Assignment 8

12310520

苟煜涵

T1. (1) Let $X=(X_1, X_2 \cdots X_{20})$ be a simple random sample from the population $X \sim N(\mu, 1)$ $\quad n=20$

$$\alpha = P(\bar{X} > 2.6 \mid H_0 \text{ is true}) = P\left(\frac{\bar{X}-2}{\sqrt{1}/\sqrt{20}} > \frac{2.6-2}{\sqrt{1}/\sqrt{20}}\right)$$

$$= P(Z > 2.683) = 1-\Phi(2.683)$$

$$= 1-0.9963 = 0.0037$$

$$\Rightarrow \alpha \approx 0.0037$$

$$1-\beta = P(\bar{X} > 2.6 \mid H_1 \text{ is true}) = P\left(\frac{\bar{X}-3}{\sqrt{1}/\sqrt{20}} > \frac{2.6-3}{\sqrt{1}/\sqrt{20}}\right)$$

$$= P(Z > -1.789) = 1-\Phi(-1.789)$$

$$\Rightarrow \beta = 0.0367$$

(2) $$1-\beta' = P(\bar{X} > 2.6 \mid H_1 \text{ is true}) = P\left(\frac{\bar{X}-3}{6/\sqrt{n}} > \frac{2.6-3}{6/\sqrt{n}}\right)$$

$$= P(Z > -0.4\sqrt{n}) = 1-P(Z \leq -0.4\sqrt{n})$$

$$\Rightarrow \beta' = \Phi(Z \leq -0.4\sqrt{n}) \leq 0.01$$

$$\Rightarrow \quad -0.4\sqrt{n} \leq -2.33 \quad \Rightarrow \quad n \geq 33.93$$

So n's minimum sample size is 34.

T2. (1) Let X be time spend on TV per week.

The testing problem: $H_0: \mu_X = 8 \leftrightarrow H_1: \mu_X < 8$

$\bar{X} = 6.5$, $s=2$, $Z_{0.05} = 1.645$ $\quad n=100$

$$T = \frac{\bar{X}-8}{S/\sqrt{n}} \overset{\text{approx}}{\sim} N(0,1) \text{ under } H_0, \quad RR: \{X : T < -1.645\}$$

t的与 Z_α 比

$t_{obs} = \frac{6.5-8}{2/10} = -7.5 < -1.645 \Rightarrow$ we would reject H_0 at significance level $\alpha = 0.05$

(2) $p\text{-value} = P(T < t_{obs} \mid H_0 \text{ is true}) = P(T < -7.5)$

你需要表示 $= \Phi(-7.5) \approx 0 < 0.05$ p-value为概率 $\Phi(\ )$ 与α比

we would reject H_0 at significance level $\alpha = 0.05$

The conclusion holds the same.

T3. $\sigma = 2$, $\mu_{\bar{X}} = 7.5$ $\quad Z_{0.05} = 1.645$

with the true value of $\mu_X = 7.5$, $\sigma = 2$

$$T \overset{approx}{\sim} N\left(\frac{7.5-8}{2/\sqrt{n}}, 1\right) = N\left(-\frac{\sqrt{n}}{4}, 1\right)$$

$$\text{power} = P(T < -1.645 \mid \mu_X = 7.5) = P\left(T + \frac{\sqrt{n}}{4} < -1.645 + \frac{\sqrt{n}}{4}\right)$$

$$= \Phi\left(-1.645 + \frac{\sqrt{n}}{4}\right) \geq 0.9$$

$$\Rightarrow -1.645 + \frac{\sqrt{n}}{4} \geq 1.29 \Rightarrow n \geq 137.82$$

n's minimum sample size is 138

T4. (1) Let X and Y be binary random variables suggesting whether participants in town A and B support the candidate.

$\hat{p_A} = \bar{X} = 0.45$, $\hat{p_B} = \bar{Y} = 0.35$

$\hat{p_A} - \hat{p_B} = \bar{X} - \bar{Y}$ is a consistent estimator of $p_A - p_B$

and $\bar{X} - \bar{Y} \overset{approx}{\sim} N\left(p_A - p_B, \frac{p_A(1-p_A)}{n} + \frac{p_B(1-p_B)}{n}\right)$

The hypothesis testing problem is $H_0: p_A - p_B = 0 \leftrightarrow H_1: p_A - p_B \neq 0$

$$T = \frac{\hat{p_A} - \hat{p_B} - 0}{\sqrt{\frac{\hat{p_A}(1-\hat{p_A}) + \hat{p_B}(1-\hat{p_B})}{n}}} \overset{approx}{\sim} N(0,1) \text{ under } H_0$$

