Department of Inter Disciplinary Studies,
Faculty of Engineering,
University of Jaffna, Sri Lanka
MC 3020 - Assignment 03

40 minutes **17 - 07 - 2024**

Important instructions:

- **Answer all the questions (1-4).**
- **If it is determined that you have violated any policies during this exam, you will receive a score of zero for this assignment, without any exceptions or considerations.**

1. Engineering students at the University of Jaffna participated in a study to evaluate the impact of a new teaching method on their academic performance. Seven students were tested before and after the implementation of the new method. The test scores (out of 100) for each student before and after the implementation are provided in the table below. Assume that the changes in test scores are approximately normally distributed.

Subject	1	2	3	4	5	6	7	8
Before	75	82	78	65	70	88	84	85
After	80	85	82	70	75	90	86	79

   (a) Can we use paired t-test for this problem? Justify your answer.

   (b) Construct a 99% confidence interval for the mean difference in test scores.

   (c) Can it be concluded that the teaching method has an impact on test scores at the 10% level of significance?

2. The administration of the University of Jaffna is conducting a study to evaluate the class attendance of engineering students in the MC3020 lecture class. They aim to determine how many students to randomly sample to estimate the proportion of students who have adequate attendance during the recent 75-day non-academic staff strike action. The administration wants the estimator to be within 0.05 of the true proportion, using a 95% confidence interval. How many students would need to be sampled to achieve this level of accuracy?

3. An experiment was conducted to evaluate the effectiveness of a new method for recording temperature at the Ariviyal Nagar weather station. A random sample of 22 days was selected, and the days were randomly divided into two groups. On twelve of the days, temperature readings were recorded using the old method, and on the remaining twelve days, the new method was used. The temperature readings (in °C) for each day are recorded in the table below:

Day	1	2	3	4	5	6	7	8	9	10	11
Old Method	28	30	29	31	27	29	30	28	31	32	29
New Method	27	29	28	30	26	28	29	27	30	31	28

   (a) Conduct an appropriate test (e.g., an F-test) to determine whether the population variances of the two methods are equal.

   (b) Based on the results from part (a), test whether the mean temperature recorded using the new method is less than the mean temperature recorded using the old method. Use an $\alpha = 0.05$ significance level.

   (c) Construct a 95% confidence interval for $\mu_1 - \mu_2$ to assess the size of the difference in the two means.

   (d) Estimate the 95% confidence interval for the ratio of standard deviations of the temperature readings from the old method to the new method.

4. As part of evaluating the impact of the recent 75-day non-academic staff strike action on engineering students at the University of Jaffna, the administration audits the number of study disruptions reported by students that were found to be unjustified. In previous years, the number of unjustified study disruptions per student had a normal distribution with a mean $\mu = 38$. Due to the significant disruption caused by the strike, the administration suspects that the mean number of unjustified disruptions has increased. An audit of 12 randomly selected engineering students is conducted to test whether there has been an increase in unjustified study disruptions. Use the sample data collected from each student:

$$39, 38, 36, 39, 39, 39, 39, 39, 39, 39, 38, 33$$

   (a) Give a point estimate of the mean number of improperly documented study sessions. Construct a 95% confidence interval for the mean number of improperly documented sessions. Interpret this interval.

   (b) Is there sufficient evidence that the mean number of improperly documented study sessions is greater than 38?. Use a $\alpha = 0.01$.

   (c) Is there sufficient evidence that the variance in the number of improperly documented study sessions is greater than 9? Use a $\alpha = 0.05$.

Formula sheet:

1. Sample size calculation based on mean confidence interval : $n = \left(Z_{\frac{\alpha}{2}} * \frac{\sigma}{E}\right)^2$ (where $E-$ margin error)

2. Sample size calculation based on proportion confidence interval : $n = \left(Z^2_{\frac{\alpha}{2}} * \frac{p(1-p)}{E^2}\right)$

3. Sample mean ($\bar{X}$) and standard deviation (s) can be estimated

$$\bar{X} = \frac{\sum X}{n}, \quad s = \sqrt{\frac{\sum\left(X - \bar{X}\right)^2}{n-1}} = \sqrt{\frac{n\sum X^2 \ - \ \left(\sum X\right)^2}{n(n-1)}}$$

4.

Parameter	$(1-\alpha) * 100\%$ confidence interval	Test statistic value
Mean (μ)	Case 1: $\left(\bar{X} \mp Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}\right)$	Case 1: $Z = \frac{\bar{X}-\mu_0}{(\frac{\sigma}{\sqrt{n}})}$
	Case 2: $\left(\bar{X} \mp t_{\frac{\alpha}{2},df} * \frac{S}{\sqrt{n}}\right)$	Case 2: $T = \frac{\bar{X}-\mu_0}{(\frac{S}{\sqrt{n}})}, df = n-1$
	Case 3: $\left(\bar{X} \mp Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}\right)$	Case 3: $Z = \frac{\bar{X}-\mu_0}{(\frac{\sigma}{\sqrt{n}})}$
	Case 4: $\left(\bar{X} \mp Z_{\frac{\alpha}{2}} * \frac{S}{\sqrt{n}}\right)$	Case 4: $Z = \frac{\bar{X}-\mu_0}{(\frac{S}{\sqrt{n}})}$
Variance (σ)	$\left(\frac{(n-1)*S^2}{\chi^2_U}, \frac{(n-1)*S^2}{\chi^2_L}\right)$	$\chi = \frac{(n-1)s^2}{\sigma_0^2}, df = n-1$

Case 1: when population is normal and σ is known, case 2: when population is normal and σ is unknown, case 3: when population is not normal and σ is known, sample size n is large and case 3: when population is not normal and σ is known, sample size n is large.

