How Far Am I Off: Estimating σ

ID: 12441

Time required
15 minutes

Activity Overview

In this activity, students will calculate a confidence interval using the chi-square distribution to estimate a population variance. The homework problems have students estimate the mean and standard deviation of the population given a sample.

Topic: Statistical Inference

- *Chi-square distribution*
- *Estimating population standard deviation*

Teacher Preparation and Notes

- *The complete data sets for the homework problems are included. Random samples were used to produce the mean and standard deviation. Other random samples could be found with other sample means and standard deviations.*
- ***To download the student worksheet, go to education.ti.com/exchange and enter "12441" in the quick search box.***

Associated Materials

- *StatWeek25_Variance_worksheet_TI84.doc*
- *STATE and SPORT lists*
- *INVERSX2 program*

Suggested Related Activities

To download any TI-84 Plus family activity listed, go to education.ti.com/exchange and enter the number in the quick search box.

- *Population Mean: σ unknown (TI-84 Plus) — 12387*
- *Chi-Square Distributions (TI-84 Plus) — 9737*
- *TI Using the Chi Square Significance Test (TI-84 Plus) — 3502*

Problem 1 – Assumptions

Explain to students that when they previously learned to estimate a population mean, it was possible because of the Central Limit Theorem (sample means follow a normal distribution).

Students will now estimate a population variance. However, the standard deviation follows a chi-square distribution. Review with students that the chi-square (χ^2) distribution is represented by the formula $\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ where n = sample size, s = standard deviation for the sample, and σ = standard deviation for the population.

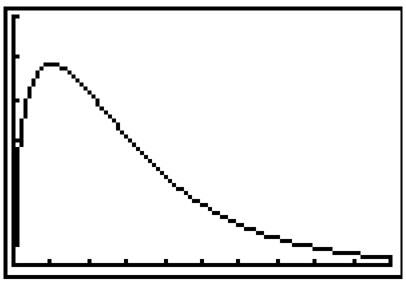
Answers

1. A 2. B 3. C 4. B 5. A

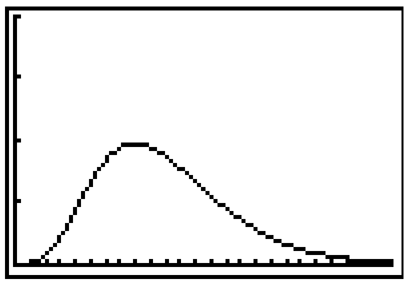
This distribution is not symmetric; it is skewed to the right. Students will graphs three chi-square distributions with 3, 10, and 25 degrees of freedom (n – 1).

Note: The graphing windows have different scales.

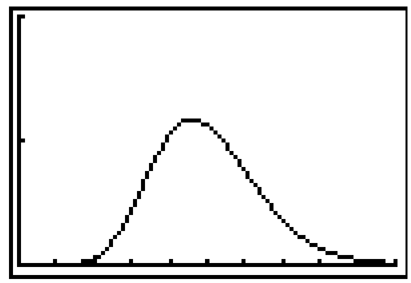
Ask students: *What do you notice about the shape of the distribution as the degrees of freedom increase?* They should see that the graph becomes more symmetric.



χ^2Pdf(X,3)



χ^2Pdf(X,10)

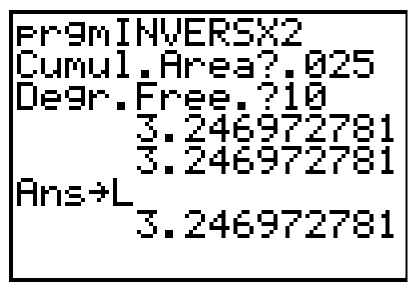


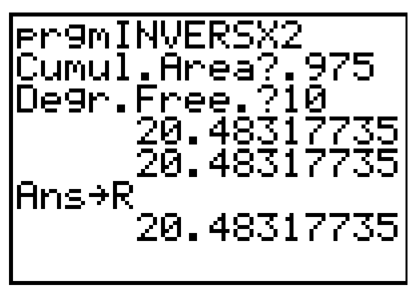
χ^2Pdf(X,25)

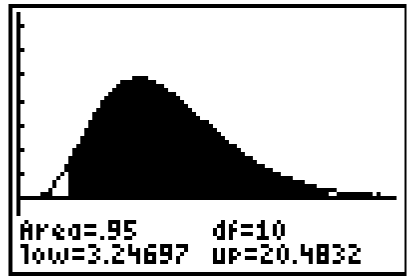
Discuss with students that confidence intervals which estimate a population variance need two different critical values as shown at the right. (This is different than with a *t*- or *z*-distribution where they are symmetric.)

Students will determine the critical values for a sample size of 11 at 95% confidence. They will need to use a Chi-square distribution chart or the **INVERSX2** program to find the left and right chi-square values.

They should store the left and right critical values as **L** and **R**.







Students will verify that the area under the curve between the two critical values is 95% or 0.95. This is to be done using the **Shade**χ^2 command (2nd [DISTR] and move to the DRAW menu). They will need to enter the lower and upper critical values and the degrees of freedom or **Shade**χ^2 **(L, R, 10)**.

