

第六章 数理统计的基本概念

1. 设总体 $\xi \sim N(5, 2)$, $\xi_1, \xi_2, \dots, \xi_9$ 为其样本, 试求样本的平均值 $\bar{\xi}$ 大于 8 的概率。
解:

$$\begin{aligned} \because \bar{\xi} \sim N\left(a, \frac{\sigma^2}{n}\right) = N\left(5, \frac{2}{3}\right) \\ \therefore P\{\bar{\xi} > 8\} = P\left\{\frac{\bar{\xi} - 5}{\sqrt{\frac{2}{3}}} > \frac{8 - 5}{\sqrt{\frac{2}{3}}}\right\} = 4.5\} \\ = 1 - \phi(4.5) = 0.598706326 \end{aligned}$$

3. 设总体 ξ 服从正态分布 $N(0, \sigma)$, $\xi_1, \xi_2, \dots, \xi_4$ 为其样本, 试问

$$\eta = \frac{(\xi_1 - \xi_2)^2}{(\xi_3 + \xi_4)^2} \text{ 服从什么分布?}$$

解:

$$\begin{array}{c} \xi_1 - \xi_2 \sim N\left(0, \frac{\sigma^2}{2}\right) \\ \xi_3 + \xi_4 \sim N\left(0, \frac{\sigma^2}{2}\right) \end{array} \Rightarrow \begin{array}{c} \frac{\xi_1 - \xi_2}{\sigma} \sim N(0, 1) \\ \frac{\xi_3 + \xi_4}{\sigma} \sim N(0, 1) \end{array}$$

$$\begin{array}{c} \left(\frac{\xi_1 - \xi_2}{\sigma}\right)^2 \sim \chi^2(1) \\ \left(\frac{\xi_3 + \xi_4}{\sigma}\right)^2 \sim \chi^2(1) \end{array} \Rightarrow \frac{(\xi_1 - \xi_2)^2}{(\xi_3 + \xi_4)^2} \sim F(1, 1)$$

4. 设总体 $\xi \sim N(1, 2)$, $\xi_1, \xi_2, \dots, \xi_4$ 为其样本, 记 $\eta = k[\sum_{i=1}^4 \xi_i - 4]^2$, 试问 k 取何值时, 使得 η 服从 $\chi^2(m)$ 分布, 自由度 m 取何值?

解:

$$\xi \sim N(1, 2)$$

$$\sum_{i=1}^4 \xi_i \sim N(4, 16) \Rightarrow \frac{\sum_{i=1}^4 \xi_i - 4}{4} \sim N(0, 1)$$

$$\Rightarrow \frac{(\sum_{i=1}^4 \xi_i - 4)^2}{16} \sim \chi^2(1)$$

$$\therefore k = \frac{1}{16}, m = 1$$

5. 设 $\xi \sim N(3, 2)$, $\xi_1, \xi_2, \dots, \xi_{16}$ 为其样本, $\bar{\xi}$ 与 S_n^2 分别为样本的均值与方差, 试建立 t 分

布的统计量。

解：

$$\sqrt{n-1} \frac{\bar{\xi} - \alpha}{S_n} = \sqrt{15} \frac{\bar{\xi} - 3}{S_n} \sim t(15)$$

6. 设正态总体 $\xi \sim N(5, 6)$, $n, \bar{\xi}$ 分别为样本容量和样本均值, 试问 n 应取多大, 才能使得 $\bar{\xi}$ 位于区间 $(3, 7)$ 概率不小于 0.90

解：

$$\xi \sim N(5, 6)$$

$$P\{3 < \bar{\xi} < 7\} = P\left\{-\frac{\sqrt{n}}{2} < \frac{3-5}{6/\sqrt{n}} < \frac{\bar{\xi}-5}{6/\sqrt{n}} < \frac{7-5}{6/\sqrt{n}} = \frac{\sqrt{n}}{2}\right\}$$

$$= \phi\left(\frac{\sqrt{n}}{2}\right) - \phi\left(-\frac{\sqrt{n}}{2}\right) = 2\phi\left(\frac{\sqrt{n}}{2}\right) - 1 \geq 0.9$$

$$\Rightarrow \phi\left(\frac{\sqrt{n}}{2}\right) \geq 0.95 \Rightarrow n \geq 25$$

7. 设总体 $\xi \sim E(\lambda)$, $\xi_1, \xi_2, \dots, \xi_n$ 为其样本, $\bar{\xi}$ 为样本均值:

1) 试求 $\eta = 2n\bar{\xi}$ 的分布。

2) 若 $n=1$, 试问 $P\{\eta > 6\}$ 是何值?

解：

$$\varphi_{\xi}(t) = (1 - \frac{1}{\lambda}it)^{-1}, \varphi_{\bar{\xi}}(t) = (1 - \frac{1}{n\lambda}it)^{-n}$$

$$\varphi_{2n\bar{\xi}}(t) = (1 - 2n\lambda \frac{1}{n\lambda}it)^{-n} = (1 - 2it)^{-n}$$

$$\Rightarrow 2n\lambda \bar{\xi} \sim G\left(\frac{1}{2}, \frac{2n}{2}\right) = \Gamma(2, \frac{2n}{2}) = \chi^2(2n)$$

$$P\{\eta > 6\} = 1 - P\{\eta \leq 6\} = 0.950212932$$

8. 设总体 $\xi \sim N(12, 2)$, 今抽取容量为 5 的样本 $\xi_1, \xi_2, \dots, \xi_5$, 试问:

1) 样本均值 $\bar{\xi}$ 大于 13 的概率是多少?

2) 样本的极小值小于 10 的概率是多少?

3) 样本的极大值大于 15 的概率是多少?

解：

$$1) P\{\bar{\xi} > 13\} = P\left\{\frac{\bar{\xi} - 12}{2/\sqrt{5}} > \frac{13 - 12}{2/\sqrt{5}} = 1.11803\right\} = 1 - \phi(1.11803)$$

$$= 0.13177709$$

$$2) P\{\xi(1) < 10\} = 1 - [1 - F_{\xi}(10)]^5$$

$$= 1 - (0.841344746)^5 = 0.57843$$

$$3) P\{\xi(5) \geq 15\} = 1 - [F_{\xi}(15)]^5 = 1 - 0.933192799^5 = 0.292287455$$

9. 设电子元件的寿命(时数) ξ 服从服从以 $\lambda = 0.0015$ 为参数的指数分布, 即有密度函数

$f(x) = 0.0015e^{-0.0015x}, x > 0$. 令测试 6 个元件，并记录它们各自失效的时间(单位: h). 试问:

- (1) 至 800 小时时没有一个元件失效的概率是多少?
- (2) 到 3000 小时时所有元件都失效的概率是多少?

解: (1) $\xi_i \sim E(0.0015), i=1, 2, \dots, 6$, 且相互独立, 则

$$P\{\min(\xi_1, \xi_2, \dots, \xi_6) > 800\} = \prod_{i=1}^6 P\{\xi_i > 800\}$$

$$= \left(e^{-0.0015 \cdot 800}\right)^6 = e^{-7.2}.$$

(2)

$$P\{\max(\xi_1, \xi_2, \dots, \xi_6) < 3000\} = \prod_{i=1}^6 P\{\xi_i < 3000\}$$

$$= \left(1 - e^{-0.0015 \cdot 3000}\right)^6 = \left(1 - e^{-4.5}\right)^6.$$

10. 设总体 $\xi \sim N(20, 3)$, 今从中抽取容量为 10 和 15 的两个独立样本, 试问这两个样本的平均之差的绝对值大于 0.3 的概率是多少?

解: 记这两个样本均值分别为:

$$\begin{cases} \bar{\xi}_1 \sim N\left(\mu, \frac{\sigma^2}{n_1}\right) = N\left(20, \frac{3}{10}\right) \\ \bar{\xi}_2 \sim N\left(\mu, \frac{\sigma^2}{n_2}\right) = N\left(20, \frac{3}{15}\right) \end{cases} \Rightarrow \bar{\xi}_1 - \bar{\xi}_2 \sim N\left(0, \frac{1}{2}\right)$$

$$P\{|\bar{\xi}_1 - \bar{\xi}_2| > 0.3\} = P\left\{ \left| \frac{\bar{\xi}_1 - \bar{\xi}_2}{\sqrt{\frac{1}{2}}} \right| > \frac{0.3}{\sqrt{\frac{1}{2}}} = 0.3\sqrt{2} \right\} = 2 * (1 - \phi(0.3\sqrt{2}))$$

$$= 0.6714$$

11. 设总体 ξ 服从正态分布 $N(\mu, \sigma^2)$, ξ_1, ξ_2 为其样本, 试求样本极差的分布, 极大值与极小值的分布。

解: (1) 当样本容量 $n=2$ 时, 极差的分布即为 $|\xi_1 - \xi_2|$ 的分布, $\because \eta = \xi_1 - \xi_2 \sim N(0, 2\sigma^2)$

$$F_{|\eta|}(x) = P\{|\eta| < x\} = P\left\{ \left| \frac{\eta}{\sqrt{2}\sigma} \right| < \frac{x}{\sqrt{2}\sigma} \right\}$$

$$= P\{|\phi| < \frac{x}{\sqrt{2}\sigma}\} = 2\phi\left(\frac{x}{\sqrt{2}\sigma}\right) - 1$$

$$\therefore f_{|\eta|}(x) = 2 \cdot \frac{1}{\sqrt{2}\sigma} \phi\left(\frac{x}{\sqrt{2}\sigma}\right) = \frac{\sqrt{2}}{\sigma} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{x^2}{2\sigma^2}}$$

$$= \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{x^2}{4\sigma^2}}$$

由公式 (6.2.16) 有

$$\begin{aligned}
f_D(v) &= \int_{-\infty}^{\infty} f(v+y)f(v)dy \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{v+y-\sigma}{\sigma}\right)^2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{v-\sigma}{\sigma}\right)^2} dy \\
&= \frac{1}{2\pi\sigma^2}
\end{aligned}$$

