Mod 21 Estimating a Population Mean

Inference about a population proportion usually arises when we study categorical variables. We learned how to construct and interpret confidence intervals for an unknown parameter *p*. To estimate a population mean, we have to record values of a quantitative variable for a sample of individuals. It makes sense to try to estimate the mean amount of sleep that students at a large high school got last night but not their mean eye color! In this section we'll examine confidence intervals for a population mean μ.

There are two types of distributions we can use to estimate the mean, either the normal distribution or the t-distribution.

We use the normal distribution when the population standard deviation is known (this is very rare) or our sample size n > 30.

1) A college admissions director wishes to estimate the mean age of all students currently enrolled. In a random sample of 20 students, the mean age is found to be 22.9 years. From past studies, the standard deviation is known to be 1.5 years and the population is normally distributed. Construct a 90% confidence interval of the population mean age.

   <u>Step 1</u>: Identify the ***parameter*** you want to estimate and the ***confidence level***.

   <u>Step 2</u>: Check ***conditions***.

   <u>Step 3</u>: If the conditions are met, perform the ***calculations***.

   <u>Step 4</u>: State the result (the confidence interval) and ***interpret*** your result in the ***context*** of the problem.

In many real-life situations, the population standard deviation is unknown. Moreover, because of various constraints such as time and cost, it is often not practical to collect sample sizes more than 30. We can use a t-distribution in these situations as long as the random variable is normally distributed (or approximately normally distributed).

2) You randomly select 16 restaurants and measure the temperature of the coffee sold at each. The sample mean temperature is 162°F with a sample standard deviation of 10°F. Find the 95% confidence interval for the mean temperature. Assume the temperatures are approximately normally distributed.

Step 1: Identify the ***parameter*** you want to estimate and the ***confidence level***.

Step 2: Check ***conditions***.

Step 3: If the conditions are met, perform the ***calculations***.

Step 4: State the result (the confidence interval) and ***interpret*** your result in the ***context*** of the problem.

DETERMINING WHICH DISTRIBUTION TO USE: NORMAL DISTRIBUTION OR T-DISTRIBUTION

Use **the normal distribution (Z test statistic)** if both of the following conditions are true.

a) The population standard deviation, σ, is known
   AND
b) Either the distribution of sample means is approximately normal OR $n \geq 30$

Use **the T-distribution (T test statistic)** if both of the following conditions are true.

a) The population standard deviation, σ, is NOT known
   AND
b) Either the distribution of sample means is approximately normal OR $n \geq 30$

So … inquiring minds want to know … which distribution do we use if $n < 30$ and we don't know whether the distribution of sampling means is approximately normal?

MOD 21 – LET'S PRACTICE (PART 1)

In exercises 1-6, use a normal distribution or a *t*-distribution to construct a 95% confidence interval for the population mean. Justify your decision. If neither distribution can be used, explain why not.

1) In a random sample of 70 bolts, the mean length was 1.25 inches and the standard deviation was 0.01 inch.

   Step 1: Identify the ***parameter*** you want to estimate and the ***confidence level***.

   Step 2: Check ***conditions***.

   Step 3: If the conditions are met, perform the ***calculations***.

   Step 4: State the result (the confidence interval) and ***interpret*** your result in the ***context*** of the problem.

2) You took a random sample of 12 two-slice toasters and found the mean price was \$61.12 and the standard deviation was \$24.62. Assume the prices are normally distributed.

Step 1: Identify the ***parameter*** you want to estimate and the ***confidence level***.

Step 2: Check ***conditions***.

Step 3: If the conditions are met, perform the ***calculations***.

Step 4: State the result (the confidence interval) and ***interpret*** your result in the ***context*** of the problem.

3) You take a random survey of 25 sports cars and record the miles per gallon for each. The data are listed below assume the miles per gallon are normally distributed.

24 24 27 20 26 23 18 29 24 22 22 27 26
20 28 30 23 24 19 22 24 26 23 24 25

<u>Step 1</u>: Identify the ***parameter*** you want to estimate and the ***confidence level***.

<u>Step 2</u>: Check ***conditions***.

<u>Step 3</u>: If the conditions are met, perform the ***calculations***.

<u>Step 4</u>: State the result (the confidence interval) and ***interpret*** your result in the ***context*** of the problem.

4) In a recent year, the standard deviation of ACT scores for all students was 4.7. You take a random survey of 20 students and determine the ACT score of each. The scores are listed below. Assume the test scores are normally distributed.

26	22	23	12	19	25	23	21	25	10
17	26	23	24	20	14	21	23	20	22

Step 1: Identify the ***parameter*** you want to estimate and the ***confidence level***.

Step 2: Check ***conditions***.

Step 3: If the conditions are met, perform the ***calculations***.

Step 4: State the result (the confidence interval) and ***interpret*** your result in the ***context*** of the problem.

For numbers 5 & 6, continue using the four-step process.

5) In a random sample of 19 patients at a hospital's minor emergency department, the mean waiting time (in minutes) before seeing a medical professional was 23 minutes and the standard deviation was 11 minutes. Assume the waiting times are not normally distributed.

6) In a random sample of 17 shoppers at a grocery store, the mean amount spent was \$28.13 and the standard deviation was \$12.05. Assume the amounts spent are normally distributed.

MOD 21 – LET'S PRACTICE (PART 2 OPTIONAL)

Size Matters: In the Module 17 activity, we used a formula to calculate the minimum sample size for estimating a population ***proportion***. Re-write that formula in the space below.

To complete this activity, we need the formula for calculating the minimum sample size for estimating a population ***mean***. Here it is.

Given a confidence level *C* and a margin of error *E*, the minimum sample size, *n*, needed to estimate the population mean is $n = \left(\frac{Z_c \cdot \sigma}{E}\right)^2$.

What if you do NOT know the population standard deviation, σ? Is there something else you could use in its place?

Try using the formula for the following problems:

1) A cheese processing company wants to estimate the mean cholesterol of all one-ounce servings of cheese. The estimate must be within 0.5 mg of the population mean.

   a) Determine the minimum required sample size to construct a 95% confidence interval for a population mean. Assume the population standard deviation is 2.8 milligrams.

   b) Repeat part a using a 99% confidence interval.

   c) Which level of confidence requires a larger sample size? Explain.

2) An admissions director wants to estimate the mean age of all students enrolled at a college. The estimate must be within 1 year of the population mean. Assume the population of ages is normally distributed.

   a) Determine the minimum required sample size to construct a 90% confidence interval for a population mean. Assume the population standard deviation is 1.2 years.

   b) Repeat part a using a 99% confidence interval.

   c) Which level of confidence requires a larger sample size? Explain.