**MATH 244 (L1)** **Applied Statistics**

# Quiz 3

Name ____________________ Student ID _______________ Tutorial section ____

**Time allowed : 45 minutes**

1. (*9 marks*) A student wanted to study the ages of couples applying for marriage licenses in Hong Kong. He studied a sample of 94 marriage licenses and found that in 50 cases the husband was older than the wife.

(a) Use this sample information to construct a 92% confidence interval for the relevant population proportion.

$$\hat{\pi} \pm Z_{0.04}\sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}} = \frac{50}{94} \pm (1.75)\sqrt{\frac{(50/94)(44/94)}{94}}$$
$$= 0.5319 \pm 0.0901 = [44.18\%\,,\,62.20\%]$$

(b) Do the sample data provide evidence that the husband is usually older than the wife among couples applying for marriage licenses in Hong Kong? Explain briefly, based on your confidence interval.

No. We have high confidence that the proportion of couples applying for marriage licenses in Hong Kong with the husband being older is from 44.18% to 62.20%. Since 50% lies within this interval, the data does not provide enough evidence that the true proportion is really greater than 50%.

(c) If you are asked to conduct the survey again so that you can produce a 95% confidence interval with margin of error at most 0.08, how many couples should you sample?

$$n = \frac{Z_{0.025}^2\hat{\pi}(1-\hat{\pi})}{D^2} = \frac{(1.96)^2(50/94)(44/94)}{(0.08)^2} = 149.45 \approx 150$$

**P. T. O**

2. (*8 marks*) A major geological survey of "The Seven Islands" was recently completed for the purpose of developing a three-dimensional gravity model of the area. Based on samples of several varieties of rock, the following information on rock density (grams per cubic centimetre) was obtained.

| **Type of Rock** | **Sample Size** | **Mean Density** | **Standard Deviation** |
|---|---|---|---|
| Late gabbro | 148 | 3.04 | 0.13 |
| Massive gabbro | 135 | 2.83 | 0.11 |
| Cumberlandite | 66 | 3.05 | 0.31 |

Let $\mu_L, \mu_M, \mu_C$ be the population density of Late gabbro, Massive gabbro and Cumberlandite respectively.

(a) Construct a 95% C.I. for $\mu_L - \mu_M$. What can be concluded from this interval?

$$\left(\overline{X}_L - \overline{X}_M\right) \pm Z_{0.025}\sqrt{\frac{S_L^2}{n_L} + \frac{S_M^2}{n_M}} = (3.04 - 2.83) \pm (1.96)\sqrt{\frac{0.13^2}{148} + \frac{0.11^2}{135}}$$
$$= 0.21 \pm 0.028 = [0.182\,,\,0.238]$$

Since all the values in this interval are positive, we have 95% confidence that the mean density of Late gabbro is higher than that of Massive gabbro for at least 0.182 g/cm$^3$.

(b) Construct a 95% C.I. for $\mu_L - \mu_C$. What can be concluded from this interval?

$$\left(\overline{X}_L - \overline{X}_C\right) \pm Z_{0.025}\sqrt{\frac{S_L^2}{n_L} + \frac{S_C^2}{n_C}} = (3.04 - 3.05) \pm (1.96)\sqrt{\frac{0.13^2}{148} + \frac{0.31^2}{66}}$$
$$= -0.01 \pm 0.078 = [-0.088\,,\,0.068]$$

Since zero is a plausible value in the interval, the data didn't provide enough evidence that the mean densities of Late gabbro and Cumberlandite are different.

(c) Write down the assumption(s) you had made in the calculations in (a) and (b).

Since all the sample sizes are large, the normal assumption is not needed. The only assumption made in the above calculations is the independence of the three random samples of the rocks.

3. (*8 marks*) From long experience with a process for manufacturing nickel-cadmium batteries it is known that the lifetime is normally distributed with a mean of 102.5 hours and a standard deviation of 9 hours. For a modified process the lifetime is claimed to have a same distribution but with increased mean lifetime of 105 hours. The research team sets up the following hypotheses to test this claim.

$H_0 : \mu = 102.5$

$H_1 : \mu = 105$ ($\mu$ is the mean lifetime of batteries produced by new process)

(a) Is this setting of hypotheses appropriate? Explain briefly.

Since we may want to show that the new process can produce batteries with longer lifetime, setting the claim as the alternative hypothesis is appropriate. For this setting, if the null hypothesis is rejected, then we can have high confidence that the new process is better and can decide to change the original manufacturing process to the new one.

(b) Construct a test with significance level 0.05 and power 0.95.

Reject $H_0$ if $\overline{X} > c$.

$$\alpha = 0.05 = \Pr\left(\overline{X} > c \mid \mu = 102.5\right) = 1 - \Phi\left(\frac{c - 102.5}{9/\sqrt{n}}\right) \Rightarrow \frac{c - 102.5}{9/\sqrt{n}} = Z_{0.05} = 1.645$$

$$\beta = 0.05 = \Pr\left(\overline{X} \le c \mid \mu = 105\right) = \Phi\left(\frac{c - 105}{9/\sqrt{n}}\right) \Rightarrow \frac{c - 105}{9/\sqrt{n}} = -Z_{0.05} = -1.645$$

Solving gives $n = 140.28.45 \approx 141$, $c = 102.5 + (1.645)\frac{9}{\sqrt{141}} = 103.75$.

Hence we should draw a sample with size 141 and reject $H_0$ if $\overline{X} > 103.75$.

(c) A sample of nickel-cadmium batteries produced by the new process was drawn. The mean lifetime was calculated with a *p-value* equal to 0.137. Based on this sample, what can be concluded?

*p-value* = 0.137 > 0.05

The null hypothesis is not rejected at significance level 0.05. Hence the data does not provide enough evidence that the new manufacturing process can produce batteries with mean lifetime 105 hours.

**<E N D >**