MIDSEM EXAMINATION
Course: AT72.9014 – Applied Data Analytics
(Time Duration: 2 hours)

Q1. (20 points)
Consider the following hypothesis testing problem:

$$H_0 : 2\mu_1 - \mu_2 = a \qquad H_a : 2\mu_1 - \mu_2 \neq a$$

a. Develop the testing procedures for large sample & small sample cases (assume that $\sigma_1^2 = \sigma_2^2$ for small sample case)?
b. Construct the confidence intervals for $2\mu_1 - \mu_2$ in large sample & small sample cases. Can we use these confidence intervals for testing purpose in question a.?.

Q2. (20 points)
Consider the following two-tailed test: $H_0 : \mu = \mu_0; H_a : \mu \neq \mu_0$
Assuming that the population is normally distributed, and the correct value of population mean is $\mu_1 = \mu_0 - 0.5\sigma$. Given the sample size $n = 25$ and significance level $\alpha = 0.05$, let determine Type II error β.

Q3. (20 points)
a. A sample of 50 lenses used in the eyeglasses yields a sample mean thickness of 3.05 mm and a sample standard deviation 0f 0.34 mm. The desired true average thickness of such lenses is 3.20 mm. Does the data strongly suggest that the true average thickness of such lenses is something other than what is desired? Test using $\alpha = 0.05$
b. In the above problem, suppose that the experimenter had believed before collecting the data that the value of σ was approximately 0.30 mm. If the experimenter wished the probability of a type II error, i.e., β, to be 0.05 when $\mu = 3.00$, was a sample size 50 unnecessarily large? If yes, what should be the sample size?

Q4. (20 points)
Consider the following one-sided test on the mean value:

$$H_0 : \quad \mu = \mu_0$$
$$H_a : \quad \mu > \mu_0$$

Suppose that the z-test can be employed, and the sample average $\bar{x}$ is known to be greater than μ_0. An analyst does not really want to reject the null hypothesis, and hence, he takes into consideration two options:

1. Reduce the normal value of α from 0.05 to 0.01
2. Keep α at 0.05, but convert to two-sided test

a. In the second option, what will be the probability that the null hypothesis is still rejected even though it is true?
b. Which option should be selected?

Q5. (20 points) Let give answers for the following theoretical questions:
a. In which situation, the test result will not be affected by the selection of significant level α?
b. Should we select the sample before establishing the hypotheses? Why or why not?
c. Can we use the t-test instead of the z-test when the sample size is large? Why?
d. The nominal mean weight of a type of toothpaste produced by a process is μ_0. A sample has been drawn and it was found that the sample average $\bar{x}$ is greater than

μ_0. Based on the sample statistics, the process engineer believes that the process mean has been shifted. However, in order to persuade the plant manager about this belief, a test should be conducted. Which test the process engineer should select? Two-tailed, left-tailed, or right-tailed? Why?