STA219: Probability and Statistics for Engineering

Assignment 8

Note: The assignment can be answered in Chinese or English, either is fine. Please provide derivation and computation details, not just the final answer. Please submit a PDF file on BB.

1. (20 points) Let $X_1, \ldots, X_n$ be a simple random sample from the population $X \sim \mathrm{N}(\mu, 1)$. Consider the testing problem $H_0: \mu = 2 \leftrightarrow H_1: \mu = 3$. We choose the rejection region to be $\{\boldsymbol{X}: \bar{X} > 2.6\}$.
   (1) When the sample size $n = 20$, calculate the Type I and Type II error rates. (10 points)
   (2) In order to control the Type II error rate below 1%, what is the minimum sample size n? (10 points)

2. (15 points) A principal of an elementary school saw a report in the newspaper stating, "Middle school students in this city watch an average of 8 hours of television per week." She believes that the students at her school spend significantly less time on watching TV than the newspaper stated. To investigate, she randomly surveyed 100 students at her school and found that the average time spent watching TV per week was $\bar{x} = 6.5$ hours, with a sample standard deviation of $s = 2$ hours. The question is whether it can be concluded that the principal's belief is correct.
   (1) Construct a rejection region at significance level of $\alpha = 0.05$, and based on the sample observed value, decide whether to reject the statement in the newspaper. (10 points)
   (2) Alternatively, compute the p-value of the observed value of the test statistics. At significance level of $\alpha = 0.05$, is the conclusion the same as in (1)? (5 points)

3. (10 points) (Continued with Problem 3.) Suppose the principal is interested in testing whether the students at her school spend significantly less time on watching TV than the newspaper stated (i.e., less than 8 hours per week). She plans to randomly select n students and perform a significance test at the 5% significance level. If the true mean time of her students watching TV is 7.5 hours and the true standard deviation is 2 hours, what is the minimum sample size n necessary to achieve at least 0.90 power in this significance test?

4. (30 points) A news agency publishes results of a recent poll. It reports that a certain candidate has a 10% stronger support in town A than in town B because 45% of the poll participants in town A and 35% of the poll participants in twon B supported the candidate. Notice that 900 randomly selected registered voters participated in the poll in each town. Let p_A and p_B be the support rates of the candidates in town A and town B.

(1) At significance level $\alpha = 0.02$, test whether there is a significant difference between the support rates of the candidate in town A and town B. (5 points)

(2) Calculate the 98% confidence intervals for each of p_A, p_B, and $p_A - p_B$. (10 points)

(3) Can we make a decision about whether the the support rates of the candidate in town A and town B are significantly different or not at level $\alpha = 0.02$ based on whether the 98% confidence intervals of p_A and p_B overlap or not? State your explanation. (10 points)

(4) Can we make a decision about whether the the support rates of the candidate in town A and town B are significantly different or not at level $\alpha = 0.02$ based on whether the 98% confidence intervals of $p_A - p_B$ contain 0 or not? State your explanation. (5 points)

5. (25 points) Suppose that $X_1, \dots, X_n$ is a simple random sample from population $X \sim \mathrm{N}(\mu_1, 4^2)$, $Y_1, \dots, Y_m$ is a simple random sample from population $Y \sim \mathrm{N}(\mu_2, 4^2)$. The two samples are independent. μ_1, μ_2 are unknown population means, and we would like to test $H_0: \mu_1 = \mu_2 \leftrightarrow H_1: \mu_1 > \mu_2$ at significance level $\alpha = 0.05$.

(1) Choose a proper test statistic T and determine its distribution under H_0, then write down the rejection region based on the test statistic. (5 points)

(2) Suppose that the underlying true mean difference between the two populations is $\mu_1 - \mu_2 = 2$, calculate the Type II error rate of the test in (1) if $n = 10$ and $m = 11$. (10 points)

(3) Continued with (2), compute the minimum total sample size $N = n + m$ needed to achieve at least 0.9 statistical power in the test. (10 points)