Note: Students need to have first graphed χ^2**pdf(X,10)**.

To calculate the area under the curve. Students will once again see that the curve is not symmetric.

Problem 2 – Estimating the Interval

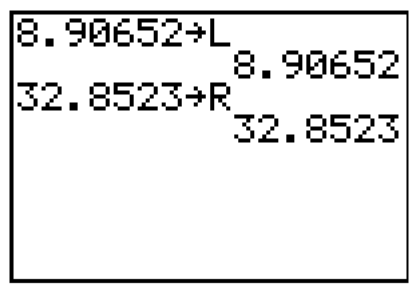
The formula for the confidence interval is given on the worksheet. Discuss with students or have them discuss with a partner how to get from the formula for the χ^2 distribution to the confidence interval formula. (It is solved for σ).

Explain to students that when the square root of both sides of an equation in Algebra is taken, one side is $\pm\sqrt{}$. In this case, the positive and negative sign is the left and right critical values.

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \Rightarrow \qquad \sigma^2 = \frac{(n-1)s^2}{\chi^2} \Rightarrow \qquad \sqrt{\frac{(n-1)s^2}{\chi^2_R}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi^2_L}}$$

Ask students: *Why are the denominators of the fractions seemingly reversed?* They should see that if not reversed, the fraction on the left would be larger than the fraction on the right.

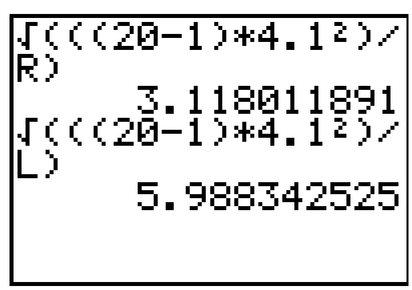
$$\left(\frac{1}{\text{large}} = \text{small} \quad \frac{1}{\text{small}} = \text{large}\right)$$



Students are given the details of a random sample. A random sample of 20 cereal boxes has a standard deviation of 4.1 grams of sugar per box.

To find the 95% confidence interval students will:

- Find χ^2_L and χ^2_R. (8.9 and 32.85) by using a chi-square distribution chart or the program. They may store these values as **L** and **R**.



- Calculate the endpoints of the interval using the formula above.

  Note: Students can press 2nd [ENTRY] to recall the entry for the right endpoint and edit it for the left endpoint.

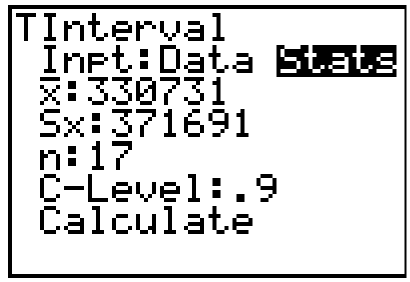
Sample Response: I am 95% confident that the population standard deviation of all the cereal boxes is between 3.12 and 5.99.

Homework

For both homework problems, students will need to use what they have previously learned to calculate a confidence interval to estimate the population mean and then also practice calculating a confidence interval to estimate the population standard deviation. Encourage students to write their findings in complete sentences.

The data for the entire population is given for each problem. Students can use the **1-Var Stats** command to find the actual mean and standard deviation. Press STAT and arrow to the CALC menu to select the command.

Students will need to transfer the lists STATE and SPORT to their calculator prior to completing the homework assigning.

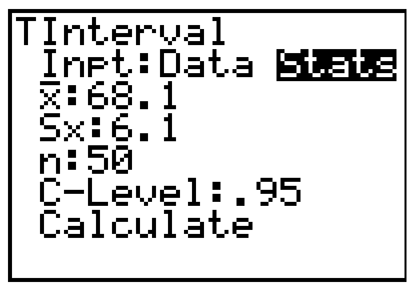


1. **a.** Mean: (Use *t*-distribution since σ is unknown.)
(173343, 488119)
Sample Response: I am 95% confident that the population mean is between 173,343 and 488,119.

b. Standard deviation: (289931, 526915)
Sample Response: I am 95% confident that the population standard deviation is between 289,931 and 526,915.

Actual: μ = 240,253; σ = 277,200

Students can access the **TInterval** command by pressing STAT and arrowing to the TESTS menu.



2. **a.** Mean: (Use *t*-distribution since σ is unknown.) (66.37, 69.83)
Sample Response: I am 95% confident that the population mean is between 66.37 and 69.83.

b. Standard deviation: (56.89, 84.86)
Sample Response: I am 95% confident that the population standard deviation is between 56.89 and 84.86.

Actual: μ = 68.7; σ = 6.82

Extension (or review):

A. The data sets could be used to review normal distributions. Students could graph the data set, find means, and standard deviations. They can verify that the data set is normal using percentages and normal probability plots.

B. Students could use the data sets to find other random samples. Have each group create a different random sample and compare the standard deviations and means. How close are they to each other? Why is there a difference? This stresses the idea that different random samples can have different results.

