# STAT 2006 Assignment 3
## Due Time and Date: 5 p.m., 23 April, 2020

1. (a) Note that $0 \le a \le b$ as they are the quantiles of $\chi^2(n-1)$. From the given constraint, $1-\alpha = Pr\left\{a \le \frac{(n-1)S^2}{\sigma^2} \le b\right\} = Pr\left\{\sqrt{\frac{(n-1)S^2}{b}} \le \sigma \le \sqrt{\frac{(n-1)S^2}{a}}\right\}$. So a confidence interval for $\sigma$ is in the form of $\left[\sqrt{\frac{(n-1)S^2}{b}}, \sqrt{\frac{(n-1)S^2}{a}}\right]$, the length of the interval is $k = \sqrt{(n-1)S^2}\left(\frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}\right)$.

   (b) Method 1: Note that the pdf of $\chi^2(n-1)$ is $g(u) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-1}{2}}} u^{\frac{n-1}{2}-1} e^{-\frac{u}{2}}, u > 0$. Differentiate both sides of the constraint with respect to $a$, we have

$$g(b)\frac{\partial b}{\partial a} - g(a) = 0 \Rightarrow \frac{\partial b}{\partial a} = \frac{g(a)}{g(b)} = \left(\frac{a}{b}\right)^{\frac{n-3}{2}} e^{-\frac{a-b}{2}}.$$

   Using results from (a),

$$\frac{\partial k}{\partial a} = \sqrt{(n-1)s^2}\left(\frac{-1}{2a^{\frac{3}{2}}} + \frac{1}{2b^{\frac{3}{2}}}\frac{\partial b}{\partial a}\right) = \frac{\sqrt{(n-1)}s^2 e^{\frac{b}{2}}}{2b^{\frac{n}{2}}a^{\frac{3}{2}}}\left(a^{\frac{n}{2}}e^{-\frac{a}{2}} - b^{\frac{n}{2}}e^{-\frac{b}{2}}\right).$$

   Therefore for any local minimum, it must satisfy the condition for critical point :

$$\frac{\partial k}{\partial a} = 0 \Rightarrow a^{\frac{n}{2}}e^{-\frac{a}{2}} - b^{\frac{n}{2}}e^{-\frac{b}{2}} = 0.$$

   Method 2: The Lagrange function is

$$L(a,b,\lambda) = \sqrt{(n-1)s^2}\left(\frac{1}{\sqrt{a}} - \frac{1}{\sqrt{b}}\right) - \lambda\left(G(a) - G(b) - (1-\alpha)\right).$$

   Differentiating the Lagrange function with respect to $a, b$ and $\lambda$ and set all of them to be zero:

$$\begin{cases} \frac{\partial L}{\partial a} = -\frac{1}{2}\sqrt{(n-1)s^2}a^{-\frac{3}{2}} - \lambda g(a) = 0 \\ \frac{\partial L}{\partial b} = \frac{1}{2}\sqrt{(n-1)s^2}b^{-\frac{3}{2}} + \lambda g(b) = 0 \\ \frac{\partial L}{\partial \lambda} = G(b) - G(a) - (1-\alpha) = 0 \end{cases}$$

   Therefore, by eliminating $\lambda$ in the first two equations, we have $a^{\frac{3}{2}}g(a) = b^{\frac{3}{2}}g(b)$. By substituting $g(a)$ and $g(b)$ in the pdf of $\chi^2(n-1)$, we have $a^{\frac{n}{2}}e^{-\frac{a}{2}} - b^{\frac{n}{2}}e^{-\frac{b}{2}} = 0$.

