Homework 3

Statistical Inference, Fall 1401

1- The average house price in New York based on the random sample of 81 houses is \$800,000. Assume that population standard deviation σ equals \$90,000.

a) Compute a 98% confidence interval for the population mean μ.
b) Compute a 95% confidence interval for the population mean μ.
c) Compute a 90% confidence interval for the population mean μ.
d) Compute a 50% confidence interval for the population mean μ.
e) Write a one-sentence interpretation of the confidence interval.
f) Express the intuition behind different confidence intervals and obtained results.
g) How many such houses should be sampled in order to obtain a 99% confidence interval with a margin of error less than or equal to \$5,000?
h) How would the required sample size change if we want to decrease the margin of the error to \$2,500?

2- Based on research, on average, high school teens spend almost seven hours each weekday on educational activities. You think that this is underestimated and want to collect your own sample for a hypothesis test. You randomly sample 70 students from your dorm and find that on average they spend 10 hours each weekday on education activity. One of your friends wants to help you and comes up with the following hypothesis. Indicate any error you see.

$$H_0: x < 7\ hours$$
$$H_A: x > 10\ hours$$

3- A mayor claims that in his small city, which is renowned for its music school, the average child takes at least 5 years of piano lessons. We have a random sample of 20 children from the city, with a mean of 4.6 years of piano lessons and a standard deviation of 2.2 years.
a) Evaluate his claim using a hypothesis test (assume $\alpha = 0.05$).
b) Construct a 95% confidence interval for the number of years students in this city take piano lessons and interpret it in the context of the data.
c) Do your results from the hypothesis test and the confidence interval agree? Explain your reasoning.

4- A study suggests that the normal body temperature in degrees Fahrenheit is 98.6^o . A doctor believes that this is an underestimate and decides to collect her own sample for a hypothesis test of 52 randomly chosen healthy adults. The body temperature of samples is measured with the following summary of the data:

$$n = 52, \bar{x} = 98.2846^o, s = 0.6824$$

a) Are the necessary conditions satisfied? Explain.
b) Write the null and alternative hypotheses.
c) Find a 98% confidence interval for the mean body temperature and make decision based on this confidence interval.
d) Give a two-side hypothesis test for a mean body temperature of 98.6^oFahrenheit and use the information above to evaluate a test with significance level $\alpha = 0.02$.

Homework 3

Statistical Inference, Fall 1401

5- Suppose that $X \sim Binomial$ (100,p). Consider the following test:

$$H_0: p = 0.5$$
$$H_1: p \neq 0.5$$

If $| X - 50 | > 10$, H_0 is rejected. Answer the following questions:

a. What is α?
b. Graph the power as a function of p.

6- A researcher designs a study to test the hypotheses H_0: $\mu \geq 28$ versus H_a: $\mu < 28$. A random sample of 50 measurements from the population of interest yields $\bar{y} = 25.9$ and $s = 5.6$.

a. Using $\alpha = .05$, what conclusions can you make about the hypotheses based on the sample information?
b. Calculate the probability of making a Type II error if the actual value of μ is at most 27.
c. Could you have possibly made a Type II error in your decision in part (a)? Explain your answer.

7- Refer to the previous exercise (Ex. 6). Sketch the power curve for rejecting H_0: $\mu \geq 28$ by determining (μ_a) for the following values of μ: 22, 23, 24, 25, 26, and 27.

a. Interpret the power values displayed in your graph.
b. Suppose we keep $n = 50$ but change to $\alpha = .01$. Without actually recalculating the values for (μ_a), sketch on the same graph as your original power curve, the new power curve for $n = 50$ and $\alpha = .01$.
c. Suppose we keep $\alpha = .05$ but change to $n = 20$. Without actually recalculating the values for $PWR(\mu_a)$, sketch on the same graph as your original power curve the new power curve for $n = 20$ and $\alpha = .05$.

8- (R) For this question, use the dataset "Galton" placed next to the exercise. This dataset includes the height of students and their fathers. (Assume that the standard deviation of the society is unknown)

Use the height of the children:

a. Generate 20000 samples of size 60 using sampling. Then calculate the 97% confidence interval for each of these samples. What percentage of the intervals include the real mean of the society?
b. Repeat the above test for 10000 samples with size 10. This time, calculate the 90% confidence interval. What is your conclusion about these two tests?

Now use the height of the father:

c. Separate a sample with the size of 70 data from the mentioned data. Then check the following assumption and calculate the power.
H_0: μ=60
H_1: $\mu \neq$60
d. Now, isolate a sample of 10 data and repeat the above test and report the results.
e. What do you get from comparing the above results?