MATH255 Autumn 2023 Computer Lab - Week 13

Note: Question 1(a) and (b) is your Lab Preparation exercise for this week. Your answers must be submitted as a pdf document before the start of your lab.

Key Results from Lectures

- $z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$, `dnorm, pnorm, qnorm, rnorm` — z-critical Value
- $t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$, `dt, pt, qt, rt` — t-critical Value
- $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ — Confidence Interval, σ Known
- $\bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ — Confidence Interval, σ Unknown
- $s^2 = \frac{(n_a-1)s_a^2+(n_b-1)s_b^2}{n_a+n_b-2}$ — Pooled Variance

1. Download the chol.csv file from Moodle and upload it to R. Remove the rows that have NA entries (`na.omit` command). The LDL cholesterol levels of the patients' first and second hospital visits are the `chol1` and `chol2` variables. You want to test, at 5% significance, whether the mean of the first-visit levels is equal to 7.5.

   (a) Formulate the null and alternative hypotheses. Write them in mathematical notation.

   (b) Is a one-tailed or two-tailed test appropriate here?

   (c) Calculate the t-statistic by hand, $t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$. You can use R to find $\bar{x}, s$ and n.

   (d) Find the p-value and use it to determine whether the mean LDL levels are significantly different from the hypothesised $\mu_0 = 7.5$ at the chosen significance level.

   (e) Use the `t.test` command to do the same test. Do the t-statistic and p-value R provides agree with the ones you calculated?

2. At 1% significance, test whether the means of the two independent samples of male and female LDL first-visit levels are the same. Assume the two population variances are equal.

3. The doctor had given some instructions to all the patients to control their cholesterol and wants to know if there is a significant difference in the average change in cholesterol from the first visit to the second. Create a new variable `CHANGE = chol2 - chol1`.

   (a) Is this a 2-sample or a 1-sample t-test? Is it 2-tailed or 1-tailed?

   (b) Draw your conclusions at the $\alpha = 0.05, 0.01$ and 0.001 significance levels.

   (c) Discuss the statistical significance you found, in the language of the problem. Can you comment on the potential clinical significance?

4. The doctor believes smokers are more inclined not to follow doctor's orders than non-smokers. Test the hypothesis of the difference you created in Question 3 between smokers and non-smokers with $\alpha = 0.01$.