(2) 由公式(6.2.12), (6.2.13), 得

$$\begin{aligned}
f_2(u) &= 2f(u)F(u) = 2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{u-\sigma}{\sigma}\right)^2} \int_{-\infty}^u \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\sigma}{\sigma}\right)^2} dx \\
&= 2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{u-\sigma}{\sigma}\right)^2} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\frac{u-\sigma}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \sigma dt \\
&= \frac{1}{\pi\sigma} e^{-\frac{1}{2}\left(\frac{u-\sigma}{\sigma}\right)^2} \int_{-\infty}^{\frac{u-\sigma}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt
\end{aligned}$$

(3) $f_1(u) = 2f(v)(1 - F(v))$

$$\begin{aligned}
f_1(v) &= 2 \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{v-\sigma}{\sigma}\right)^2} - 2f(v)F(v) \\
&= \frac{2}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{v-\sigma}{\sigma}\right)^2} - f_2(v) \text{(由 (2))}
\end{aligned}$$

12. 设总体 ξ 服从参数为 λ 的指数分布, ξ_1, ξ_2 为其样本, 试求样本的极大值、极小值与极差的分布。

解: 参见 P13 页例 6.2.5 当 $n=2$ 的情形:

$$f_2(u) = \begin{cases} 2\lambda e^{-\lambda u}(1 - e^{-\lambda u}), & 0 < u < \infty \\ 0, & \text{其他} \end{cases}$$

$$f_1(v) = \begin{cases} 2\lambda e^{-2\lambda v}, & 0 < v < \infty \\ 0, & \text{其他} \end{cases}$$

$$f_{D_2}(y) = \begin{cases} \lambda e^{-\lambda y}, & 0 < y < \infty \\ 0, & \text{其他} \end{cases}$$

13. 设 $\xi_1, \xi_2, \dots, \xi_n$ 是相互独立的且都是服从正态 $N(0,1)$ 分布的随机变量, $\xi_1, \xi_2, \dots, \xi_n$ 到 $\eta_1, \eta_2, \dots, \eta_n$ 的变换为正交变换, 试证: $\eta_1, \eta_2, \dots, \eta_n$ 是 n 个相互独立的且都服正态 $N(0,1)$ 分布的随机变量。

证明: 因为 $f_{\xi}(x_1, x_2, \dots, x_n) = (2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}x'x\right\}$, 其中 $x = (x_1, x_2, \dots, x_n)'$. 设正交变换为: $x = Ty$, 则其雅可比行列式 $\|T\|=1$, 且 $xx' = y'T'Ty = y'y$, 由随机向量函数的密度公式, 得

$$f_{\eta}(\gamma_1, \gamma_2, \dots, \gamma_n) = (2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}(\mathcal{I}\gamma)'(\mathcal{I}\gamma)\right\} |\mathcal{I}|$$

$$= (2\pi)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2}y'y\right\}$$

$$\therefore \eta \sim N_n(0, I_n)$$

其中, I_n 为单位矩阵。

14. 设总体 ξ 服从正态 $N(\mu, \sigma^2)$, $\xi_1, \xi_2, \dots, \xi_n$ 为其样本, $\bar{\xi}$ 与 S_n^2 分别为样本均值及方差。又设 ξ_{n+1} 服从正态 $N(\mu, \sigma^2)$, 且与 $\xi_1, \xi_2, \dots, \xi_n$ 相互独立, 试求统计量

$$\eta = \frac{\xi_{n+1} - \bar{\xi}}{S_n} \sqrt{\frac{n-1}{n+1}}$$

的抽样分布.

解:

$$\begin{aligned} \xi_{n+1} - \bar{\xi} &\sim N\left(0, \frac{n+1}{n}\sigma^2\right) \Rightarrow \frac{\xi_{n+1} - \bar{\xi}}{\sqrt{\frac{n+1}{n}\sigma^2}} \sim N(0, 1) \\ \text{又 } \because \frac{nS_n^2}{\sigma^2} &\sim \chi^2(n-1) \end{aligned} \quad \left. \right\} \Rightarrow$$

$$\frac{\xi_{n+1} - \bar{\xi}}{\sqrt{\frac{n+1}{n}\sigma^2}} = \frac{\xi_{n+1} - \bar{\xi}}{S_n} \sqrt{\frac{n-1}{n+1}} \sim t(n-1)$$

15. 设 $\xi_1, \xi_2, \dots, \xi_n$ 相互独立且服从正态分布 $N(\mu_i, \sigma_i^2), i=1, 2, \dots, n$, 试证明: $\eta = \sum_{i=1}^n c_i \xi_i$

$$N\left(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2\right)$$

服从正态分布

证明: 应用特征函数。 $\varphi_{\xi_i}(t) = e^{it\mu_i - \frac{1}{2}\sigma_i^2 t^2}, i=1, 2, \dots, n$.

$$\begin{aligned} \varphi_{\eta}(t) &= Ee^{\theta\eta t} = Ee^{i\sum_{k=1}^n c_k \xi_k t} = \prod_{k=1}^n Ee^{ic_k \xi_k t} = \prod_{k=1}^n e^{ic_k \mu_k t - \frac{1}{2}c_k^2 \sigma_k^2 t^2} \\ &= e^{i\sum_{k=1}^n c_k \mu_k t - \frac{1}{2}\sum_{k=1}^n c_k^2 \sigma_k^2 t^2} \Rightarrow \eta \sim N\left(\sum_{k=1}^n c_k \mu_k, \sum_{k=1}^n c_k^2 \sigma_k^2\right) \end{aligned}$$

16. 设总体在 $(\theta - \frac{1}{2}, \theta + \frac{1}{2})$ 上服从均匀分布, $\xi_1, \xi_2, \dots, \xi_n$ 为其样本, $\xi_{(1)}, \xi_{(2)}, \dots, \xi_{(n)}$ 为顺序统计量, 试求 $\xi_{(1)}, \xi_{(n)}$ 及 $(\xi_{(1)}, \xi_{(n)})$ 的分布。

解:

$$f_1(v) = nf(v)[1 - F(v)]^{n-1} \quad (\text{公式6.2.13})$$

$$= n[1 - (v - (\theta - \frac{1}{2}))]^{n-1} = n(\frac{1}{2} + \theta - v)$$

$$f_n(u) = nf(u)[F(u)]^{n-1} \quad (\text{公式6.2.12})$$

$$= n(u + \frac{1}{2} - \theta)^{n-1}$$

$$\begin{aligned} F(u, v) &= P\{\xi_{(1)} \leq u, \xi_{(n)} \leq v\} \\ &= 1 - P\{\xi_{(n)} > u\} - P\{\xi_{(1)} > v, \xi_{(n)} < u\} \\ &= 1 - P\{\xi_{(n)} > u\} - P(v < \xi_{(1)} \leq \xi_{(n)} < u) \\ &= 1 - P\{\xi_{(n)} > u\} - [F(u) - F(v)]^n \\ &= 1 - P\{\xi_{(n)} > u\} - (u - v)^n \\ \Rightarrow f(u, v) &= \frac{\partial^2}{\partial u \partial v} F(u, v) = n(n-1)(u-v)^{n-2} \end{aligned}$$

第七章 参数估计

课后习题详解：

1.解：

$$\xi \quad \Gamma(\alpha, \beta) = \Gamma(3, \frac{1}{9}), E\xi = \alpha\beta = \frac{1}{3}, D\xi = \alpha\beta^2 = \frac{1}{27}$$

总体 ξ 的c.f. $\varphi_0(t) = (1 - i\beta t)^{-\alpha}$

子样 $\bar{\xi}$ 的c.f. $\varphi_1(t) = (1 - i\beta \frac{t}{n})^{-n\alpha}$

记 $\eta = \frac{2n\bar{\xi}}{\beta}$, 则 η 的c.f. $\varphi_2(t) = (1 - 2it)^{-n\alpha}$

$$\therefore \eta \quad \Gamma(n\alpha, 2) = \chi^2(2n\alpha) (\chi^2(n) = \Gamma(\frac{n}{2}, 2))$$

$$\Rightarrow P\{\bar{\xi} > M\} = P\{\eta > \frac{2nM}{\beta}\} = 0.01$$

$$\frac{2nM}{\beta} = \chi^2_{2n\alpha}(0.001) = \chi^2_{30}(0.01) = 50.89$$

$$\Rightarrow M = 50.89 \cdot \frac{\beta}{2n} = 50.89 \cdot \frac{1}{10} \cdot \frac{1}{9} = 0.56544$$

2.解：

$$\eta(y) = e^{\alpha-\beta y}, \alpha > 0, \beta > 0, y \geq 0$$

震级 $\xi < m$ 的概率

$$P\{\xi < m\} = F(m) = \frac{\int_0^m \eta(y) dy}{\int_0^\infty \eta(y) dy} = \frac{\int_0^m e^{\alpha-\beta y} dy}{\int_0^\infty e^{\alpha-\beta y} dy} = 1 - e^{-\beta m}, m > 0$$

即 $\xi \sim E(\beta)$

3. 解:

$$\xi \sim N(\alpha, 1), P\{\xi < 0\} = \frac{14}{20} = 0.7 \text{ (用频率估计概率)}$$

$$\because \xi - \alpha \sim N(0, 1)$$

$$\therefore P(\xi < 0) = P(\xi - \alpha < -\alpha) = \phi(-\alpha)$$

$$= 1 - \phi(\alpha) = 0.7$$

$$\phi(\alpha) = 0.3 \Rightarrow \hat{\alpha} = -0.2544.$$

4. 解: 求 θ 的极大似然估计量

$$(1) f(x; \theta) = \frac{1}{2} e^{-|x-\theta|}, |x| > \theta, -\infty < \theta < 0, -\infty < x < \infty$$

$$\prod_{i=1}^n f(x_i; \theta) = \frac{1}{2^n} e^{-\sum_{i=1}^n |x_i - \theta|}$$

