16}{0.95}=\frac{320}{19} \approx 16.8421052631579$$

因为 n 为整数，所以 n 至少为 17 时才能使得 $P(|\bar{x}-\mu|<1) \geqslant 0.95$ 成立。 □

4. 由正态总体 $N(\mu, \sigma^2)$ 抽取容量为 20 的样本，试求 $P\left(10\sigma^2 \leqslant \sum_{i=1}^{20}(x_i-\mu)^2 \leqslant 30\sigma^2\right)$。

解：

由于 $s^2=\frac{1}{19}\sum_{i=1}^{20}(x_i-\mu)^2$, $\frac{19s^2}{\sigma^2} \sim \chi^2(n-1)$， 所以 $\frac{1}{\sigma^2}\sum_{i=1}^{20}(x_i-\mu)^2 \sim \chi^2(n-1)$。 所以

$$\begin{aligned} &P\left(10\sigma^2 \leqslant \sum_{i=1}^{20}(x_i-\mu)^2 \leqslant 30\sigma^2\right)=P\left(10 \leqslant \frac{1}{\sigma^2}\sum_{i=1}^{20}(x_i-\mu)^2 \leqslant 30\right) \\ &=\int_{10}^{30} \frac{\left(\frac{1}{2}\right)^{\frac{20}{2}}}{\Gamma\left(\frac{20}{2}-1\right)} y^{\frac{20}{2}} e^{-\frac{y}{2}} \mathrm{d}y=\int_{10}^{30} \frac{\left(\frac{1}{2}\right)^{10}}{9!} y^9 e^{-\frac{y}{2}} \mathrm{d}y \approx 0.898318281994385 \end{aligned}$$

$$\int_{10}^{30} \frac{\left(\frac{1}{2}\right)^{10}}{9!} x^9 e^{-\frac{x}{2}} \mathrm{d}x$$

0.898318282

□

6. 设 $x_1, x_2, \cdots, x_n$ 是来自 $N(\mu, 1)$ 的样本，试确定最小的常数 c，使得对任意的 $\mu \geqslant 0$，有 $P(|\bar{x}|<c) \leqslant \alpha$。

解：

这题什么意思？当 $\mu=0$ 时，当 $n \to \infty$ 时 $\bar{x} \to 0$, $P(|\bar{x}|<c) \to 1$，怎么可能 $P(|\bar{x}|<c) \leqslant \alpha$ 呢？ □

8. 设随机变量 $X \sim F(n,m)$，证明 $Z = \frac{n}{m}X \Big/ \left(1+\frac{n}{m}X\right)$ 服从贝塔分布，并指出其参数。

证明：

设 $Y = \frac{n}{m}X$，则

$$p_Y(y) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} y^{\frac{m}{2}-1}(1+y)^{-\frac{m+n}{2}}$$

$$Z = \frac{Y}{1+Y} = 1 - \frac{1}{1+Y} \Longrightarrow Y = \frac{1}{1-Z} - 1 = \frac{Z}{1-Z}$$

所以

$$\begin{aligned} p_Z(z) &= \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\left(\frac{z}{1-z}\right)^{\frac{m}{2}-1}\left(\frac{1}{1-z}\right)^{-\frac{m+n}{2}} \\ &= \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}\left(\frac{z}{1-z}\right)^{\frac{m}{2}}\left(\frac{1-z}{z}\right)(1-z)^{\frac{m}{2}}(1-z)^{\frac{n}{2}} \\ &= \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} z^{\frac{m}{2}}(1-z)^{\frac{n}{2}}\left(\frac{1-z}{z}\right) \\ &= \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} z^{\frac{m}{2}-1}(1-z)^{\frac{n}{2}+1} \end{aligned}$$

所以 Z 服从贝塔分布，其参数为 $\frac{m}{2}$ 和 $\frac{n}{2}$。 □

9. 设 x_1, x_2 是来自 $N(0,\sigma^2)$ 的样本，试求 $Y = \left(\frac{x_1+x_2}{x_1-x_2}\right)^2$ 的分布。

解：

$$Y = \left(\frac{x_1+x_2}{x_1-x_2}\right)^2 = \left(\frac{\frac{x_1}{x_2}+1}{\frac{x_1}{x_2}-1}\right)^2 = \left(1+\frac{2}{\frac{x_1}{x_2}-1}\right)^2$$

其中 $\frac{x_1}{x_2}$ 的概率密度函数为

$$\begin{aligned} p_{\frac{x_1}{x_2}}(t) &= \int_{-\infty}^{+\infty} |x|\, p(x,tx)\mathrm{d}x = \int_{-\infty}^{+\infty} |x|\, \phi(x)\phi(tx)\mathrm{d}x = \int_{-\infty}^{+\infty} |x| \frac{1}{\sqrt{2\pi}} e^{\frac{-\sigma^2 x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-\sigma^2 t^2 x^2}{2}}\mathrm{d}x \\ &= \frac{1}{\sigma^2\pi}\int_0^{+\infty} x e^{\frac{-\sigma^2 x^2}{2}(1+t^2)}\mathrm{d}x = \frac{1}{\pi\sigma^4(t^2+1)} \end{aligned}$$

设随机变量 $Z = 1+\frac{2}{\frac{x_1}{x_2}-1}$，则 $\frac{x_1}{x_2} = 1+\frac{2}{Z-1}, Y = Z^2$，所以

$$p_Z(z) = \frac{1}{\pi\sigma^4\left[\left(1+\frac{2}{z-1}\right)^2+1\right]} = \frac{(z-1)^2}{\pi\sigma^4((z-1)^2+(z+1)^2)}$$

$$\begin{aligned} p_Y(y) &= p_Z(\sqrt{y}) + p_Z(-\sqrt{y}) = \frac{(\sqrt{y}-1)^2}{\pi\sigma^4((\sqrt{y}-1)^2+(\sqrt{y}+1)^2)} + \frac{(\sqrt{y}-1)^2}{\pi\sigma^4((\sqrt{y}-1)^2+(\sqrt{y}+1)^2)} \\ &= \frac{-2\sqrt{y}+y+1}{\pi\sigma^4(y+1)} \end{aligned}$$

此即为 Y 的概率密度函数。 □