Exercises - Lecture 5

Exercise 1:

1) $P(|X-3|>6) = P(-6 > X-3) + P(X-3>6) = P\left(-2 > \frac{X-3}{3}\right) + P\left(\frac{X-3}{3} > 2\right)$

with $\frac{X-3}{3} \hookrightarrow \mathcal{N}(0,1)$ $\quad P(|X-3|>6) = \Phi(-2) + (1-\Phi(2)) = 2 - 2\Phi(2)$

$P(|X-3|>6) = 2 - 2 \times 0.9772 = 0.0456$

2) $P(2.5 < X < 6.5) = P\left(\frac{2.5-5}{2} < Z < \frac{6.5-5}{2}\right) = P(Z < 0.75) - P(-1.25 < Z)$

$= \Phi(0.75) - (1 - \Phi(1.25)) = 0.7734 - (1 - 0.8944)$

$= 0.6678$

3) $P(X < x) = 0.42 \iff P\left(Z < \frac{x-3}{\sqrt{2.25}}\right) = 0.42$ and $0.42 = 1 - 0.58$

$\left.\begin{array}{l} 0.42 = 1 - \Phi(0.21) \\ 0.42 = \Phi(-0.21) \end{array}\right\}$ so $\quad \frac{x-3}{\sqrt{2.25}} = 0.21 \iff x = 3 - 0.21\sqrt{2.25} = 2.68$

4) $P(\mu - x < X < \mu + x) = 0.9$

$P\left(\frac{-x}{2} < Z < \frac{x}{2}\right) = 0.9 \iff \Phi\left(\frac{x}{2}\right) - \Phi\left(-\frac{x}{2}\right) = 0.9$

$\iff 2\Phi\left(\frac{x}{2}\right) - 1 = 0.9 \iff \Phi\left(\frac{x}{2}\right) = 0.95$

$\iff \frac{x}{2} = 1.65 \iff x = 3.30$

Exercise 2:

1) One sample t-test: $\mu = 2.5$, here $M = 2.6$ and $S^2 = 0.16$

$Z = \sqrt{n}\,\frac{M-\mu}{S} \sim t(n-1) \sim \mathcal{N}(0,1)$ $\quad$ with $\begin{cases} H_0 : \mu = 2.5 \\ H_1 : \mu \neq 2.5 \end{cases}$

with a risk $\alpha = 5\%$, the limits are $\Phi(z_{\alpha/2}) = \frac{\alpha}{2}$ and $\Phi(z'_{\alpha/2}) = 1 - \frac{\alpha}{2}$

ie: -1.96 and $+1.96$: zone of acceptance of H_0.

Here, the observation is $\sqrt{100}\,\frac{0.1}{\sqrt{0.16}} = 2.5 \gg 1.96$: H_0 is rejected.

level of significance (p-value): $P(M-\mu \geqslant 0.1) = P(Z \geqslant 2.5) = 1 - \Phi(2.5)$

$= 6.2 \cdot 10^{-3}$ if unilateral test

$P(|M-\mu| \geqslant 0.1) = P(Z \geqslant 2.5 \cup Z \leqslant -2.5) = 1.24 \cdot 10^{-2}$ (bilateral)

2) Unpaired sample t-test. observations (M_1, S_1) and (M_2, S_2) with effectives n_1 and n_2.

$$Z = \frac{M_1 - M_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim t(n_1+n_2-2) \sim \mathcal{N}(0,1) \qquad \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

here $z_{obs} = \frac{-30}{\sqrt{\frac{400}{100} + \frac{100}{150}}} = -13.9$ with $\alpha = 5\%$ the interval of acceptance of H_0 is $[-1.96 ; 1.96]$ so H_0 is rejected.

p-value : $P(M_1 - M_2 \leqslant -30) = P(Z \leqslant -13.9) = 10^{-10}$ (unilateral)

$P(|M_1 - M_2| \geqslant 30) = P(Z \leqslant -13.9) + P(Z \geqslant 13.9) = 2 \cdot 10^{-10}$ (bilateral)

3) Paired sample t-test : D: difference between the 2 diagnostics.

$$Z = \sqrt{n}\,\frac{D}{S} \sim t(n-1) \sim \mathcal{N}(0,1) \qquad \begin{cases} H_0 : E(D) = 0 \\ H_1 : E(D) \neq 0 \end{cases}$$

here $z_{obs} = \sqrt{36}\,\frac{0.5}{\sqrt{1}} = 3$ with $\alpha = 5\%$ H_0 is rejected ($z_{obs} >> 1.96$)

p-value : $P(D \geqslant 0.5) = \mathbb{P}(Z \geqslant 3) = 1 - \mathbb{P}(Z \leqslant 3) = 0.0013$ (unilateral)

$P(|D| \geqslant 0.5) = 2\,\mathbb{P}(Z \geqslant 3) = 0.0026$ (bilateral)

Exercice 3: Confidence interval.