当 θ 取 x_1, x_2, \dots, x_n 的中位数时, $\sum_{i=1}^n |x_i - \theta|$ 取到最小值。

(2) θ 的似然函数为

$$L(x; \theta) = \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1}$$

对数似然方程为

$$\frac{\partial \ln L(x; \theta)}{\partial \theta} = \frac{\partial}{\partial \theta} (n \ln \theta + (\theta - 1) \sum_{i=1}^n \ln x_i)$$

$$= \frac{n}{\theta} + \sum_{i=1}^n \ln x_i = 0 \therefore \theta = \frac{-n}{\sum_{i=1}^n \ln x_i}$$

$$\therefore \hat{\theta} = \frac{-n}{\sum_{i=1}^n \ln \xi_i} = -\frac{1}{\bar{\xi}}$$

(3) θ 的似然函数为

$$L(x; \theta) = \frac{1}{\theta^n} \leq \frac{1}{x_{(n)}} (0 < x_{(n)} \leq \theta)$$

$$\therefore \hat{\theta}_{MLE} = \xi_{(n)}$$

又 $E\xi = \frac{\theta}{2}$, $\Rightarrow \hat{\theta}_1 = 2\bar{\xi}$ 是 θ 的矩法估计量 (不同于极大似然估计量)。

(4) θ 的似然函数为

$$L(x; \theta) = \frac{1}{\theta^n} e^{-\frac{\sum x_i}{\theta}}$$

对数似然方程为

$$\frac{\partial \ln L(x; \theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left(-n \ln \theta - \frac{\sum x_i}{\theta} \right)$$

$$= -\frac{n}{\theta} + \frac{\sum x_i}{\theta^2} = 0$$

$$\theta = \frac{\sum x_i}{n} \Rightarrow \hat{\theta}_{MLE} = \bar{x}.$$

5.解:

$$\begin{cases} \bar{\xi} = Np \\ S^2 = Npq \end{cases} \Rightarrow \begin{cases} p = 1 - \frac{S^2}{\bar{\xi}} = \frac{\bar{\xi} - S^2}{\bar{\xi}} \\ N = \frac{\bar{\xi}}{p} = \frac{\bar{\xi}^2}{\bar{\xi} - S^2} \end{cases} \Rightarrow$$

$$\begin{cases} p = \frac{\bar{\xi} - S^2}{\bar{\xi}} \\ N = \frac{\bar{\xi}^2}{\bar{\xi} - S^2} \end{cases}$$

6.解:

$$\eta = \ln \xi \sim N(\alpha, \sigma^2), \xi = e^\eta,$$

$$E\xi = Ee^\eta$$

7.解: (1)

β 的似然方程为

$$L(x; \beta) = \beta^n e^{-\beta \sum_{i=1}^n (t_i - t_0)}$$

对数似然方程为

$$\frac{\partial}{\partial \beta} \ln L(x; \beta) = \frac{\partial}{\partial \beta} \left(n \ln \beta - \beta \sum_{i=1}^n t_i + n \beta t_0 \right)$$

$$= \frac{n}{\beta} - \sum_{i=1}^n t_i + nt_0$$

$$\Rightarrow \beta = \frac{n}{\sum_{i=1}^n t_i - nt_0} \Rightarrow \beta = \frac{1}{T - t_0}$$

(2) t_0 的似然方程为:

$$L(x; t_0) = \beta^n e^{-\beta \sum_{i=1}^n (t_i - t_0)}$$

对数似然方程为

$$\begin{aligned}\frac{\partial}{\partial t_0} \ln L(x; t_0) &= \frac{\partial}{\partial t_0} (n \ln \beta - \beta \sum_{i=1}^n t_i + n \beta t_0) \\ &= n \beta \neq 0\end{aligned}$$

$\because t_0 < t_i, i = 1, 2, \dots, n$

$$\therefore \min_{t_i > t_0} \sum_{i=1}^n (t_i - t_0) \geq \sum_{i=1}^n (t_i - \min t_i)$$

$$\therefore t_0 = \min_{1 \leq i \leq n} t_i$$

8.解：

(1):

$$\begin{aligned}E\sigma &= \frac{1}{k} \sum_{i=1}^n E|\xi_i - a| = \frac{1}{k} \sum_{i=1}^n \sqrt{\frac{2}{\pi}} \sigma (ch6, Ex7) \\ &= \frac{n}{k} \sqrt{\frac{2}{\pi}} \sigma, \text{令 } \frac{n}{k} \sqrt{\frac{2}{\pi}} = 1 \Rightarrow k = n \sqrt{\frac{2}{\pi}}\end{aligned}$$

(2):

$$\begin{aligned}E\sigma^2 &= \frac{1}{k} \sum_{i=1}^{n-1} E(\xi_{i+1} - \xi_i)^2 \\ &= \frac{1}{k} \sum_{i=1}^{n-1} (E\xi_{i+1}^2 - 2E\xi_i\xi_{i+1} + E\xi_i^2) = \frac{n-1}{k} 2\sigma^2 \\ &\Rightarrow k = 2(n-1)\end{aligned}$$

9.解：

$$T = \sum_{i=1}^n c_i \xi_i \Rightarrow ET = \sum_{i=1}^n c_i E\xi_i = \sum_{i=1}^n c_i a = a \Rightarrow \sum_{i=1}^n c_i = 1$$

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i$$

$$\therefore \rho_{T, \bar{\xi}} = \frac{\text{cov}(\sum_{i=1}^n c_i \xi_i, \frac{\sum_{i=1}^n \xi_i}{n})}{\sqrt{DT D \bar{\xi}}} = \frac{\frac{1}{n} \sum_{i=1}^n c_i D \xi_i}{\sqrt{DT D \bar{\xi}}}$$

$$\text{令 } \frac{D \xi / n}{\sqrt{DT D \bar{\xi}}} = \frac{D \bar{\xi}}{\sqrt{DT D \bar{\xi}}} = \sqrt{\frac{D \bar{\xi}}{DT}}$$

10.解：

$E\alpha_1 = E\alpha_2 = E\alpha_3 = \alpha$, 但

$$D\alpha_1 = \frac{1}{25} + \frac{9}{100} + \frac{1}{4} = 0.38$$

$$D\alpha_2 = \frac{1}{9} + \frac{1}{16} + \frac{25}{144} = 0.3472$$

$$D\alpha_3 = \frac{1}{9} + \frac{1}{36} + \frac{1}{4} = 0.388$$

$\Rightarrow \alpha_2$ 的方差最小。

11. 解:

$$\rho = \frac{\text{cov}(\alpha_1, \alpha_2)}{\sigma_1 \sigma_2} \Rightarrow \text{cov}(\alpha_1, \alpha_2) = \rho \sigma_1 \sigma_2$$

$$\therefore D(c_1 \alpha_1 + c_2 \alpha_2) = c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 + 2c_1 c_2 \text{cov}(\alpha_1, \alpha_2)$$

$$= c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 + 2\rho c_1 c_2 \sigma_1 \sigma_2$$

$$= (\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2) c_1^2 + (2\rho \sigma_1 \sigma_2 - 2\sigma_2^2) c_1 + \sigma_2^2$$

$$\Rightarrow \text{当 } c_1 = -\frac{2\rho \sigma_1 \sigma_2 - 2\sigma_2^2}{2(\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2)}$$

$$= \frac{\sigma_2^2 - \rho \sigma_1 \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho \sigma_1 \sigma_2} \text{ 时, 取最小值.}$$

12. 解:

$$(1) \hat{\alpha} = \bar{\xi} - \bar{\eta}, E\hat{\alpha} = E(\bar{\xi} - \bar{\eta}) = E\bar{\xi} - E\bar{\eta} = \alpha_1 - \alpha_2$$

(2)

$$D\hat{\alpha} = D(\bar{\xi} - \bar{\eta}) = D\bar{\xi} + D\bar{\eta} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

$$= \frac{1}{n_1} + \frac{4}{n_2},$$

即 $z = f(x, y) = \frac{1}{x} + \frac{1}{y}$, 在 $x + y = n$ 条件下的极值,

由 **Langrange** 乘数法, 令

$$L(x, y, \lambda) = f(x, y) + \lambda(x + y - n) = \frac{1}{x} + \frac{4}{y} + \lambda(x + y - n)$$

$$\left. \begin{aligned} L_x &= -\frac{1}{x^2} + \lambda = 0 \\ L_y &= -\frac{4}{y^2} + \lambda = 0 \\ L_\lambda &= x + y - n = 0 \end{aligned} \right\} \Rightarrow x = \frac{n}{3}$$

故 $n_1 = \frac{n}{3}$, $n_2 = n - \frac{n}{3}$ 时, 可使 $\hat{\alpha}$ 的方差达到最小.

13. 解:

$\because E(\xi - \eta) = \alpha_1 - \alpha_2, D(\xi - \eta) = \sigma_1^2 + \sigma_2^2$, 由 P48 的推论知, $\overline{\xi - \eta} = \bar{\xi} - \bar{\eta}$ 为 $\alpha_1 - \alpha_2$ 的最小线性无偏估计。

14.解:

$$P\{\xi_n^* < x\} = F^n(x) = \frac{x^n}{\theta^n} = \frac{x^3}{\theta^3}$$

$$f_{\xi_n^*}(x) = \frac{3x^2}{\theta^3}, E\xi_n^* = \frac{\int_0^\theta 3x^3 dx}{\theta^3} = \frac{3}{4}\theta$$

$\therefore \frac{4}{3}\xi_n^*$ 为 θ 的无偏估计.

$$E\left(\frac{4}{3}\xi_n^*\right)^2 = \frac{16}{9} \int_0^\theta x^2 \frac{3x^2}{\theta^3} dx = \frac{16}{15}\theta^2$$

$$D\left(\frac{4}{3}\xi_n^*\right) = E\left(\frac{4}{3}\xi_n^*\right)^2 - E^2\left(\frac{4}{3}\xi_n^*\right)$$

$$= \frac{16}{15}\theta^2 - \theta^2 = \frac{1}{15}\theta^2.$$

$$\forall P\{\xi_1^* < x\} = 1 - (1 - \frac{x}{\theta})^3 = 1 - \frac{(\theta - x)^3}{\theta^3}$$

$$f_{\xi_1^*}(x) = \frac{3(\theta - x)^2}{\theta^3}$$

$$E\xi_1^* = \int_0^\theta \frac{3(\theta - x)^2}{\theta^3} dx = \frac{\theta}{4}, \therefore 4\xi_1^* \text{ 为 } \theta \text{ 的无偏估计.}$$

$$E(4\xi_1^*)^2 = 16E\xi_1^{*2} = 16 \int_0^\theta x^2 \frac{3(\theta - x)^2}{\theta^3} dx = \frac{8}{5}\theta^2$$

$$\therefore D(4\xi_1^*) = E(4\xi_1^*)^2 - E^2(4\xi_1^*) = \frac{8}{5}\theta^2 - \theta^2 = \frac{3}{5}\theta^2$$

$$D\left(\frac{4}{3}\xi_n^*\right) = \frac{1}{15}\theta^2 < D(4\xi_1^*) = \frac{3}{5}\theta^2$$

$\therefore \frac{4}{3}\xi_n^*$ 较为有效.