2. (a) A two-sided 0.95 confidence interval for $p_1 - p_2$ is

$$\left[\hat{p_1} - \hat{p_2} - z_{0.025}\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}, \hat{p_1} - \hat{p_2} + z_{0.025}\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}\right]$$
$$\approx [-0.1844, -0.0938].$$

   (b) Since $n = 2000, y = y_1 + y_2 = 1075, \hat{p} = y/n = 0.5375$, a two-sided confidence interval for $p$ is

$$\left[\hat{p} - z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{0.025}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] \approx [0.5156, 0.5594].$$

3. (a)

$$\alpha = P(Y \le 7|H_0) = \sum_{i=1}^{7} \binom{50}{y} 0.08^y 0.92^{50-y} = 0.9562.$$

(b)

$$P(Y \le 7|p = 0.05) = \sum_{i=1}^{7} \binom{50}{y} 0.05^y 0.95^{50-y} = 0.9968.$$

4. (a) The test statistic is $\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{(81-1)516^2}{580^2} = 63.32 > \chi^2_{0.95}(80) = 60.39$, so don't reject $H_0$.
   (b) $P = P(s \le 516|\sigma = 580) = P(\chi^2(80) \le 63.32) \approx 0.085$.

5. (a) The test statistic $Z = \frac{\bar{X}-110}{10\sqrt{16}} = 1.6 < Z_{0.01} = 2.33$, so we don't reject $H_0$.
   (b) $1.6 < Z_{0.05} = 1.645$, so we don't reject $H_0$.
   (c) $p$-value $= P(Z \ge 1.6) = 0.0548$.

6. (a) Let $Y$ denote hte number of students who find 6F's, then by CLT, $Y/n \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right)$, under $H_0$, $\frac{Y/n - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \le -Z_{0.05} = -1.645$, so we get the critical region $Y \le 102$.
   (b) $110 > 102$, so we don't reject $H_0$.
   (c) $\frac{Y/n - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = -0.66$, so p-value$=P(Z \le -0.66) = 0.2546$.

7. Let $p$ be the probability of getting a head. Observing that $560 > 500$, we have $H_0 : p = 0.5, H_1 : P > 0.5$.
   Let $X$ be the number of heads out of 1000. If the coin is fair, then $X \sim Bin(1000, 0.5)$.
   Method 1, direct calculation (by programming):

$$P(X \ge 560) = \sum_{x=560}^{1000} \binom{1000}{x} 0.5^x 0.5^{n-x} \approx 0.0000825.$$

   Method 2, by CLT, approximately, $X \sim N(500, 250)$, then

$$P(X \ge 560) = P\left(\frac{X - 500}{\sqrt{250}} \ge \frac{559.5 - 500}{\sqrt{250}}\right) = P(Z \ge 3.763) \approx 0.0000839.$$

   Since p-value is very close to 0, we reject $H_0$ and tend to believe that the coin is not fair.

8. By CLT, $Z = \left(\sum_i X_i - np\right)/\sqrt{np(1-p)}$ is approximately $N(0,1)$. For a test that rejects $H_0$ when $\sum_i X_i - np > c$, we need to find $c$ and $n$ to satisfy

$$P\left(Z > \frac{c - 0.49n}{\sqrt{n \times 0.49 \times 0.51}}\right) = 0.01 \text{ and } P\left(Z > \frac{c - 0.51n}{\sqrt{n \times 0.49 \times 0.51}}\right) = 0.99$$

   Thus we want $\frac{c-0.49n}{\sqrt{n \times 0.49 \times 0.51}} = 2.33$ and $\frac{c-0.51n}{\sqrt{n \times 0.49 \times 0.51}} = -2.33$. Solving these equations gives $n = 13567$ and $c = 6783.5$.

9. Under $H_0, Y \sim U(0,1)$, so $P(Y_n \ge 1) = 0$,

$$\alpha = P(Y_1 \ge k|\theta = 0) = (1-k)^n.$$

   Thus, $k = 1 - \alpha^{1/n}$.

10. Let $X \sim Poisson\left(\lambda\right)$, and we observe $X = 10$. To assess if the accident rate has dropped, we would calculate

$$P\left(X < 10 | \lambda = 15\right) = \sum_{x=0}^{10} \frac{e^{-15} 15^x}{x!} \approx 0.11846.$$

This is a fairly large value, so not overwhelming evidence that the accident rate has dropped (A normal approximation with continuity correction gives a value of 0.12264).