V: percentage of persons immunized. $E(V) = \mu$: value to estimate.

$$P(\mu - 1 < V < \mu + 1) = 0.95$$

$$\Leftrightarrow P\left(-\sqrt{\frac{n}{\sigma^2}} < \frac{V-\mu}{\sqrt{\sigma^2/n}} < \sqrt{\frac{n}{\sigma^2}}\right) = 0.95 \quad \text{with } \frac{V-\mu}{\sqrt{\sigma^2/n}} \hookrightarrow \mathcal{N}(0,1)$$

$$\Leftrightarrow P\left(Z < \sqrt{\frac{n}{\sigma^2}}\right) - P\left(Z < -\sqrt{\frac{n}{\sigma^2}}\right) = 0.95$$

$$\Leftrightarrow 2\,\Phi\left(\sqrt{\frac{n}{\sigma^2}}\right) - 1 = 0.95$$

$$\Leftrightarrow \Phi\left(\sqrt{\frac{n}{\sigma^2}}\right) = 0.975 \Leftrightarrow \sqrt{\frac{n}{\sigma^2}} = 1.96 \Leftrightarrow n = (1.96)^2 \times \sigma^2$$

Here, if $\sigma = 2.5$, $n = 24$.

Exercise 4: Chi-2 test of independence

Three categories: C, I, S and 2 treatments A and B

Contingency table in the case of independence:

Treat.	C	I	S	Sum
A	300	180	120	600
B	200	120	80	400
Sum	500	300	200	1000

$$Y = \sum_{cat.} \sum_{treat} \frac{(O-E)^2}{E} \sim \chi^2((3-1)(2-1))$$

$\begin{cases} H_0: \text{categories and treatment are independent} \\ H_1: \text{they are dependent.} \end{cases}$

Here $y_{obs} = \frac{(280-300)^2}{300} + \frac{(220-200)^2}{200} + \dots = 17.9.$

with $\alpha = 5\%$ $\quad y_{lim} = 5.99$ with the $\chi^2_{(dff=2)}$ distribution.

p-value: $P(y \geq 17.9) \ll 0.01$

Exercise 5: ANOVA

3 varieties $\begin{cases} H_0: \mu_A = \mu_B = \mu_C \\ H_1: \text{at least one mean is different} \end{cases}$

$$F = \frac{SS_{factor}/(3-1)}{SS_{residual}/(15-3)} \sim \mathcal{F}(2,12)$$

Here, $M = 5$, $SS_{factor} = \sum_i n_i (M_i - M)^2$ and $SS_{residual} = \sum_i \sum_{j=1}^{5} (x_j^i - M_i)^2$

with $i = \{A, B, C\}$, we have: $\begin{cases} SS_{factor} = 40 \\ SS_{residual} = 16 \end{cases}$

$$f_{obs} = \frac{40/2}{16/12} = 15.04$$

with $\alpha = 5\%$ $\quad f_{lim} = 3.4$: H_0 is rejected.

Exercise 6: Power of a t-test

Let's consider the random variable T describing the temperature of the human body after drug treatment and Δ be the difference in temperature

H_0 : the treatment has no effect : $E(T) = 37$. i.e, $E(\Delta) = 0$

We expect an effect of the treatment with an increase of 1.5°C

i.e.: $E(T|H_1) - E(T|H_0) = 1.5 = \mu$

1) The power of a statistical test is given by:

$$1-\beta = 1 - P(\text{type II error}) = 1 - P(\text{no reject } H_0 \mid H_1 \text{ true})$$
$$= P(\text{reject } H_0 \mid H_1)$$
$$= P\left(\frac{\Delta}{\sigma}\sqrt{n} \geq t_{n-1,\alpha} \mid H_1\right)$$
$$= P\left(\Delta \geq t_{n-1,\alpha}\frac{\sigma}{\sqrt{n}} \mid H_1\right)$$
$$= P\left(\frac{\Delta-\mu}{\sigma}\sqrt{n} \geq \left(t_{n-1,\alpha}\frac{\sigma}{\sqrt{n}} - \mu\right)\frac{\sqrt{n}}{\sigma}\right)$$
$$= P\left(\frac{\Delta-\mu}{\sigma}\sqrt{n} \geq t_{n-1,\alpha} - \mu\frac{\sqrt{n}}{\sigma}\right)$$
$$= P(Z \geq 2.13 - 1.5\sqrt{5}/1)$$
$$= 1 - P(Z \leq -1.22) = 0.86$$

2) See on R.