15.解:

$$E\theta_1 = E\theta_2 = \theta, D\theta_1 = \sigma_1^2 = 2D\theta_2 = 2\sigma_2^2$$

$$E[c_1\theta_1 + c_2\theta_2] = (c_1 + c_2)\theta = \theta \Rightarrow c_1 + c_2 = 1$$

$$\text{且 } D[c_1\theta_1 + c_2\theta_2] = c_1^2 D\theta_1 + c_2^2 D\theta_2 = c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2$$

$$= 2c_1^2 \sigma_1^2 + c_2^2 \sigma_2^2 = (2c_1^2 + c_2^2) \sigma_2^2$$

$$\text{且 } 2c_1^2 + c_2^2 = 2c_1^2 + (1 - c_1)^2 = 2c_1^2 + 1 - 2c_1 + c_1^2$$

$$= 3c_1^2 - 2c_1 + 1$$

当 $c_1 = \frac{1}{3}$ 时, $c_1\theta_1 + c_2\theta_2 = \frac{1}{3}\theta_1 + \frac{2}{3}\theta_2$, $D[c_1\theta_1 + c_2\theta_2]$ 取最小值.

16. 证:

$\forall I_0 \in U_0$, $\because I_1, I_2$ 分别是 $g_1(\theta), g_2(\theta)$ 的最优无偏估计量, $\therefore EI_1I_0 = EI_2I_0 = 0$ 。

又因为 $E(b_1I_1 + b_2I_2) = b_1g_1(\theta) + b_2g_2(\theta)$, $D(b_1I_1 + b_2I_2) = b_1^2 DI_1 + b_2^2 DI_2 < \infty$

$$\therefore b_1I_1 + b_2I_2 \in U$$

$$\text{且 } EI_0(b_1I_1 + b_2I_2) = Eb_1I_1I_0 + Eb_2I_2I_0 = 0$$

$\therefore b_1I_1 + b_2I_2$ 是 $b_1g_1(\theta) + b_2g_2(\theta)$ 的最优无偏估计量。

17. 解:

$$\frac{(n_1-1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1-1), \frac{(n_2-1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2-1)$$

且 $\frac{(n_1-1)S_1^2}{\sigma_1^2}, \frac{(n_2-1)S_2^2}{\sigma_2^2}$ 独立, 由 F 分布的定义, 有

$$(1) \quad F = \frac{(n_1-1)S_1^2}{\sigma_1^2(n_1-1)} / \frac{(n_2-1)S_2^2}{\sigma_2^2(n_2-1)} = \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} \sim F(n_1-1, n_2-1)$$

则由

$$1 - \alpha = P\{k_1 < F < k_2\} = P\left\{k_1 < \frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2} < k_2\right\}$$

$$= P\left\{\frac{S_1^2}{k_2 S_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{k_1 S_2^2}\right\}$$

通常取 $k_1 = F_{1-\alpha/2}(n_1-1, n_2-1)$, $k_2 = F_{\alpha/2}(n_1-1, n_2-1)$,

即 $\frac{\sigma_1^2}{\sigma_2^2}$ 的 $1 - \alpha$ 置信区间为 $(\frac{S_1^2}{k_2 S_2^2}, \frac{S_1^2}{k_1 S_2^2})$.

(2).

$$\because \bar{\xi} - \bar{\eta} \sim N(a_1 - a_2, \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2})$$

$$\therefore \frac{\bar{\xi} - \bar{\eta} - (a_1 - a_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \sim N(0, 1)$$

$\Rightarrow a_1 - a_2$ 的置信区间为

$$[\bar{\xi} - \bar{\eta} \pm u_{\alpha/2} \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}]$$

18. 解:

$\because \sigma_1^2 = \sigma_2^2 = \sigma^2$, 但 σ^2 未知, 则

$$T = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \cdot \frac{(\bar{\xi} - \bar{\eta}) - (\alpha_1 - \alpha_2)}{\sqrt{n_1 S_1^2 + n_2 S_2^2}} \sim t(n_1 + n_2 - 2)$$

$\Rightarrow \alpha_1 - \alpha_2$ 的 $1 - \alpha$ 置信区间为

$$[\bar{\xi} - \bar{\eta} \pm t_{\alpha/2} \frac{(n_1 + n_2 - 2) \sqrt{n_1 S_1^2 + n_2 S_2^2}}{\sqrt{n_1 + n_2}}, \sqrt{\frac{n_1 + n_2}{n_1 n_2 (n_1 + n_2 - 2)}}]$$

19. 解:

$$\bar{\xi} = \frac{\sum_{i=1}^{15} x_i}{15} = \frac{8.7}{15} = 0.58, S_n^2 = \sum_{i=1}^{15} x_i^2 - \bar{\xi}^2 = 24.7136,$$

α 的置信水平为 $1 - \alpha = 0.95$ 的区间估计为

$$\begin{aligned} & [\bar{\xi} - t_{\alpha/2}(n-1) \frac{S_n}{\sqrt{n-1}}, \bar{\xi} + t_{\alpha/2}(n-1) \frac{S_n}{\sqrt{n-1}}] \\ & = [0.58 - 22 * \frac{\sqrt{24.7136}}{\sqrt{14}}, 0.58 + 22 * \frac{\sqrt{24.7136}}{\sqrt{14}}] \\ & = [-28.6499, 29.8099] \end{aligned}$$

σ^2 的置信水平为 $1 - \alpha = 0.95$ 的区间估计为

$$\begin{aligned} & \left[\frac{n S_n^2}{\chi_{\alpha/2}^2(n-1)}, \frac{n S_n^2}{\chi_{1-\alpha/2}^2(n-1)} \right] \\ & = \left[\frac{15 * 24.7136}{26.1189}, \frac{15 * 24.7136}{5.6287} \right] \\ & = [14.1929, 65.8596] \end{aligned}$$

20. 解: n=10, $\alpha = 0.1$

$$\hat{\alpha} = \bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i = 2, \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (\xi_i - \bar{\xi})^2 = S^2 = 5.778 \Rightarrow S = 2.4037.$$

$$\alpha = 1 - 0.90 = 0.1, t_{\alpha/2}(10-1) = t_{0.05}(9) = 1.833$$

$$\begin{aligned} & \alpha \text{ 的区间估计} [\bar{\xi} - t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}, \bar{\xi} + t_{\alpha/2}(n-1) \frac{S}{\sqrt{n}}] \\ & = [2 - 1.833 \cdot \frac{2.4037}{\sqrt{10}}, 2 + 1.833 \cdot \frac{2.4037}{\sqrt{10}}] = [0.6067, 3.3933]. \end{aligned}$$

$$\chi_{\alpha/2}^2(n-1) = \chi_{0.05}^2(9) = 16.919$$

$$\chi_{1-\alpha/2}^2(n-1) = \chi_{0.95}^2(9) = 3.325.$$

$$\begin{aligned} & \sigma^2 \text{ 的区间估计} \left[\frac{(n-1)S^2}{\chi_{\alpha/2}^2(n-1)}, \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2(n-1)} \right] = \left[\frac{9 * 5.778}{16.919}, \frac{9 * 5.778}{3.325} \right] \\ & = [3.0735, 15.6391] \end{aligned}$$

第八章 假设检验 课后习题参考答案

1. 解一： 这是成对数据下均值差的假设检验， $\because X \sim N(\mu_1, \sigma^2), Y \sim N(\mu_2, \sigma^2)$ 分别有样本 $x = [0.19, 0.16, 0.21, 0.30, 0.66, 0.42, 0.08, 0.12, 0.30, 0.27]$, $y = [0.19, 0.24, 1.04, 0.08, 0.20, 0.12, 0.31, 0.29, 0.13, 0.07]$; 令 $Z = X - Y, \Rightarrow Z \sim N(\mu_1 - \mu_2, 2\sigma^2)$, 令 $z_i = x_i - y_i, i = 1, 2, \dots, 10$. 因此问题归结为检验假设 $H_0: \mu = \mathbf{0}$, $H_1: \mu \neq \mathbf{0}$ 单个正态总体均值(方差未知)的假设检验。检验的拒绝域为

$$\begin{aligned} W &= \{(z_1, \dots, z_n) : \left| \frac{\bar{z} - \mu_0}{s / \sqrt{n}} \right| > t_{\alpha/2}(n-1)\} \\ &= \{(z_1, \dots, z_n) : \left| \frac{\bar{z}}{s / \sqrt{10}} \right| > t_{0.025}(9)\} (H_0 \text{ 成立}) \end{aligned}$$

计算得： z 的均值 $\bar{z} = 0.006$, z 的样本标准差为 $S = 0.3641$. 又 $n=10$, 所以

$$U = \frac{\bar{z}}{s / \sqrt{10}} = \frac{0.006}{0.1325} \cdot \sqrt{10} = 0.0521.$$

又计算可得 $t_{\frac{\alpha}{2}}(n-1) = t_{0.025}(9) = 2.2622$.

