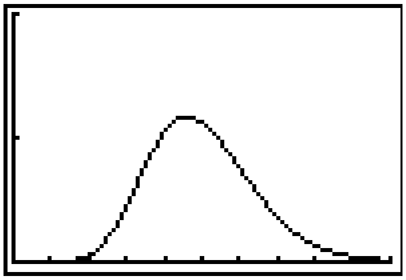
Problem 1 – Assumptions

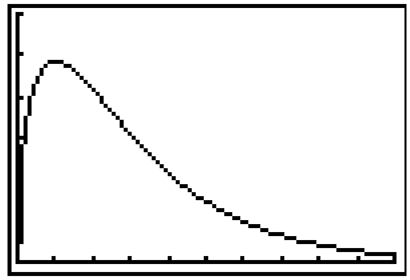
Determine whether each graph is

A) Symmetric **B)** Skewed Right **C)** Skewed Left **D)** Uniform

1.



2.



3. With a symmetric distribution, like the *t*-distribution, the critical value for 50% is

A) –1 **B)** –0.5 **C)** 0 **D)** 0.5 **E)** 1

4. For a symmetric distribution, like the *t*-distribution, the critical values for what two areas are negative?

A) 1 and –1

B) 0.025 and 0.975

C) 0.5 and –0.5

D) 0.3 and 0.97

E) It's impossible to know.

5. For a non-symmetric distribution, the lack of symmetry makes two critical values an equal distance from the mean to be different.

A) True **B**) False

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.

This distribution will be used to estimate the standard deviation for the population.

One at a time, graph the three chi-square (χ^2) distributions shown at the right. Each distribution has a different degree of freedom. To enter the χ^2 function, press 2nd [DISTR].

$Y_1 = \chi^2\text{Pdf}(X,3)$

$Y_1 = \chi^2\text{Pdf}(X,10)$

$Y_1 = \chi^2\text{Pdf}(X,25)$

6. What do you notice about the shape of the distribution as the degrees of freedom increase?

When constructing a confidence interval for the variance, it is necessary to find two critical values due to the lack of symmetry in the chi-square (χ^2) distribution.

A 95% confidence interval has a low percentage of 2.5% and a high percentage (area) of 97.5%. The values on the *x*-axis (the critical values) that correspond with these percentages can be found using a chi-squared distribution chart or the **INVERSX2** program.

- The critical value on the left is χ_L^2. It uses the low percentage of area.
- The critical value of the right is χ_R^2. It uses the high percentage of area.

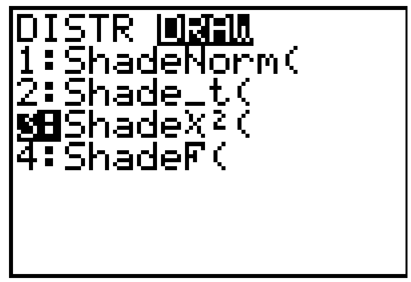
7. Find the χ_L^2 and χ_R^2 values for a 95% confidence interval with 10 degrees of freedom. Store χ_L^2 as **L** and χ_R^2 as **R**.

Note: Use the [STO▸] key to store a value.

Verify that the area between these two values is 95% of the area under the entire curve with the **Shade**χ^2 command ([2nd] [DISTR] and arrow to the DRAW menu).

First graph **Y1 =** χ^2**pdf(X,10)**.

Then on the home screen enter **Shade**χ^2 **(L, R, 10).**



Problem 2 – Estimating the Interval

Goal: Estimate the true variance (σ) of the population from a sample.

Confidence Interval
$\sqrt{\dfrac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\dfrac{(n-1)s^2}{\chi_L^2}}$

8. Why is the right χ^2 value used in the left bound of the interval and vice versa?

A random sample of 20 cereal boxes has a mean of 7.45 grams of sugar and a standard deviation of 4.1 grams of sugar per box. Assume that the population is normally distributed. Find a 95% confidence interval for the standard deviation for the population.

9. Find χ_R^2 and χ_L^2 and store as **R** and **L**.

10. Calculate the endpoints of the interval. (Hint: Use the formula above.)

11. Interpret the interval in as it applies to the problem.

Homework

Import the included data sets to a list and use them to complete the homework problems. Use the STATE list for Problem 1and the SPORT list for Problem 2.

1. A random sample of the population of 17 U.S. state capitals has a mean of 330,731 and a standard deviation of 371,691. Assume that the population is normally distributed.

a. Find 90% confidence intervals for the mean and standard deviation of all U.S. state capitals. Interpret the intervals.

b. The list STATE gives the complete set of US capital data. Calculate the actual mean and standard deviation. Do the actual values fall within the calculated confidence intervals? Why do you think this happens?

2. There are 343 teams in Division I College Men's Basketball. A random sample of 50 teams has a mean of 68.1 and a standard deviation of 6.1 for the average number of points scored per game. Assume that the population is normally distributed.

a. Find a 95% confidence interval for the mean and standard deviation of the population. Interpret the intervals.

b. The list SPORT gives the complete set of basketball data. Calculate the actual mean and standard deviation. Do the actual values fall within the calculated confidence intervals? Why do you think this happens?