5.

Parameter	$(1-\alpha) * 100\%$ confidence interval	Test statistic value
Dependent populations	Case 1: $\left(\bar{D} \mp Z_{\frac{\alpha}{2}} * \frac{\sigma_D}{\sqrt{n}}\right)$	Case 1: $Z = \frac{\bar{D}-\mu_0}{(\frac{\sigma_D}{\sqrt{n}})}$
Mean difference ($\mu_1 - \mu_2 = \mu_D$)	Case 2: $\left(\bar{D} \mp t_{\frac{\alpha}{2},df} * \frac{s_D}{\sqrt{n}}\right)$	Case 2: $T = \frac{\bar{D}-\mu_0}{(\frac{s_D}{\sqrt{n}})}, df = n-1$
	Case 3: $\left(\bar{D} \mp Z_{\frac{\alpha}{2}} * \frac{\sigma_D}{\sqrt{n}}\right)$	Case 3: $Z = \frac{\bar{D}-\mu_0}{(\frac{\sigma_D}{\sqrt{n}})}$
	Case 4: $\left(\bar{D} \mp Z_{\frac{\alpha}{2}} * \frac{s_D}{\sqrt{n}}\right)$	Case 4: $Z = \frac{\bar{D}-\mu_0}{(\frac{s_D}{\sqrt{n}})}$

Case 1: when population distribution of the differences is normal and σ_D is known, case 2: when population distribution of the differences is normal and σ_D is unknown, case 3: when population distribution of the differences is not normal and σ_D is known, sample size n is large and case 4: when population distribution of the differences is not normal and σ_D is known, sample size n is large.

6. Sample mean ($\bar{D}$) and standard deviation (s_D) for the differences can be estimated

$$\bar{D} = \frac{\sum D}{n}, \quad s_D = \sqrt{\frac{\sum\left(D - \bar{D}\right)^2}{n-1}} = \sqrt{\frac{n\sum D^2 \ - \ \left(\sum D\right)^2}{n(n-1)}}$$

7.

Parameter	$(1-\alpha)*100\%$ **confidence interval**	**Test statistic value**
Independent populations Mean difference $(\mu_1-\mu_2)$	Case 1: $\left(\bar{X}-\bar{Y} \mp Z_{\frac{\alpha}{2}} * \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$ Case 2: $\left(\bar{X}-\bar{Y} \mp t_{\frac{\alpha}{2}} * \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)$ Case 3: $\left(\bar{X}-\bar{Y} \mp t_{\frac{\alpha}{2}} * s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$ where $s_p^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ Case 4: $\left(\bar{X}-\bar{Y} \mp Z_{\frac{\alpha}{2}} * \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$ Case 5: $\left(\bar{X}-\bar{Y} \mp Z_{\frac{\alpha}{2}} * \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)$	Case 1: $Z = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ Case 2: $T = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$ $df = \frac{(A+B)^2}{\frac{A^2}{n_1-1}+\frac{B^2}{n_2-1}}; A = \frac{s_1^2}{n_1},\ B = \frac{s_2^2}{n_2},$ Case 3: $T = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ $df = n_1+n_2-2$ Case 4: $Z = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ Case 5: $Z = \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$
Proportions difference (p_1-p_2)	$\left(\hat{p_1}-\hat{p_2} \mp Z_{\frac{\alpha}{2}} * \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1}+\frac{\hat{p_2}(1-\hat{p_2})}{n_2}}\right)$	$Z = \frac{\hat{p_1}-\hat{p_2}-(p_1-p_2)}{\bar{p}(1-\bar{p})(\frac{1}{n_1}+\frac{1}{n_2})}$ $\bar{p} = \frac{X_1+X_2}{n_1+n_2}$
Variances ratio $(\sigma_1^2/{\sigma_2}^2)$	$\left(\frac{S_1^2}{S_2^2} * \frac{1}{F_{\alpha/2,n_1-1,n_2-1}}, \frac{S_1^2}{S_2^2} * F_{\alpha/2,n_2-1,n_1-1}\right)$	$F = \frac{S_1^2}{S_2^2}$ $df = n_1-1, n_2-1$

Case 1: When the two independent population distributions are normal and the population variances σ_1^2 and σ_2^2 are known, case 2: When the two independent population distributions are normal and the population variances σ_1^2 and σ_2^2 are unknown and unequal, case 3: When the two independent population distributions are normal and the population variances σ_1^2 and σ_2^2 are unknown but equal, case 4: When the two independent population distributions are not normal and the population variances σ_1^2 and σ_2^2 are known, and the sample size n_1 and n_2 are large and case 5: When the two independent population distributions are not normal and the population variances σ_1^2 and σ_2^2 are known, and the sample size n_1 and n_2 are large.