由于 $0.0521 < 2.2622$, 则不能否定原假设, 即可认为处理前后含脂率的平均值无显著变化。

解二： 用 t 检验法. 设 $H_0: \mu_1 = \mu_2 \Leftrightarrow H_1: \mu_1 \neq \mu_2$, 拒绝域为

$$\begin{aligned} W &= \left\{ \left| \frac{(\bar{x} - \bar{y})}{s_{\sqrt{1/n+1/m}}} \right| > t_{\alpha/2}(n+m-2) = t_{0.025}(18) \right\}, \quad t_{0.025}(18) = 2.10 \\ S_1^2 &= 0.0281, S_2^2 = 0.0806 \\ S_p^2 &= \sqrt{\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}} = \sqrt{\frac{9S_1^2 + 9S_2^2}{18}} = 0.2331 \end{aligned}$$

所以统计量的值

$$t_0 = \frac{(\bar{x} - \bar{y})}{s_{\sqrt{1/n+1/m}}} = \frac{\bar{z}}{0.2331 \sqrt{\frac{1}{10} + \frac{1}{10}}} = \frac{0.006}{0.2331} \sqrt{5} = 0.0576$$

因为 $0.0576 < 2.10$, 所以不能拒绝原假设, 即认为处理前后含脂率的平均值无显著变化。

解三： 使用 Matlab 统计工具箱求解: 使用函数函数 `ttest2()`

调用格式:

`[h,sig,ci]=ttest2(X,Y,alpha,tail)` % sig 为当原假设为真时得到观察值的概率, 当 sig 为

小概率时则对原假设提出质疑，ci 为真正均值 μ 的 1-alpha 置信区间。

说明 若 h=0, 表示在显著性水平 alpha 下, 不能拒绝原假设;

若 h=1, 表示在显著性水平 alpha 下, 可以拒绝原假设。

原假设: $H_0: \mu_1 = \mu_2$, (μ_1 为 X 为期望值, μ_2 为 Y 的期望值)

若 tail=0, 表示备择假设: $H_1: \mu_1 \neq \mu_2$ (默认, 双边检验);

tail=1, 表示备择假设: $H_1: \mu_1 > \mu_2$ (单边检验);

tail=-1, 表示备择假设: $H_1: \mu_1 < \mu_2$ (单边检验)。

程序代码:

```
x=[0.19 0.18 0.21 0.30 0.66 0.42 0.08 0.12 0.30 0.27];  
y=[0.19 0.24 1.04 0.08 0.20 0.12 0.31 0.29 0.13 0.07];  
[h,sig,ci]=ttest2(x,y,0.05)
```

执行结果:

h = 0% 不否定原假设。

Sig=0.9547;% p 值太大, 不能否定原假设。

ci = [-0.2130 0.2250] % 真正均值 0.95 的置信区间。

2. 解: 属于两个正态总体方差的假设检验, 用 F 检验法。

$H_0: \sigma_1^2 = \sigma_2^2 \Leftrightarrow H_1: \sigma_1^2 \neq \sigma_2^2$, 由于当 H0 成立时,

$$\frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} = \frac{S_1^2}{S_2^2} \sim F(n-1, m-1) = F(5, 8); \text{ 拒绝域的形式为}$$

$$\begin{aligned} W &= \{(x_1, \dots, x_n) : s_1^2 / s_2^2 < F_{1-\alpha/2}(n-1, m-1) \\ &\quad \text{or } s_1^2 / s_2^2 > F_{\alpha/2}(n-1, m-1)\} \\ &= \{(x_1, \dots, x_n) : s_1^2 / s_2^2 < F_{0.975}(5, 8) = 0.1480 \\ &\quad \text{or } s_1^2 / s_2^2 > F_{0.025}(5, 8) = 4.8173\} \end{aligned}$$

$$\text{又 } F = \frac{S_1^2}{S_2^2} = \frac{0.3450}{0.3570} = 0.9664 \notin W, \text{ 既不在拒绝域之中, 故可认为方差无显著性差异。}$$

3. 解: H0: 电话呼叫次数服从泊松分布。

设 i 为呼叫次数, v_i 为对应的频数, $i=0, 1, \dots, 7$, $m=8$, 则总观察次数 $n = \sum_{i=1}^7 v_i = 60$

参数 λ (每分钟的平均呼叫次数) 的极大似然估计为

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^7 i v_i = 2, \text{ 从而可计算理论频数值}$$

$$p_k = \frac{\hat{\lambda}^k e^{-\hat{\lambda}}}{k!}, k = 0, 1, \dots, 7,$$

默认情形下给定显著水平为 $\alpha = 0.05$, 得卡方分布的上侧 0.05 分位数

$\chi^2_{0.05}(m-1-1) = \chi^2_{0.05}(6) = 12.5916$, 又在给定样本的情形下, 皮尔逊 χ^2 检验统计量的值

$$\chi^2(6) = \sum_{i=0}^7 \frac{v_i^2}{np_i} - n = 0.5595 < 12.5916, \text{故不能拒绝原假设,此分布可看作为泊松分布。}$$

4. 解: H0: 记录的汽车辆数服从poisson分布。

$$v_i = [92 \ 68 \ 28 \ 11 \ 1 \ 0], m=6, n=200, \alpha = 0.05, \chi^2_{0.05}(6-1-1) = \chi^2_{0.05}(4) = 9.4877$$

λ 的极大似然估计为

$$\lambda = \frac{1}{n} \sum_{i=1}^m nv_i = 0.8050, \text{计算理论频数得}$$

$$p = n^\lambda v = (0.447 \ 0.359 \ 0.1448 \ 0.03887 \ 0.00782 \ 0.001267)$$

从而皮尔逊 χ^2 检验统计量的值

$$\chi^2(6) = \sum_{i=0}^5 \frac{v_i^2}{np_i} - n = 2.1596 < 9.4877$$

故不能拒绝原假设,此分布可看作为泊松分布。

5. 解: H0: 假设此分布服从均匀分布, n=800, m=10,

$$\text{实际频数 } V = (74 \ 92 \ 83 \ 79 \ 80 \ 73 \ 77 \ 75 \ 76 \ 91)$$

$$\text{理论频数 } np = (80 \ 80 \ 80 \ 80 \ 80 \ 80 \ 80 \ 80 \ 80 \ 80)$$

$$\chi^2_{0.05}(10-1) = \chi^2_{0.05}(9) = 16.919$$

根据样本计算皮尔逊 χ^2 检验统计量的值

$$\chi^2(9) = \sum_{i=0}^9 \frac{v_i^2}{np_i} - n = 5.125 < 16.919$$

故拒绝原假设,认为此分布不服从均匀分布。

6. 解: 设 ξ 为每次抽取10个产品中的次品数, H0: $\xi \sim b(10, p)$.

$$P \text{ 的极大似然估计为: } p = \frac{\sum_{i=0}^{10} iv_i}{10 \times 100} = 0.1, \text{ 可计算得理论频数为}$$

$$P = (0.3487 \ 0.3874 \ 0.1937 \ 0.0574 \ 0.0112 \ 0.0015 \ 0.0001 \ 0.0000 \\ 0.0000 \ 0.0000 \ 0.0000).$$

$$\chi^2_{0.05}(9) = 16.919$$

根据样本计算皮尔逊 χ^2 检验统计量的值

$$\chi^2(11-1-1) = \chi^2(9) = \sum_{i=0}^{10} \frac{v_i^2}{np_i} - n = 5.1295 < 16.919$$

故不能拒绝原假设，认为此分布是服从二项分布。

7. 解：这是成对数据下正态总体均值差的假设检验，并且具有共同的方差（未知），设 $X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, \sigma^2)$ $\alpha = 0.05$, $t_{\frac{\alpha}{2}}(n-1) = t_{0.025}(5) = 2.5706$ 令

$$Z = X - Y \Rightarrow Z \sim N(\mu_1 - \mu_2, 2\sigma^2) = N(\mu, 2\sigma^2), \text{ 检出原假设: } H_0: \mu_1 = \mu_2, \text{ 即 } \mu = 0.$$

转化为单个正态总体方差未知情形下均值是否为零的假设检验，我们可采用t检验法，建立统计量：

$$t = \frac{\bar{z} - \mu_0}{s / \sqrt{n}} = t(n-1) = t(6-1) = t(5)$$

$$N=6, \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = -0.1667, s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (z_i - \bar{z})^2} = 5.8793.$$

在 H_0 成立时，t的统计量的值为

$$t(5) = \frac{\bar{z} - \mu_0}{s / \sqrt{n}} = \frac{-0.1667}{5.8793 / \sqrt{6}} = -0.0694. |t(5)| < 2.5706, \text{ 故不能拒绝原假设.}$$

另解：采用两个正态总体方差相等时的均值差的t检验法. 检验统计量：

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{1/n + 1/m}} \sim t(2n-2) = t(10),$$

$$t_{\frac{\alpha}{2}}(n+m-2) = t_{0.025}(10) = 2.2281$$

故不能拒绝原假设，认为此分布是服从二项分布。

7. 解：这是成对数据下正态总体均值差的假设检验，并且具有共同的方差（未知），设 $X \sim N(\mu_1, \sigma^2)$, $Y \sim N(\mu_2, \sigma^2)$ $\alpha = 0.05$, $t_{\frac{\alpha}{2}}(n-1) = t_{0.025}(5) = 2.5706$ 令

$$Z = X - Y \Rightarrow Z \sim N(\mu_1 - \mu_2, 2\sigma^2) = N(\mu, 2\sigma^2), \text{ 检出原假设: } H_0: \mu_1 = \mu_2, \text{ 即 } \mu = 0.$$

转化为单个正态总体方差未知情形下均值是否为零的假设检验，我们可采用t检验法，建立统计量：

$$t = \frac{\bar{z} - \mu_0}{s / \sqrt{n}} = t(n-1) = t(6-1) = t(5)$$

$$N=6, \bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = -0.1667, s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (z_i - \bar{z})^2} = 5.8793.$$

在 H_0 成立时，t的统计量的值为

$$t(5) = \frac{\bar{z} - \mu_0}{s / \sqrt{n}} = \frac{-0.1667}{5.8793 / \sqrt{6}} = -0.0694. |t(5)| < 2.5706, \text{ 故不能拒绝原假设.}$$

