Department of Inter Disciplinary Studies,
Faculty of Engineering,
University of Jaffna, Sri Lanka

MC 3020 - Assignment 03

55 minutes **27 - 06 - 2023**

Important instructions:

- **Answer all the questions (1-6).**
- **If it is determined that you have violated any policies during this exam, you will receive a score of zero for this assignment, <u>without</u> any exceptions or considerations.**

1. A dietitian wishes to see if a person's cholesterol level will change if the diet is supplemented by a certain mineral. Seven subjects were pretested, and then they took the mineral supplement for a 6-week period. The results are shown in the table. (Cholesterol level is measured in milligrams per deciliter.)

Subject	1	2	3	4	5	6	7
Before	210	235	208	190	172	244	232
After	190	170	210	188	173	228	232

   Assume the variable (person's cholesterol level changes) is approximately normally distributed.

   (a) Construct 99% confidence interval for the mean difference of person's cholesterol level changes.

   (b) Can it be concluded that the cholesterol level has been changed at the 10% level of significance?

2. The designer of the new operating system has decided to conduct a more extensive study. She wants to determine how many programs to randomly sample in order to estimate the proportion of Microsoft Windows–compatible programs that would perform adequately using the new operating system. The designer wants the estimator to be within 0.03 of the true proportion using a 95% confidence interval as the estimator. How many programs would need to be tested?

3. An experiment was conducted to evaluate the effectiveness of a treatment for tapeworm in the stomachs of sheep. A random sample of 24 worm-infected lambs of approximately the same age and health was randomly divided into two groups. Twelve of the lambs were injected with the drug and the remaining twelve were left untreated. After a 6-month period, the lambs were slaughtered and the worm counts were recorded in below Table:

Number	1	2	3	4	5	6	7	8	9	10	11	12
Treated Sheep	18	43	28	50	16	32	13	35	38	33	6	7
Untreated Sheepp	40	54	26	63	21	37	39	23	48	58	28	39

   Assume that the population distributions of the measurements are normal with equal variances.

   (a) Test whether the mean number of tapeworms in the stomachs of the treated lambs is less than the mean for untreated lambs. Use an a $\alpha = 0.05$ test.

   (b) Place a 95% confidence interval on to $\mu_1 - \mu_2$ assess the size of the difference in the two means.

4. An educational researcher designs a study to compare the effectiveness of teaching English to non-English-speaking people by a computer software program and by the traditional classroom system. The researcher randomly assigns 125 students from a class of 300 to instruction using the computer. The remaining 175 students are instructed using the traditional method. At the end of a 6-month instructional period, all 300 students are given an examination with the results reported in given Table

Exam Results	**Computer Instruction**	**Traditional Instruction**
Pass	94	113
Fail	31	62

   Does instruction using the computer software program appear to increase the proportion of students passing the examination in comparison to the pass rate using the traditional method of instruction? Use a $\alpha = 0.01$.

5. As a part of the evaluation of Colombo municipal council employees, the city manager audits the parking tickets issued by city parking officers to determine the number of tickets that were contested by the car owner and found to be improperly issued. In past years, the number of improperly issued tickets per officer had a normal distribution with mean $\mu = 380$. Because there has recently been a change in the city's parking regulations, the city manager suspects that the mean number of improperly issued tickets has increased. An audit of 12 randomly selected officers is conducted to test whether there has been an increase in improper tickets. Use the sample data given here which is audit collected from each officers:

$$390, 380, 369, 392, 398, 393, 392, 396, 399, 391, 387, 393$$

   (a) Give a point estimate of the mean number of improperly issued tickets. Construct a 95% confidence interval on the mean number of improperly issued tickets. Interpret this interval.

   (b) Is there sufficient evidence that the mean number of improperly issued tickets is greater than 380? Use a $\alpha = 0.01$.

   (c) Is there sufficient evidence that the variance number of improperly issued tickets is greater than 35? Use a $\alpha = 0.05$.

6. An industry pays an average wage rate of \$9.00 per hour. A sample of 36 workers from one company showed a mean wage of \$8.50 and a sample standard deviation of \$0.60. But a general perception is that the true or population standard deviation is above \$0.58. Test using a 5% level of significance whether the sample evidence is significant to justify the general perception. Assume that the sample is from a normal population.

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

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

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

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

	Parameter	$(1-\alpha)*100\%$ **confidence interval**	**Test statistic value**
7.	Independent populations	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 1: $Z=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$
	Mean difference $(\mu_1-\mu_2)$	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 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: $\left(\bar{X}-\bar{Y}\mp t_{\frac{\alpha}{2}}*s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$ where $s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$	Case 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: $\left(\bar{X}-\bar{Y}\mp Z_{\frac{\alpha}{2}}*\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$	Case 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: $\left(\bar{X}-\bar{Y}\mp Z_{\frac{\alpha}{2}}*\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)$	Case 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)	$\left(\hat{p_1}-\hat{p_2}\mp Z_{\frac{\alpha}{2}}*\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1}+\frac{\hat{p_2}(1-\hat{p_2})}{n_2}}\right)$	$Z=\frac{\hat{p_1}-\hat{p_2}-(p_1-p_2)}{\bar{p}(1-\bar{p})(\frac{1}{n_1}+\frac{1}{n_2})}$ $\bar{p}=\frac{X_1+X_2}{n_1+n_2}$

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.

	Parameter	$(1-\alpha)*100\%$ **confidence interval**	**Test statistic value**
7.	Independent populations	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 1: $Z=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$
	Mean difference $(\mu_1-\mu_2)$	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 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: $\left(\bar{X}-\bar{Y}\mp t_{\frac{\alpha}{2}}*s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$ where $s_p^2=\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$	Case 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: $\left(\bar{X}-\bar{Y}\mp Z_{\frac{\alpha}{2}}*\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$	Case 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: $\left(\bar{X}-\bar{Y}\mp Z_{\frac{\alpha}{2}}*\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right)$	Case 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)	$\left(\hat{p_1}-\hat{p_2}\mp Z_{\frac{\alpha}{2}}*\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1}+\frac{\hat{p_2}(1-\hat{p_2})}{n_2}}\right)$	$Z=\frac{\hat{p_1}-\hat{p_2}-(p_1-p_2)}{\bar{p}(1-\bar{p})(\frac{1}{n_1}+\frac{1}{n_2})}$ $\bar{p}=\frac{X_1+X_2}{n_1+n_2}$

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.