BIOS: 4120 Lab 9

March 31 2020

Prior to break, we covered one-sample categorical data, and today we discussed one-sample continuous data. In this lab, we will be conducting hypothesis tests and creating confidence intervals for both categorical and continuous data.

Note: A hat on a Greek letter indicates an estimator, so for example, when you see $\hat{\mu}$, this is the same thing as $\bar{x}$.

Normal Distribution and T Distribution Key Formulas:

1. $Z = \frac{x-\mu}{\sigma} \sim N(0,1)$,

   where μ, and σ are known population parameters and the distribution of X is normal.

2. $Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$,

   where μ, and σ are known population parameters and the distribution of $\bar{X}$ is normal.

   Note: The underlying distribution of X does not affect the distribution of $\bar{X}$

3. $T = \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t_{n-1}$,

   where μ is a known parameter and s is the sample standard deviation of observed $x_i's$.

   Note: As n get larger, the t-distribution will be equal to the normal distribution.

Example 1: Z-test for categorical data

Suppose the incidence rate of myocardial infarction per year was 0.005 among males age 45-54 in 1970. For 1 year starting in 1980, 5000 males age 45-54 were followed, and 15 new myocardial infarction cases were observed.

From the central limit theorem, we know that the sample proportion approximately follows a normal distribution (if the sample size is reasonably large), so we can perform a z-test on this data.

Conduct a hypothesis test to determine if true myocardial infarction rate changed from 1970 to 1980. *How would you interpret the result?*

Creating a confidence interval (z)

Now we want to create a 95% confidence interval for π. *Interpret the interval.*

Now that we've seen an example of categorical data, let's look at a continuous data example.

Example 2: T-test Continuous

The distribution of weights for the population of males in the United States is approximately normal. We believe the mean $\mu = 172.2$. We conduct an experiment with a sample size of 50, and we find our sample mean to be 180 and the sample standard deviation to be 30. Conduct a hypothesis test to determine if the true mean is 172.2 based on our data. *How would you interpret the result?*

Practice Problem 1:

Suppose that the average IQ is 95. Perform a test to see if the children in the lead-IQ dataset have an average IQ. Also, create a 95% confidence interval for the mean IQ based on this data.

Practice Problem 2:

A patient recently diagnosed with Alzheimer's disease takes a cognitive abilities test. The mean of this test is $\mu = 52$ and the variance was $\sigma^2 = 25$. Assume the cognitive abilities test scores are normally distributed. Find the answers to the following questions with the Z distribution table, your calculators, or in R. Remember the Z table gives you the left-tailed probability.

a) What percent of individuals scored between a 47 and a 56?

b) Suppose we have a sample of 9 individuals. Calculate the probability that the sample mean test score is greater than 60.

c) Patients can be considered for an alternative treatment if they score below a 43 on this test. What percent of patients can be considered for this treatment?

d) Find the test score where 27.1% of patients lie above.

e) What is the probability that at least 2 patients of 25 sampled Alzheimer's patients will be considered for the alternative treatment?

Practice Problem 3:

Wilson's orchard's pumpkins' weights are known to follow a normal distribution with population mean $\mu = 18lbs.$ and variance $\sigma^2 = 16lbs.$ Each year Wilson's orchard randomly selects 4 pumpkins and measures the mean weight of the pumpkins.

a) What distribution do the sample means follow?

b) Using this distribution, calculate the probability that this year's sample mean weight is less than 16 lbs.

c) What is the probability that this year's sample mean weight is greater than 21 lbs?

d) What is the probability that at least 2 of the next 5 years' sample means are between 14 and 20 lbs?