另解：采用两个正态总体方差相等时的均值差的t检验法. 检验统计量：

$$t = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w \sqrt{1/n + 1/m}} \sim t(2n-2) = t(10),$$

$$t_{\frac{\alpha}{2}}(n+m-2) = t_{0.025}(10) = 2.2281$$

$$S_w = \sqrt{\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}} = 3.0469$$

$$H_0: \mu_1 = \mu_2 \Leftrightarrow H_1: \mu_1 \neq \mu_2$$

当 H_0 成立时，统计量的值

$$t(10) = \frac{(\bar{X} - \bar{Y})}{S_w \sqrt{1/n + 1/m}} = \frac{\bar{z}}{S_w \sqrt{2/n}} = \frac{-0.1667}{3.0469 \sqrt{2/6}} = -0.0947$$

因为 $|t(10)| = 0.0947 < t_{0.025}(10) = 2.2281$

故不能拒绝原假设。

8. 解： $H_0: P_{ij} = p_i \cdot p_j, i, j = 1, 2, n=385$.

$$\begin{aligned} (\bar{P}_{ij}) &= (\bar{P}_i \bar{P}_{j,i}) = \left(\frac{n_i}{n}, \frac{n_j}{n} \right) \\ &= \begin{pmatrix} 0.0793 & 0.3825 & 0.0836 \\ 0.0661 & 0.3188 & 0.0697 \end{pmatrix} \end{aligned}$$

\hat{P}_v 是在 H_0 成立时， p_v 的极大似然估计，此时相应统计量的值为

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(n_{ij} - np_{ij})^2}{np_{ij}} = 3.0762 < \chi^2_{0.95}(2) = 5.9915$$

故不能拒绝原假设，即认为慢性气管炎与吸烟量是相关的。

9. 解：当 $a_1 > a_0$ 时，设 $\xi_1, \xi_2, \dots, \xi_n$ 为总体 ξ 的样本，由于似然比

$$\begin{aligned} \frac{L(a_1)}{L(a_0)} &= \exp\left\{-\frac{1}{2\sigma^2}[-2n\bar{\xi}(a_1 - a_0) + n(a_1^2 - a_0^2)]\right\} \\ &= \exp\left\{\frac{\sqrt{n}(a_1 - a_0)}{\sigma}\left[\frac{\sqrt{n}(\bar{\xi} - a_0)}{\sigma} - \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right]\right\} \end{aligned}$$

所以由奈曼—皮尔逊基本引理，知 H_0 的最佳否定域为

$$\begin{aligned} X_0 &= \left\{(\xi_1, \xi_2, \dots, \xi_n) : \frac{L(a_1)}{L(a_0)} \geq k\right\} \\ &= \left\{(\xi_1, \xi_2, \dots, \xi_n) : \frac{\sqrt{n}(\bar{\xi} - a_0)}{\sigma} \geq \frac{\sigma \ln k}{\sqrt{n}(a_1 - a_0)} + \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right\} \end{aligned}$$

从而犯第一类错误的概率为

$$\alpha = P\{X_0 | H_0\} = P_{a_0}\left\{\frac{\sqrt{n}(\bar{\xi} - a_0)}{\sigma} \geq \frac{\sigma \ln k}{\sqrt{n}(a_1 - a_0)} + \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right\}$$

$$= 1 - \phi\left(\frac{\sigma \ln k}{\sqrt{n}(a_1 - a_0)} + \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right) \rightarrow 0(n \rightarrow \infty)$$

犯第二类错误的概率为

$$\begin{aligned}\beta &= P\{\bar{X}_0 \mid H1\} = P_{a_1}\left\{\frac{\sqrt{n}(\bar{\xi} - a_0)}{\sigma} < \frac{\sigma \ln k}{\sqrt{n}(a_1 - a_0)} + \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right\} \\ &= P_{a_1}\left\{\frac{\sqrt{n}(\bar{\xi} - a_1)}{\sigma} < \frac{\sigma \ln k}{\sqrt{n}(a_1 - a_0)} - \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right\} \\ &= \phi\left(\frac{\sigma \ln k}{\sqrt{n}(a_1 - a_0)} - \frac{\sqrt{n}(a_1 - a_0)}{2\sigma}\right) \rightarrow 0(n \leftrightarrow \infty)\end{aligned}$$

10. 解: $n = 16, \sigma = 2$. 属于单个正态总体均值的假设检验,

$$H0: \alpha = 0, H1: \alpha \neq 0$$

构造统计量

$$U = \frac{\bar{\xi} - \alpha}{\sigma / \sqrt{n}} \sim N(0, 1), \text{ 当 } H0 \text{ 成立时}$$

$$U = \frac{\bar{\xi} - \theta}{\sigma / \sqrt{n}} = \frac{\sqrt{n}\bar{\xi}}{\sigma} = \frac{\sqrt{16}\bar{\xi}}{2} = 2\bar{\xi} \sim N(0, 1)$$

$$1. \alpha_1 = P\{2\bar{\xi} \leq -1.645\} = \phi(-1.645) = 0.05$$

$$\begin{aligned} 2. \alpha_2 &= P\{1.50 \leq 2\bar{\xi} \leq 2.125\} = \phi(2.125) - \phi(1.50) \\ &= 0.9832 - 0.9332 = 0.05 \end{aligned}$$

$$\begin{aligned} 3. \alpha_3 &= P\{2\bar{\xi} \leq -1.96\} + P\{2\bar{\xi} \geq 1.96\} \\ &= 0.025 + 0.025 = 0.05 = 0.05 \end{aligned}$$

$$\therefore \alpha_1 = \alpha_2 = \alpha_3$$

即它们有相同的显著水平0.05.

1.1. 解：由于似然比

$$\frac{L(\theta_1)}{L(\theta_0)} = \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{(\frac{1}{\theta_0} - \frac{1}{\theta_1})\bar{x}\right\}, \text{ 由奈曼-皮尔逊基本引理知}$$

(1) H0的最佳否定域为

$$\begin{aligned} X_0 &= \left\{ (x_1, x_2, \dots, x_n) : \frac{L(\theta_1)}{L(\theta_0)} \geq k \right\} = \left\{ (x_1, x_2, \dots, x_n) : \left(\frac{\theta_0}{\theta_1}\right)^n \exp\left\{(\frac{1}{\theta_0} - \frac{1}{\theta_1})\bar{x}\right\} \geq k \right\} \\ &= \left\{ (x_1, x_2, \dots, x_n) : \left(\frac{1}{2}\right)^n \exp\left\{\frac{n}{4}\bar{x}\right\} \geq k \right\} = \left\{ (x_1, x_2, \dots, x_n) : \sum_{i=1}^n x_i \geq c_1 \right\}, c_1 = 4 \ln(2^n k) \end{aligned}$$

因为H0成立时，

$$\begin{aligned} \zeta &= \sum_{i=1}^n \xi_i \sim \Gamma(n, \frac{1}{2}) = \chi^2(2n) \\ \therefore c_1 &= \chi_a^2(2n) \end{aligned}$$

故最佳否定域为

$$X_0 = \left\{ (x_1, x_2, \dots, x_n) : \sum_{i=1}^n x_i \geq \chi_a^2(2n) \right\}$$

(2) H0的最佳否定域为

$$\left[\frac{L(\theta)}{L(\theta_0)} = 2^n e^{\frac{1}{2}n\bar{\xi}} \geq 2^n e^{\frac{1}{2}k} \Leftrightarrow n\bar{\xi} \leq k \right]$$

$$X_0 = \left\{ (x_1, x_2, \dots, x_n) : \frac{L(\theta)}{L(\theta_0)} \geq k \right\} = \left\{ (x_1, x_2, \dots, x_n) : \sum_{i=1}^n x_i \leq -2 \ln \left(\frac{k}{2^n} \right) \right\}$$

记 $\zeta = \sum_{i=1}^n x_i$, $c_2 = -2 \ln \left(\frac{k}{2^n} \right)$, 当 H_0 成立时, $\zeta \sim \chi^2(2n)$, 当 $\alpha = 0.05$ 时

$$X_0 = \left\{ (x_1, x_2, \dots, x_n) : \sum_{i=1}^n x_i \leq -\chi^2_{0.95}(2n) \right\}$$

1.2. 解:

(1) $p_0 = \frac{1}{2}$, $p_1 = \frac{1}{3}$, $\frac{L(p_1)}{L(p_0)} = \left[\frac{4}{3} \right]^n \left[\frac{1}{2} \right]^{n-k}$. 当 H_0 成立时, $n\bar{\xi} \sim b(n, \frac{1}{2})$. 设 $k = \left[\frac{4}{3} \right]^n \left[\frac{1}{2} \right]^c$, 则由 $\frac{L(p_1)}{L(p_0)} \geq k$ 得 $n\bar{\xi} \leq C$. 对给定的 $\alpha = 0.05$ 和 n , 如果存在非负整数 C_* 使得

$$0.05 = P_{\frac{1}{2}}(n\bar{\xi} \leq C_*) = \sum_{k=0}^{C_*} C_*^k \left[\frac{1}{2} \right]^k \left[\frac{1}{2} \right]^{n-k} = \sum_{k=0}^{C_*} C_*^k \left[\frac{1}{2} \right]^n \quad (\star)$$

成立, 则 H_0 的最佳否定域为

$$E_0 = \left\{ (x_1, \dots, x_n) : \sum_{i=1}^n x_i \leq C_* \right\}.$$

从而最佳检验为

$$\phi(X) = \begin{cases} 1, & \sum_{i=1}^n x_i \leq C_*, \\ 0, & \sum_{i=1}^n x_i > C_*. \end{cases}$$

但是, 对于 $\alpha = 0.05$ 和确定的 n , 可能找不到非负整数 C_* 使得 (\star) 式成立, 这时所得的最佳检验将是随机化检验. 例如, 当 $n=10$ 时,

$$\begin{aligned} \alpha_1 &\triangleq \sum_{k=0}^1 C_{10}^k \left[\frac{1}{2} \right]^{10} = 0.0107 < \alpha = 0.05 \\ &< \sum_{k=0}^2 C_{10}^k \left[\frac{1}{2} \right]^{10} = 0.0547. \end{aligned}$$

