14.12.2021

# Seminar W11 – 831

**Exercise 3.** A vending machine contains the following numbers of bills:

| 1 RON | 5 RON | 10 RON |
|---|---|---|
| 25 | 15 | 10 |

Assume that each client pays the vending machine with only one bill. At a 5% significance level, test the following hypotheses:

(a) the mean value of the amount of money paid by a client is less than 5 RON.

(b) the standard deviation of the amount of money paid by a client is larger than 5 RON.

(c) the proportion of clients that pay the vending machine with a 5 RON bill is 40%.

*Proof:* We want to find $\mu$.

$\alpha = 0.05$

| *Parameter $\theta$ to be tested* | *Known* | *TS (test statistic)* |
|---|---|---|
| $\mu$ | $\sigma$ | $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \in \mathcal{N}(0,1)$ |
| $\mu$ | | $Z = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \in T(n-1)$ |
| $\sigma$ | | $Z = \frac{(n-1)s^2}{\sigma^2} \in \chi^2(n-1)$ |
| $p$ | | $Z = \frac{\overline{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \in \mathcal{N}(0,1)$ |
| $p_1 - p_2$ | | $Z = \frac{\overline{p}_1-\overline{p}_2-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}} \in \mathcal{N}(0,1)$ |

We don't know $\sigma$

$$\Rightarrow Z = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}}$$

$$\overline{X} = \frac{25 \cdot 1 + 15 \cdot 5 + 10 \cdot 10}{50} = 4$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2} = \sqrt{\frac{1}{49} \cdot \left(25 \cdot (1-4)^2 + 15 \cdot (5-4)^2 + 10\,(10-4)^2\right)}$$

$$= \frac{10\sqrt{6}}{7}$$

$$\begin{cases} H_0: \mu = 5 \\ H_1: \mu < 5 \end{cases} \Rightarrow \text{left-tailed test}$$

$$RR = \begin{cases} (-\infty, z_\alpha], \text{ for a left-tailed test} \\ [z_{1-\alpha}, \infty), \text{ for a right-tailed test} \\ (-\infty, z_{\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty), \text{ for a two-tailed test} \end{cases}$$

left tailed test $\Rightarrow$ $RR = (-\infty, z_\alpha]$

$z \in T(n-1) \Rightarrow z_\alpha = \text{tinv}(\alpha, n-1) = \text{tinv}(0.05, 49) = -1.6766$

$$Z = \frac{\overline{X} - \mu}{\frac{s}{\sqrt{n}}} \qquad \overline{X} = 4$$

$$s = \frac{10\sqrt{6}}{7}$$

$$z_0 = \frac{4-5}{\frac{10\sqrt{6}}{7 \cdot \sqrt{49}}} = \frac{-1 \cdot 49}{10\sqrt{6}} = -2.0004$$

We reject the hypothesis $\Leftrightarrow z_0 \in RR$

$z_0 \in RR$?

$-2.0004 \in (-\infty, -1.6766)$? YES!

$\Rightarrow H_0$ is rejected $\Rightarrow H_1$ is accepted $\Rightarrow \mu < 5$

(b) $\begin{cases} H_0 : \sigma = 5 \\ H_1 : \sigma > 5 \end{cases} \Rightarrow$ right-tailed test

| Parameter $\theta$ to be tested | Known | TS (test statistic) |
|---|---|---|
| $\mu$ | $\sigma$ | $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \in \mathcal{N}(0,1)$ |
| $\mu$ | | $Z = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \in T(n-1)$ |
| $\sigma$ | | $Z = \frac{(n-1)s^2}{\sigma^2} \in \chi^2(n-1)$ $\leftarrow$ |
| $p$ | | $Z = \frac{\overline{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \in \mathcal{N}(0,1)$ |
| $p_1 - p_2$ | | $Z = \frac{\overline{p}_1-\overline{p}_2-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}} \in \mathcal{N}(0,1)$ |

$$Z = \frac{(n-1)\cdot s^2}{\sigma^2} \in \chi^2(n-1)$$

$$\alpha = 0.05$$

$$RR = [z_{1-\alpha}, \infty)$$

$$z_{1-\alpha} = \text{chi2inv}(1-\alpha, n-1) = \text{chi2inv}(0.95, 49) = 66.339$$

$$\Rightarrow RR = [66.339, \infty)$$

$$Z = \frac{(n-1)\cdot s^2}{\sigma^2} \qquad z_0 = \frac{49\cdot\left(\frac{10\sqrt{6}}{7}\right)^2}{25} = \frac{600}{25} = 40$$

$$z_0 = 40 \notin [66.339, \infty) = RR$$

$\Rightarrow$ $H_0$ is not rejected $\Rightarrow$ $H_0$ is accepted.