H_1 is two-sided $\Rightarrow$ RR : $\{X-Y : |T| > Z_{\alpha/2}\}$ $\alpha = 0.02$

$Z_{0.01} = 2.33$ $\Rightarrow$ RR : $\{X-Y : |T| > 2.33\}$

$$t_{obs} = \frac{0.1}{\sqrt{\frac{0.45\times0.55+0.35\times0.65}{400}}} = 2.90 \text{ , } |t_{obs}| > 2.33$$

the evidence against H_0 is sufficient, we would reject H_0 at significance level $\alpha=0.02$. There is significance difference between the support rates of candidate in town A and town B

(2) $$\hat{\sigma}_A = \sqrt{\frac{\hat{P}_A(1-\hat{P}_A)}{n}} = \sqrt{\frac{0.45\times0.55}{400}} = 0.0249$$

$$\hat{\sigma}_B = \sqrt{\frac{\hat{P}_B(1-\hat{P}_B)}{n}} = \sqrt{\frac{0.35\times0.65}{400}} = 0.0238$$

$\alpha = 0.02$ $Z_{\alpha/2} = Z_{0.01} \doteq 2.33$

CI of P_A : $\hat{P}_A \pm Z_{0.01}\sigma_A = 0.45 \pm 2.33\times0.0249 = (0.392, 0.508)$

CI of P_B : $\hat{P}_B \pm Z_{0.01}\sigma_B = 0.35 \pm 2.33\times0.0238 = (0.295, 0.405)$

$$\hat{\sigma}_{A-B} = \sqrt{\frac{\hat{P}_A(1-\hat{P}_A)+\hat{P}_B(1-\hat{P}_B)}{n}} = 0.0345$$

CI of $P_A - P_B$: $\hat{P}_A - \hat{P}_B \pm Z_{0.01}\hat{\sigma}_{A-B} = 0.1 \pm 2.33\times0.0345$
$= (0.020, 0.180)$

(3) No, CI of P_A and CI of P_B are independent results of estimation. Even if the two CIs overlaps, the CI of $P_A - P_B$ can still not contain 0, which means they can still be significantly different.

(4) Yes, the CI of $P_A - P_B$ is $(0.020, 0.180)$ which doesn't contain 0, meaning they are significantly different at significant level $\alpha=0.02$. It's equivalent two-tailed test.

T5. (1) $T = \dfrac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$, under $H_0: \mu_1 - \mu_2 = 0$

$\Rightarrow T \overset{approx}{\sim} N(0,1)$ under H_0

$\alpha = 0.05$ RR: $\{X-Y: T > 1.645\}$

(2) true mean difference: $\mu_1 - \mu_2 = 2$ $n=10$, $m=11$

$$T \sim N\left(\frac{\mu_1-\mu_2}{\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}}, 1\right) = N(1.144, 1)$$

$$1-\beta = P(T > 1.645 \mid H_1 \text{ is true})$$

$$\Rightarrow \beta = P(T \le 1.645 \mid H_1 \text{ is true}) = P\left(\frac{T-1.144}{\sqrt{1}} \le \frac{1.645-1.144}{\sqrt{1}}\right)$$

$$= \Phi(0.501) = 0.695 \quad \Rightarrow \beta = 0.695$$

(3) power $= 1-\beta = P(T > 1.645 \mid H_1 \text{ is true}) \geq 0.9$

$$\Rightarrow P(T \le 1.645 \mid H_1 \text{ is true}) \le 0.1 \quad \mu_T = \frac{\mu_1-\mu_2}{\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}}$$

$$= P\left(\frac{T-\mu_T}{\sqrt{1}} \le \frac{1.645-\mu_T}{\sqrt{1}}\right) = \Phi(1.645-\mu_T) \le 0.1$$

$$1.645 - \mu_T \le -1.29 \quad \mu_T = \frac{2}{\sqrt{\frac{16}{n}+\frac{16}{m}}} \geq 2.935$$

when $n=m$ we can get $\min(n+m)$

$$\frac{2}{\sqrt{\frac{32}{n}}} \geq 2.935 \Rightarrow \frac{\sqrt{2n}}{4} \geq 2.935 \Rightarrow n \geq 68.9$$

if $n=68, m=69$ $\mu_T = \dfrac{2}{\sqrt{\frac{16}{68}+\frac{16}{69}}} = 2.926 < 2.935$

$\Rightarrow n = m = 69$ $N_{min} = 138$