记 $\delta = (\alpha - \alpha_1) / P_{\frac{1}{2}}\{n\bar{\xi} = 2\} = 0.8932$, 于是得最佳检验

$$\phi(X) = \begin{cases} 1, & \sum_{i=1}^{10} x_i < 2 \\ \delta = 0.8932, & \sum_{i=1}^{10} x_i > 2 \\ 0, & \sum_{i=1}^{10} x_i = 2 \end{cases}$$

(2) 当 $p_0 = \frac{1}{3}, p_1 = \frac{1}{2}, \frac{L(p_1)}{L(p_0)} = \left(\frac{3}{4}\right)^2 2^2,$

记 $k = \left[\frac{3}{4}\right]^n 2^c$, 则由 $\frac{L(p_1)}{L(p_0)} \geq k$ 得

$n\bar{\xi} \geq c$, 当且仅当成立时

$n\bar{\xi} \sim b\left(n, \frac{1}{3}\right)$. 对于给定的 $\alpha = 0.05$ 和 n 如果存在非负整数

C_α , 使得: $0.05 = P_1(n\bar{\xi} \geq C_\alpha) = \sum_{k=C_\alpha}^n C_n^k \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k}$

C_α , 使得: $0.05 = P_1(n\bar{\xi} \leq C_\alpha) = \sum_{k=C_\alpha}^n C_n^k \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k}$

则所求最佳否定域为

$$\mathcal{L}_0 = \{(x_1, \dots, x_n) : \sum_{i=1}^n x_i \geq C_\alpha\}.$$

但是对于给定的 α 和 n , 可能找不到非负整数 C_α 使上式成立, 类似于 (1), 这时要求的最佳检验将是随机化检验.

1.3. 解:

由于似然比为

$$\frac{L(\lambda_1)}{L(\lambda_0)} = \prod_{i=1}^{10} \left[e^{-1} \frac{1}{\xi_i!} \right] / \prod_{i=1}^{10} \left[e^{-0.1} \frac{(0.1)^{\xi_i}}{\xi_i!} \right] = e^{-9} 10^{\xi_{10}}$$

其中, $\xi_{10} = \sum_{i=1}^{10} \xi_i$. 设 $k = e^{-\alpha} 10^C$, 则由 $\frac{L(\lambda_0)}{L(\lambda_k)} \geq k$ 得 $\xi_{10} \geq C$. 当 H_0 成立时, $\xi_{10} \sim P(10\lambda_0) = P(1)$. 所以由给定的 $\alpha = 0.05$.

$$P_{0.1}\{\xi_{10} \geq 3\} = 1 - P_{0.1}\{\xi_{10} \leq 2\} = 0.0803 > \alpha = 0.05,$$

$$P_{0.1}\{\xi_{10} \geq 4\} = 1 - P_{0.1}\{\xi_{10} \leq 3\} = 0.019 < \alpha = 0.05.$$

记

$$\delta = (0.05 - 0.019) / P_{0.1}\{\xi_{10} = 3\} = 0.506,$$

故所求最佳检验为如下随机化检验

$$\phi(X) = \begin{cases} 1, & \sum_{i=1}^{10} X_i > 3 \\ \delta = 0.506, & \sum_{i=1}^{10} X_i = 3 \\ 0, & \sum_{i=1}^{10} X_i < 3 \end{cases}$$

14. 解:

$$\bar{\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i, \Rightarrow E \bar{\xi} = E \xi_i = 0.5 \quad , \text{由切比雪夫不等式, 得}$$

$$P\{0.4 \leq \bar{\xi} \leq 0.6\} = P\{|\bar{\xi} - E \bar{\xi}| \leq 0.1\}$$

$$P\{|\bar{\xi} - E \bar{\xi}| \leq 0.1\} \geq 1 - \frac{D \bar{\xi}}{0.1^2} = 0.9$$

$$\Rightarrow 100 \cdot \frac{1}{4n} = 0.1 \Rightarrow n = 250$$

15. 解:

$$P\left\{ \left| \frac{\bar{\xi} - \mu}{\sigma / \sqrt{n}} \right| \leq \mu \right\} = 1 - \alpha \Rightarrow \mu \in \left[\bar{\xi} - \frac{\sigma}{\sqrt{n}} \mu, \bar{\xi} + \frac{\sigma}{\sqrt{n}} \mu \right]$$

$$\Rightarrow \text{区间长度为} \frac{2\sigma}{\sqrt{n}} \mu$$

$$\frac{2\sigma}{\sqrt{n}} \mu \leq L \Rightarrow n \geq \frac{4\sigma^2 \mu^2}{L^2}$$

1.6. 证明：设 $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$, 而且 X, Y 相互独立，则

$$\frac{X/n_1}{Y/n_2} \sim F(n_1, n_2), \quad \frac{Y/n_2}{X/n_1} \sim F(n_2, n_1)$$

于是对任意的 $0 < \alpha < 1$, 存在上侧位数 $F_\alpha(n_2, n_1)$, 使得

$$P\left[\frac{Y/n_2}{X/n_1} \geq F_\alpha(n_2, n_1)\right] = \alpha$$

$$\text{而 } P\left[\frac{Y/n_2}{X/n_1} > F_\alpha(n_2, n_1)\right] = P\left[\frac{X/n_1}{Y/n_2} < \frac{1}{F_\alpha(n_2, n_1)}\right]$$

$$\begin{aligned} P\left[\frac{Y/n_2}{X/n_1} > F_\alpha(n_2, n_1)\right] \\ = P\left[\frac{X/n_1}{Y/n_2} < \frac{1}{F_\alpha(n_2, n_1)}\right] \\ = 1 - P\left[\frac{X/n_1}{Y/n_2} \geq \frac{1}{F_\alpha(n_2, n_1)}\right] \end{aligned}$$

故 $P\left[\frac{X/n_1}{Y/n_2} \geq \frac{1}{F_\alpha(n_2, n_1)}\right] = 1 - \alpha$, 又 $\frac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$, 所以 $\frac{1}{F_\alpha(n_2, n_1)}$

表示 $F(n_1, n_2)$ 的上侧 $1 - \alpha$ 分位数 $F_{1-\alpha}(n_1, n_2)$. 从而 $F_{1-\alpha}(n_1, n_2) = \frac{1}{F_\alpha(n_2, n_1)}$.

1.7. 解: H0: $\hat{p}_{ij} = p_i \cdot p_j, i, j = 1, 2$, n=520.

$$\begin{aligned} (\hat{p}_{ij}) &= (\hat{p}_i \hat{p}_j) = \left(\frac{n_{ij}}{n}, \frac{n_{ij}}{n} \right) \\ &= \begin{pmatrix} 0.352 & 0.106 \\ 0.417 & 0.125 \end{pmatrix} \end{aligned}$$

\hat{p}_{ij} 是在 H0 成立时, p_{ij} 的极大似然估计, 此时相应统计量的值为

$$\chi^2 = \sum_{i=1}^2 \sum_{j=1}^2 \frac{(n_{ij} - n\hat{p}_{ij})^2}{n\hat{p}_{ij}} = 0.1614 < \chi^2_{0.95}(1) = 3.8415$$

故不能拒绝原假设, 即认为获利与逃税有关联。

1.8. 解:

$H_0: \sigma_1^2 = \sigma_2^2 \Leftrightarrow H_1: \sigma_1^2 \neq \sigma_2^2$, 当 H0 成立时, 统计量

$$F = \frac{S_1^2}{S_2^2} \sim F(n-1, m-1) = F(4, 6),$$

$$\text{又 } F_{0.005}(4,6)=12.0275, \quad F_{0.995}(4,6)=0.0455$$

拒绝域为

$$W = \{(x_1, \dots, x_n) : s_1^2 / s_2^2 < F_{0.005}(4,6)=0.0455 \text{ or } s_1^2 / s_2^2 > F_{0.995}(4,6)=12.0275\}$$

由样本计算得

$$S_1^2 = 78.8, S_2^2 = 944.476, F(4,6) = \frac{S_1^2}{S_2^2} = \frac{78.8}{944.476} = 0.0834$$

因为 $F(4,6) = \frac{S_1^2}{S_2^2} = \frac{78.8}{944.476} = 0.0834 \notin W$, 即不能拒绝原假设, 即可认为两个公司的影片长方差相等.

第九章 方差分析与回归分析参考答案

1. 解:

$$(1): X = \begin{pmatrix} 1 & 0.5000 \\ 1 & -0.8000 \\ 1 & 0.9000 \\ 1 & -2.8000 \\ 1 & 6.5000 \\ 1 & 2.3000 \\ 1 & 1.6000 \\ 1 & 5.1000 \\ 1 & -1.9000 \\ 1 & -1.5000 \end{pmatrix}, Y = \begin{pmatrix} -0.3 \\ -1.2 \\ 1.1 \\ -3.5 \\ 4.6 \\ 1.8 \\ 0.5 \\ 3.8 \\ -2.8 \\ 0.5 \end{pmatrix}, L = X'X = \begin{pmatrix} 10.00 & 9.90 \\ 9.90 & 91.51 \end{pmatrix}.$$

$$\hat{\beta} = L^{-1} X'Y = \begin{pmatrix} -0.349 \\ 0.807 \end{pmatrix}$$

(2). 用 Matlab 软件可求解:

alpha=0.05;

[b, bint, r, rint, stats]=regress(Y, X, alpha)

得到

$$bint = \begin{pmatrix} -1.0676 & 0.3698 \\ 0.5694 & 1.0445 \end{pmatrix}$$

从而 β_1 的 0.95 置信区间是 [0.5694 1.0445].

(3) stats = (0.8846 61.3433 0.0001 0.8674)

F=61.3433, p 值为 0.0001 < 0.05, 故拒绝 H0.

