HSCI 50 LAB 4: Confidence Intervals and Hypothesis Testing

INSTRUCTIONS

The data set that we will be using for the entire course is resampled data from the MIT COVID-19 Beliefs, Behaviors & Norms Survey (https://covidsurvey.mit.edu/api.html). This is a multi-country, online survey that examined different COVID-19 perceptions across time, from July 6, 2020 to March 28, 2021. We will be using data from the Philippines aged 20 to 60.

In this lab, you will practice calculating confidence intervals and conducting hypothesis tests for one-sample proportions and means.

You have many options to submit this worksheet. Either you work on this by hand and scan/take a clear photo of your submission and save as PDF, or type your responses in a Word processor or PowerPoint presentation. You do not have to copy the questions again, but please number them accordingly.

Part A. One Sample Proportions

The Philippine government has a target vaccination rate of 70% of the population. This means that the government can only afford to have at most 30% of the population who are hesitant of taking the COVID-19 vaccine. The government decides to conduct a survey of 1,145 individuals to check on the situation.

1. Show that the sampling distribution of the sample proportion $\hat{p}$ approximates a normal distribution.

2. What is the mean and the variance of the sampling distribution of the sample proportion $\hat{p}$?

Now, the government received the results of the survey and the results are summarized below. Note that we are using the same table as in Lab 3.

Table 1: COVID-19 vaccine hesitancy in the Philippines, March 14-28, 2021

Vaccine hesitant	Count	Percent
No	612	53.4
Yes	533	46.6

3. Now, we want to test whether the sample proportion $\hat{p}$ is different from the target of 30%. What is the null and alternate hypothesis? Calculate the *z*-statistic and the corresponding two-sided *p*-value. What is your conclusion based on the *p*-value and α = 0.05? You may use the University of Iowa Normal Distribution Applet to calculate the *p*-value (https://homepage.divms.uiowa.edu/~mbognar/applets/normal.html).

4. Now, we want to be more specific with our hypothesis test and check whether the sample proportion $\hat{p}$ is *lower* than the target of 30%. What are your new hypotheses, the *z*-statistic, and corresponding one-sided *p*-value? Does your conclusion change?

5. Now, estimate the 95% confidence interval of the sample proportion. Compare the confidence interval to the government target of 30%. Is your conclusion consistent with your answers in Questions 3 and 4?

PART B. One Sample Means

Let us revisit the vaccine norms data from Lab 3. Remember that the survey had results of mean = 60.7 and standard deviation = 12.5. In Lab 3, we assumed this was the population mean and population standard deviation, respectively. This time around, we will assume them to be the sample mean and sample standard deviation on a total sample of 1,145 respondents sampled nationally using simple random sampling.

This time around, the government has set a target of a mean of 70 for vaccine norms. There is no target standard deviation.

1. If the objective is to estimate the current population mean vaccine norm, does it matter that the sample data is normally distributed?

2. Estimate the variance of the sampling distribution of the sample mean.

3. Now, we want to test whether the sample mean $\bar{x}$ is different from the target of 70. What is the null and alternate hypothesis? Calculate the *t*-statistic and the corresponding two-sided *p*-value. What is your conclusion based on the *p*-value and α = 0.05? You may use the University of Iowa Student's t Distribution Applet to calculate the *p*-value (https://homepage.divms.uiowa.edu/~mbognar/applets/t.html).

4. Now, we want to be more specific with our hypothesis test and check whether the sample mean $\bar{x}$ is *higher* than the target of 70. What are your new hypotheses, the *t*-statistic, and corresponding one-sided *p*-value? Does your conclusion change?

5. Now, estimate the 95% confidence interval of the sample mean. Compare the confidence interval to the government target of 70. Is your conclusion consistent with your answers in Questions 3 and 4?

END OF LAB