《概率论与数理统计》习题

第十五讲　统计量

1. 设 x_1, x_2, x_n 和 $y_1, y_2, \cdots, y_n$ 是两组样本观测值，且关系如下：

$$y_i = ax_i + b, i = 1, 2, \cdots, n$$

其中，a 和 b 为非零常数。试求：

(a) 样本均值 $\bar{x}$ 和 $\bar{y}$ 间的关系；

(b) 样本方差 s_x^2 和 s_y^2 间的关系。

2. 设总体 X 的 3 阶矩存在，若 $x_1, x_2, \cdots, x_n$ 是取自该总体的简单随机样本，$\bar{x}$ 为样本均值，s^2 为样本方差，试证：

$$\mathrm{Cov}(\bar{x}, s^2) = \frac{\nu_3}{n},$$

其中 $\nu_3 = E(x - E(x))^3$。

3. 设 $\bar{x}_1$ 和 $\bar{x}_2$ 是从同一正态总体 $N(\mu, \sigma^2)$ 独立抽取的容量相同的两个样本均值。试确定样本容量 n，使得两样本均值的差超过 σ 的概率不超过 0.01。

4. 从指数总体 $Exp(1/\theta)$ 抽取了 40 个样品，试求 $\bar{x}$ 的渐近分布。

5. 设总体 X 的分布函数 $F(x)$ 是连续的，$x_{(1)}, x_{(2)}, \cdots, x_{(n)}$ 为取自此总体的次序统计量，设 $\eta_i = F(x_{(i)})$，试证：

(a) $\eta_1 \le \eta_2 \le \cdots \le \eta_n$，$\eta_i$ 是来自均匀分布 $U(0,1)$ 总体的次序统计量；

(b) $E(\eta_i) = \frac{i}{n+1}$，$\mathrm{Var}(\eta_i) = \frac{i(n+1-i)}{(n+1)^2(n+2)}, 1 \le i \le n$；

(c) η_i 和 η_j 的协方差矩阵为

$$\begin{pmatrix} \frac{a_1(1-a_1)}{n+2} & \frac{a_1(1-a_2)}{n+2} \\ \frac{a_2(1-a_1)}{n+2} & \frac{a_2(1-a_2)}{n+2} \end{pmatrix}$$

其中，$a_1 = \frac{i}{n+1}, a_2 = \frac{j}{n+1}$.

6. 假定总体的密度函数为

$$p(x) = 6x(1-x), 0 < x < 1.$$

求样本量为 n 的样本中位数 $m_{0.5}$ 的渐近分布。

第十六讲　三大抽样分布

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

2. 设随机变量 $X \sim F(n, m)$，证明：

$$Z = \frac{\frac{n}{m}X}{(1 + \frac{n}{m}X)}$$

服从贝塔分布，并指出其参数。

3. 设 $x_1, x_2, \cdots, x_n$ 是来自 $N(\mu_1, \sigma^2)$ 的样本，$y_1, y_2, \cdots, y_m$ 是来自 $N(\mu_2, \sigma^2)$ 的样本，c, d 是任意两个不为 0 的常数，证明

$$t = \frac{c(\bar{x} - \mu_1) + d(\bar{y} - \mu_2)}{s_w\sqrt{\frac{c^2}{n} + \frac{d^2}{n}}} \sim t(n + m - 2)$$

其中 $s_w^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$。

(编程题) 使用随机模拟的方法得到样本峰度和样本偏度的分布（输出分位数）。