(4) 残差向量

$$e = \begin{pmatrix} -0.3546 \\ -0.2056 \\ 0.7226 \\ -0.8917 \\ -0.2963 \\ 0.2929 \\ -0.4422 \\ 0.0334 \\ -0.9179 \\ 2.0593 \end{pmatrix} \Rightarrow Q_e = e'e = 6.9388$$
$$\Rightarrow \hat{\sigma}^2 = \frac{Q_e}{n-k-1} = \frac{6.9388}{8} = 0.8674.$$

(5) 给定具体的 x 值可得相应预测区间.

2. 解:

将 x_1 取倒数, 再利用线性回归进行计算, 得

$$X = \begin{pmatrix} 1.0000 & 32.1000 & 0.3333 \\ 1.0000 & 32.1000 & 0.2000 \\ 1.0000 & 32.1000 & 0.1429 \\ 1.0000 & 32.1000 & 0.0833 \\ 1.0000 & 32.1000 & 0.0500 \\ 1.0000 & 32.1000 & 0.0333 \\ 1.0000 & 33.0000 & 0.3333 \\ 1.0000 & 33.0000 & 0.2000 \\ 1.0000 & 33.0000 & 0.1429 \\ 1.0000 & 33.0000 & 0.0833 \\ 1.0000 & 33.0000 & 0.0500 \\ 1.0000 & 33.0000 & 0.0333 \\ 1.0000 & 27.6000 & 0.3333 \\ 1.0000 & 27.6000 & 0.2000 \\ 1.0000 & 27.6000 & 0.1429 \\ 1.0000 & 27.6000 & 0.0833 \\ 1.0000 & 27.6000 & 0.0500 \\ 1.0000 & 27.6000 & 0.0333 \end{pmatrix}, Y = \begin{pmatrix} 17.8000 \\ 22.9000 \\ 25.9000 \\ 29.9000 \\ 32.9000 \\ 35.4000 \\ 18.2000 \\ 22.9000 \\ 25.1000 \\ 28.6000 \\ 31.2000 \\ 34.1000 \\ 16.8000 \\ 20.0000 \\ 23.6000 \\ 28.0000 \\ 30.0000 \\ 33.1000 \end{pmatrix}$$

$$\Rightarrow \hat{\beta} = L^{-1} X' Y = \begin{pmatrix} 23.4813 \\ 0.3387 \\ -53.2528 \end{pmatrix} (L = X' X)$$

3. 解: (1)

令 $x_1 = t, x_2 = t^2$, 转化为线性回归进行统计分析, 得

$$X = \begin{pmatrix} 1.0000 & 0.0333 & 0.0011 \\ 1.0000 & 0.0667 & 0.0044 \\ 1.0000 & 0.1000 & 0.0100 \\ 1.0000 & 0.1333 & 0.0178 \\ 1.0000 & 0.1667 & 0.0278 \\ 1.0000 & 0.2000 & 0.0400 \\ 1.0000 & 0.2333 & 0.0544 \\ 1.0000 & 0.2667 & 0.0711 \\ 1.0000 & 0.3000 & 0.0900 \\ 1.0000 & 0.3333 & 0.1111 \\ 1.0000 & 0.3667 & 0.1344 \\ 1.0000 & 0.4000 & 0.1600 \\ 1.0000 & 0.4333 & 0.1878 \\ 1.0000 & 0.4667 & 0.2178 \\ 1.0000 & 0.5000 & 0.2500 \end{pmatrix}, Y = \begin{pmatrix} 11.8600 \\ 15.6700 \\ 20.6000 \\ 26.6900 \\ 33.7100 \\ 41.9300 \\ 51.1300 \\ 61.4900 \\ 72.9000 \\ 85.4400 \\ 99.0800 \\ 113.7700 \\ 129.5400 \\ 146.4800 \\ 165.0600 \end{pmatrix}, e = \begin{pmatrix} -0.0884 \\ -0.0593 \\ -0.0071 \\ 0.1081 \\ 0.0563 \\ 0.1074 \\ 0.0416 \\ 0.0387 \\ -0.0111 \\ -0.0280 \\ -0.0419 \\ -0.1028 \\ -0.1806 \\ -0.1855 \\ 0.3526 \end{pmatrix}$$

$$\hat{\beta} = L^{-1} X' Y = \begin{pmatrix} 9.2646 \\ 64.0590 \\ 493.6532 \end{pmatrix}$$

$$Q_e = e'e = 0.2456, \widehat{\sigma^2} = \frac{Q_e}{n} = 0.0164$$

(2) 用 SPSS 统计软件分析，得系数检验表如下：

Coefficients^a

| Model | Unstandardized Coefficients | | Beta | t | Sig. |
|--------------|-----------------------------|------------|------|---------|------|
| | B | Std. Error | | | |
| 1 (Constant) | 9.268 | .128 | | 72.305 | .000 |
| x1 | 64.082 | 1.106 | .192 | 57.951 | .000 |
| x2 | 493.576 | 2.016 | .812 | 244.824 | .000 |

a. Dependent Variable: y

P 值很小，故拒绝原假设。

(3) y_0 的置信度为 $1-\alpha$ 的预测区间为

$$\hat{y}_0 \pm t_{\frac{\alpha}{2}}(n-k-1)\hat{\sigma}_e \sqrt{1+GCG'}$$

4. 解：

a) 双曲线方程： $\frac{1}{y} = a + \frac{b}{x}$ 倒数变换

- b) 幂函数方程: $y = ax^b$, 对数变换
- c) 指数曲线方程: $y = ae^{bx}$, $y' = \ln y, x' = x$
- d) 指数曲线方程: $y = ae^{\frac{b}{x}}$, $y' = \ln y, x' = \frac{1}{x}$
- e) 对数曲线方程: $y = a + b \ln x$, $y' = y, x' = \ln x$
- f) S型曲线方程: $y = \frac{1}{a + be^{-x}}$, $y' = \frac{1}{y}, x' = e^{-x}$
- g) 抛物线方程: $y = b_0 + b_1 x + b_2 x^2$, $x_1 = x, x_2 = x^2$.

5. 解:

6. 解:

$$\begin{cases} -1 = \beta_0 - \beta_1 + \beta_2 \\ 0 = \beta_0 - 2\beta_2 \\ 1 = \beta_0 + \beta_1 + \beta_2 \end{cases} \Rightarrow X = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}, Y_0 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$Y_0 = X\beta$$

$$L = X'X = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}, L^{-1} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{6} \end{pmatrix}$$

$$\hat{\beta} = L^{-1}X'Y = \begin{pmatrix} \bar{y} \\ \frac{y_3 - y_1}{2} \\ \frac{y_1 + y_3 - 2y_2}{6} \end{pmatrix}$$

当 $\beta_2 = 0$ 时,

$$X = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}, X' = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}, L = X'X = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix},$$

$$L^{-1} = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = L^{-1}X'Y = \begin{pmatrix} \bar{y} \\ \frac{y_3 - y_1}{2} \end{pmatrix}, \text{不变.}$$

7. 解：

因为

$$X' = \begin{bmatrix} 1 \cdots 1 & 1 \cdots 1 & 1 \cdots 1 \\ \underbrace{0 \cdots 0}_m & \underbrace{1 \cdots 1}_m & \underbrace{-2 \cdots -2}_n \end{bmatrix},$$

故

$$X'X = \begin{bmatrix} 2m+n & m-2n \\ m-2n & m+4n \end{bmatrix},$$

$$(X'X)^{-1} = \begin{bmatrix} m+4n & 2n-m \\ 2n-m & 2m+n \end{bmatrix} \frac{1}{m^2+13mn}, \quad X'Y = \begin{bmatrix} \sum_{i=1}^m (y_i + y_{m+i}) + \sum_{i=1}^n y_{2m+i} \\ \sum_{i=1}^m y_{m+i} - 2 \sum_{i=1}^n y_{2m+i} \end{bmatrix},$$

$$\hat{\beta} = \begin{bmatrix} \hat{\theta} \\ \hat{\phi} \end{bmatrix} = (X'X)^{-1} X'Y$$

$$= \frac{1}{m^2+13mn} \begin{bmatrix} (m+4n) \sum_{i=1}^m y_i + 6n \sum_{i=1}^n y_{m+i} + 3m \sum_{i=1}^n y_{2m+i} \\ (2n-m) \sum_{i=1}^m y_i + (3n+m) \sum_{i=1}^n y_{m+i} - 5m \sum_{i=1}^n y_{2m+i} \end{bmatrix}.$$

因为当 $m=2n$ 时，

$$(X'X)^{-1} = \frac{1}{30n^2} \begin{bmatrix} 6n & 0 \\ 0 & 5n \end{bmatrix} = \begin{bmatrix} \frac{1}{5n} & 0 \\ 0 & \frac{1}{6n} \end{bmatrix},$$

所以

$$\text{Cov}(\hat{\beta}, \hat{\beta}) = \sigma^2 \begin{bmatrix} \frac{1}{5n} & 0 \\ 0 & \frac{1}{6n} \end{bmatrix} = \begin{bmatrix} \text{Cov}(\hat{\theta}, \hat{\theta}) & \text{Cov}(\hat{\theta}, \hat{\phi}) \\ \text{Cov}(\hat{\theta}, \hat{\phi}) & \text{Cov}(\hat{\phi}, \hat{\phi}) \end{bmatrix}.$$

即 $\text{Cov}(\hat{\theta}, \hat{\phi}) = 0$. 故 $\hat{\theta}$ 与 $\hat{\phi}$ 这时不相关. 又因诸 ϵ_i 相互独立同服从正态分布 $N(0, \sigma^2)$, 所以 $\hat{\theta}, \hat{\phi}$ 均服从正态分布, 从而知 $\hat{\theta}$ 与 $\hat{\phi}$ 相互独立.

8. 解：

$$X = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}, L^{-1} = (X'X)^{-1} = \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{6} \end{pmatrix}$$

$$L^{-1}X' = \begin{pmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ 0 & -\frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

$$\hat{\beta} = \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{y_1+2y_2+y_3}{6} \\ \frac{2y_3-y_1}{3} \end{pmatrix}$$