HW4: Normal Distribution & Z-Test

1. Suppose that, untreated, the mean length of illness for a certain type of influenza is 15 days, with a standard deviation of 2. Suppose further that the distribution of the lengths of illness is normal. Assume that these are parameters. Answer the following question in terms of these data:
   (1) What proportion of those persons who contract the disease and received no treatment are ill between 12 and 14 days? (1 pt)
   (2) What is the probability that for the 25 people in Taipei who contracted the disease and were not treated for it, the mean length of their illnesses was between 14 and 16 days? (1 pt)

   Ans:
   (1) Area below 12: Z = (12-15)/2 = -1.5
   ➔ Standard normal curve areas below $Z_{-1.5}$ = 0.0668
   Area below 14: Z = (14-15)/2 = -0.5
   ➔ Standard normal curve areas below $Z_{-0.5}$ = 0.3085
   Therefore, area between 12 and 14 = 0.3085 – 0.0668 = 0.2417

   (2) Area below 14: $Z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}} = \frac{(14-15)}{2/\sqrt{25}} = -2.5$
   ➔ Standard normal curve areas below $Z_{-2.5}$ = 0.0062
   Area below 16: $Z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}} = \frac{(16-15)}{2/\sqrt{25}} = 2.5$
   ➔ Standard normal curve areas below $Z_{2.5}$ = 0.9938
   Therefore, area between 14 and 16 = 0.9938 – 0.0062 = 0.9876

2. Suppose that entering college freshmen in a large number of colleges and universities complete a test to measure their English proficiency. Suppose further that this test has been used over a period of years and the mean score received by all entering freshmen over the last 10 years was 65.5 with a standard deviation of 10.5. The distribution of scores is approximately normal. Assume that these are parameters.
   You select a random sample of 49 entering freshmen this year at NTU and collect data on their performance on this test. Their mean score is 68.0.
   (1) Does this mean score differ significantly from the overall mean reported above? Use the 0.05 level to determine statistical significance. (Total 2 pts-- State H_0 & H_a (1 pt), Z-value (0.5 pt), Test result: Reject or Accept H_0 (0.5 pt))

(2) Precisely interpret the meaning of the statistical findings to question (1) above. What is your probability of having made Type I and Type II error. (Total 2 pts: Conclusion (1 pt), Type I error (0.5 pt), Type II error (0.5 pt)) (1 pt)

Now suppose that this sample of 49 students had a mean score of 70.0.

(3) Does this mean score differ significantly from the overall mean? Use the 0.05 level to determine statistical significance. (Total 2 pts--State H_0 & H_a (1 pt), Z-value (0.5 pt), Test result: Reject or Accept H_0 (0.5 pt))

(4) Precisely interpret the meaning of the statistical findings in question (3) above. What is your probability of having made Type I and Type II error. (Total 2 pts: Conclusion (1 pt), Type I error (0.5 pt), Type II error (0.5 pt))

Ans:

(1) H_0: $\mu_{General\ population} = \mu_{NTU}$

H_a: $\mu_{General\ population} \neq \mu_{NTU}$

$\bar{y}$: *the mean score of NTU*

μ: *the overall mean*

$$Z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}} = \frac{(68.0-65.5)}{10.5/\sqrt{49}} \cong 1.67 < Z_{0.05,two-tailed} = 1.96$$

Difference between population $\mu = 65.5$ and $\mu_{NTU} = 68.0$ is not significant at 0.05 level. ➔ Do not reject H_0

(2) The findings of Z = 1.67 are not significant at 0.05 level ($Z_{0.05,two-tailed} = 1.96$); therefore, we fail to reject the null hypothesis. The data do not support the research hypothesis, and we cannot conclude that there is a difference between the population we sampled at NTU and the population of the general freshmen in the entire colleges and universities. The probability of having made a Type I error, α=0. However, we may have made a Type II error, as the probability of a Type II error, β, is greater than zero ($\beta > 0$).

(3) H_0: $\mu_{General\ population} = \mu_{NTU}$

H_a: $\mu_{General\ population} \neq \mu_{NTU}$

$\bar{y}$: *the mean score of NTU*

μ: *the overall mean*

$$Z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}} = \frac{(70.0-65.5)}{10.5/\sqrt{49}} = 3 > Z_{0.05,two-tailed} = 1.96$$

Difference between population $\mu = 65.5$ and $\mu_{NTU} = 70.0$ is statistically significant at 0.05 level. ➔ Reject H_0

(4) The data support the research hypothesis, and we reject the null hypothesis.

We conclude that there is a difference between the NTU sample and the general college and university population. The population we sampled at NTU had a mean ($\bar{y}$ =70) greater than the mean of general college and university population ($\mu =$ 65.5). However, we may have made a Type I error. The probability of a Type I error, α, is calculated as $\alpha = (1 - 0.9987) * 2 = 0.0026$, meaning that there is a 0.26% chance that we are wrong in rejecting H_0. The probability of having made a Type II error, β=0.

3. A drug manufacturer asserts that a fixed dosage of a certain drug causes an average (mean) increase in pulse rate of 10 beats per minute with a standard deviation of 4. A group of 20 patients given the same dosage, showed a mean increase of 11 beats per minute with a variance, $S^2 = 12.46$. Use the 0.05 level to determine statistical significance. Is there evidence to contradict the drug manufacturers' claim?
   (1) State the Null Hypothesis and the Alternative Hypothesis. Show your work. (Total 2 pts: State H_0 & H_a (1 pt), Z-value (0.5 pt), Result: Reject or Accept H_0 (0.5 pt))
   (2) Make your conclusion. What is your probability of having made Type I and Type II error. (Total 2 pts: Conclusion (1 pt), Type I error (0.5 pt), Type II error (0.5 pt))

Ans:

(1) Given population: $\mu = 10, \sigma = 4$

Population we sampled: $n = 20, \bar{y} = 11, S^2 = 12.46$

H_0: $\mu_{Given\ population} = \mu_{Population\ we\ sampled}$

H_a: $\mu_{Given\ population} < \mu_{Population\ we\ sampled}$

Statement of the null hypothesis: In the population, there is no difference between the mean of the given population and the mean of the population we sampled.

Statement of the research hypothesis: In the population, the mean of population we sampled is greater than the mean of the given population.

Use 0.05 level, one-tailed test:

$$Z = \frac{\bar{y}-\mu}{\sigma/\sqrt{n}} = \frac{(11-10)}{4/\sqrt{20}} \cong 1.12 < Z_{0.05,one-tailed} = 1.645$$

Difference between population $\mu = 10$ and $\bar{y} = 11$ is not significant at 0.05 level. ➔ Do not reject H_0

(2) The findings of Z = 1.12 are not significant at 0.05 level ($Z_{0.05,one-tailed} =$

1.645); therefore, we fail to reject the null hypothesis. The data do not support the research hypothesis, and we cannot conclude that there is a difference between the population we sampled and the given population. The probability of having made a Type I error, α=0. However, we could have made a Type II error, as the probability of a Type II error, β, is greater than zero ($\beta > 0$).