(c) $\begin{cases} H_0: p = 0.4 \\ H_1: p \neq 0.4 \end{cases}$ $\Rightarrow$ two-tailed test

| Parameter $\theta$ to be tested | Known | TS (test statistic) |
|---|---|---|
| $\mu$ | $\sigma$ | $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \in \mathcal{N}(0,1)$ |
| $\mu$ | | $Z = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \in T(n-1)$ |
| $\sigma$ | | $Z = \frac{(n-1)s^2}{\sigma^2} \in \chi^2(n-1)$ |
| $p$ | | $Z = \frac{\overline{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \in \mathcal{N}(0,1)$ |
| $p_1 - p_2$ | | $Z = \frac{\overline{p}_1-\overline{p}_2-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}} \in \mathcal{N}(0,1)$ |

$$TS = Z = \frac{\overline{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \in \mathcal{N}(0,1)$$

$$\overline{p} = \frac{15}{50} = 0.3$$

$$z_0 = \frac{0.3 - 0.4}{\sqrt{\frac{0.4 \cdot 0.6}{50}}} = \frac{-0.1 \cdot \sqrt{50}}{\sqrt{0.24}} = \frac{-\frac{1}{10} \cdot \sqrt{50}}{\sqrt{0.24}} = -1.4434$$

$$RR = (-\infty, z_{\frac{\alpha}{2}}] \cup [z_{1-\frac{\alpha}{2}}, \infty)$$

$$z_{\frac{\alpha}{2}} = \text{norminv}\left(\frac{\alpha}{2}\right) = -1.96$$

$$z_{1-\frac{\alpha}{2}} = \text{norminv}\left(1-\frac{\alpha}{2}\right) = 1.96$$

$$RR = (-\infty, -1.96] \cup [1.96, \infty)$$

$$z_0 = -1.4434$$

$z_0 \notin RR \Rightarrow$ we do not reject $H_0 \Rightarrow$ we accept $H_0$

| | $H_0$ | $H_1$ |
|---|---|---|
| reject $H_0$ | false negative $P(Z_0 \in RR \mid H_0) = \alpha$ | ✓ |
| accept $H_0$ | ✓ | false positive $\beta$ |

**Exercise 2.** In a pre-election poll, we are interested in the proportion $p_A$ of people who plan to vote for candidate $A$ against candidate $B$.

(a) Given that 530 persons out of a random sample of 1000 persons support $A$, is, at 5% significance level, candidate $A$ favourite to win the elections? Find the corresponding $p$-value of the test.

(b) Estimate the minimum number $n_A$ of persons that must support $A$ out of a random sample of 900 persons, in order to conclude that candidate $A$ is favourite to win the elections, at 5% significance level.

Sol. : (a)

| *Parameter $\theta$ to be tested* | *Known* | *TS (test statistic)* |
|---|---|---|
| $\mu$ | $\sigma$ | $Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \in \mathcal{N}(0,1)$ |
| $\mu$ | | $Z = \frac{\overline{X}-\mu}{\frac{s}{\sqrt{n}}} \in T(n-1)$ |
| $\sigma$ | | $Z = \frac{(n-1)s^2}{\sigma^2} \in \chi^2(n-1)$ |
| $p$ | | $Z = \frac{\overline{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \in \mathcal{N}(0,1)$ |
| $p_1 - p_2$ | | $Z = \frac{\overline{p}_1-\overline{p}_2-(p_1-p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1}+\frac{p_2(1-p_2)}{n_2}}} \in \mathcal{N}(0,1)$ |

We want to find $p$, $\quad TS = Z = \dfrac{\overline{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$

$\overline{p} = 0.53$

$$\begin{cases} H_0 : p = 0.5 \\ H_1 : p > 0.5 \end{cases}$$

(b) *For **significance testing**: Let $F_Z$ be the cdf of the test statistic $Z$. Find the P-value:*

$$P = \begin{cases} F_Z(Z_0), \text{ for a left-tailed test} \\ 1 - F_Z(Z_0), \text{ for a right-tailed test} \\ 2 \cdot \min\{F_Z(Z_0), 1-F_Z(Z_0)\}, \text{ for a two-tailed test} \end{cases}$$

*If $P \le \alpha$, then the hypothesis $H_0$ is rejected, otherwise it is accepted.*

$$Z_0 = \frac{0.53 - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{1000}}} = \frac{0.03}{\frac{0.5}{10\sqrt{10}}} = \frac{0.3}{0.5} \cdot \sqrt{10} = \frac{3}{5}\sqrt{10} = 1.8974$$

Test is right tailed $\Rightarrow$ $P = 1 - F_Z(Z_0) = 1 - \text{normcdf}(1.8974) = 0.0288$

If $p \le \alpha \Rightarrow H_0$ is rejected

$0.0288 \le 0.05 \Rightarrow H_0$ is rejected $\Rightarrow H_1$ is accepted $\Rightarrow p > 0.5$

(b) $$z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$ $n = 900$

$$\begin{cases} H_0: & p = 0.5 \\ H_1: & p > 0.5 \end{cases}$$

We have to find $\bar{p}$ so that $z_0 \in RR$

$$z_0 = \frac{\bar{p} - 0.5}{\sqrt{\frac{0.5 \cdot 0.5}{900}}} = \frac{\bar{p} - 0.5}{\frac{0.5}{30}}$$

$$RR = [z_{1-\alpha}, \infty)$$

$$z_{1-\alpha} = \text{norminv}(1-\alpha) = \text{norminv}(0.95) = 1.6449$$

We must find $\bar{p}$ so that

$$\frac{\bar{p} - 0.5}{\frac{0.5}{30}} \in [1.6449, \infty)$$

$$\Leftrightarrow \quad \bar{p} \ge 0.5 + \frac{0.5}{30} \cdot 1.6449 = 0.52741$$

$\Leftrightarrow$ # of people voting for A in this sample $\ge 900 \cdot 0.52741 = 474.67$