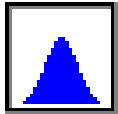# CHAPTER 5

Visualizing Continuous Distributions

CONCEPTS
- Continuous Distribution, Normal Distribution, Chi-Square Distribution, Student's t Distribution, F Distribution, Normal Approximations

OBJECTIVES
- Recognize common continuous distributions and their cumulatives

- Identify the parameters of common continuous distributions and how they affect the distribution

- Recognize shape measures for common continuous distributions

- Understand when common continuous distributions can be approximated by a normal distribution

- Understand the relation between a value of a distribution and the area in the distribution's tail

Overview of Concepts

Continuous distributions are used to describe variables that can take on an infinite number of values or at least a very large number of values. Unlike a discrete distribution, we cannot easily list the characteristics of a data-generating process (e.g., binomial) that will guarantee a particular continuous distribution. However, there are variables that we believe can be well approximated by a continuous distribution. For example, the height of seven-year old girls (in the U.S.) appears to be normally distributed with a mean of 120.5 cm and a standard deviation of 4.3 cm. We also know that process errors often are normally distributed. For example, if a door is supposed to be 32 inches wide, its actual width might be normally distributed with a mean of 32 inches and a standard deviation of 0.1 inches. But continuous distributions are primarily used to describe the distribution of sample statistics (e.g., the sample mean or the sample variance) or some transformation of these sample statistics.

The best-known distribution is the **normal distribution**. It describes the distribution of the sample mean if the statistic is based upon a large number of observations. This is proven by the Central Limit Theorem, which you will learn later in this book. This distribution has the classic symmetric bell-shape that you have probably seen elsewhere. It is defined by two parameters: its mean μ (mu) and its standard deviation σ (sigma).

If a normal distribution is sampled, the sample variance multiplied by its degrees of freedom is distributed as a **chi-square distribution**. This distribution is defined by one parameter called degrees of freedom. Unlike the normal distribution, the chi-square distribution is always positively skewed, with the degree of skewness decreasing with the degrees of freedom. For very large degrees of freedom, the chi-square distribution becomes normally distributed.

While the normal and chi-square distributions describe the behavior of the sample mean and variance, two other well-known distributions describe transformations of the sample mean and variance. The **Student's t distribution** describes the ratio of the sample mean divided by the square root of the sample variance divided by the sample size. This distribution has one parameter called the degrees of freedom. Like the normal distribution the t distribution is symmetric. For very large degrees of freedom increase, the t distribution becomes normally distributed.

The **F distribution** describes the ratio of two sample variances from independent samples. It is defined by two parameters, the degrees of freedom in the numerator and the degrees of freedom in the denominator. As both the numerator and denominator degrees of freedom get very large, the F distribution becomes normally distributed.

Under certain conditions, the chi-square, Student's t, and F distributions can be approximated by a normal distribution. The **normal approximation** improves in every case as the number of degrees of freedom increase.

Illustration of Concepts

Continuous distributions have been derived by mathematicians and statisticians in order to solve real-world problems. Abraham De Moivre (1667–1754) was investigating an approximation to the binomial distribution when he derived an expression that would later be known as the normal curve (i.e., the first occurrence of the **normal approximation**). His paper of 1738 is the first appearance in the English language of such an expression. Karl Gauss (1777–1855) was investigating the mathematics of planetary orbits when he derived what we now know as the **normal distribution** function (1809). His investigation was motivated by a need to explain why planetary orbits did not precisely correspond to the mathematical expressions that he derived. This need caused him to turn his attention to the distribution of errors. Independently, Pierre Laplace (1749–1827) was studying the anomalies in the orbits of Jupiter and Saturn when he derived an extension of De Moivre's limit theorem: "Any sum or mean will, if the number of terms is large, be approximately normally distributed." This is a simplified version of what we now call the Central Limit Theorem (1810).

The first derivation of a **chi-square distribution** was by Ernst Abbé in a paper written for his appointment as professor in the Faculty of Philosophy (1863). He derived the distribution while studying the distribution of the sum of squared errors. These were the same errors studied by Gauss, the difference being that Gauss studied their distribution while Abbé squared the errors and studied the distribution of their sum. Independently, Ludwig Boltzmann studied the distribution of kinetic energy of molecules (1878) when he derived the chi-square distribution for two and three degrees of freedom. However, it wasn't until Karl Pearson (1857–1936) provided numerous examples of how the chi-square statistic could be used (e.g., to study the heights of schoolgirls, barometric pressure, and pauperism) that the distribution became useful to most statisticians.

William Gosset (1876–1937) became a brewer for Messrs. Arthur Guinness Sons and Co., Ltd. in 1899. He was hired to study how the quality of hops and barley as well as the production process itself affected the quality of the beer produced. To study these relations he examined the ratio of $\overline{X}$ to s. However, because each experiment took a day to complete, very small samples were the norm. He noticed that comparing the ratio to the normal distribution seemed to produce incorrect results. Based upon this need, he developed a different approach that corrected these problems, which he published (1908) under the pseudonym "Student." Later, Harold Hotelling would derive the t distribution that "Student" first developed. It has become known as **Student's t distribution** in honor of William Gosset.

Sir Ronald Fisher (1890–1962) was one of the most important statisticians of all time. In one of his studies he examined differences in crop yields. However, to study these differences he had to develop a new distribution based upon the ratio of independent chi-square distributions. His variance ratio test later became known as an **F distribution,** in his honor.

Sources: M.G. Kendall and R.L. Plackett, *Studies in the History of Statistics and Probability, Volume II* (New York: Macmillan, 1977); E.S. Pearson and M.G. Kendall, *Studies in the History of Statistics and Probability, Volume I* (New York: Macmillan, 1970); and Stephen M. Stigler, *The History of Statistics: The Measurement of Uncertainty before 1900* (Cambridge, MA: Harvard University Press, 1986).

Orientation to Basic Features

This module illustrates the normal, chi-square, Student's t, and F probability distributions and their cumulatives. The distributions can also be compared visually to a normal distribution.

1. **Opening Screen**
   Start the module by clicking on the module's icon, title, or chapter number in the *Visual Statistics* menu and pressing the Run Module button. When the module is loaded, you will be on the introduction page of the Notebook. Read the questions this module covers and then click the Concepts tab to see the concepts that you will learn. Click on the Scenarios tab. Select Normal Distribution from the list of choices. Select a scenario and press OK. The upper left of the screen depicts the normal distribution. The module's Control Panel appears on the right. On the bottom left is the table of shape measures. On the bottom right are spin buttons that show the areas in each tail of the distribution. Other features are controlled from the menu bar at the top of the screen. A flashing Update button will indicate when you have changed one or more control settings.
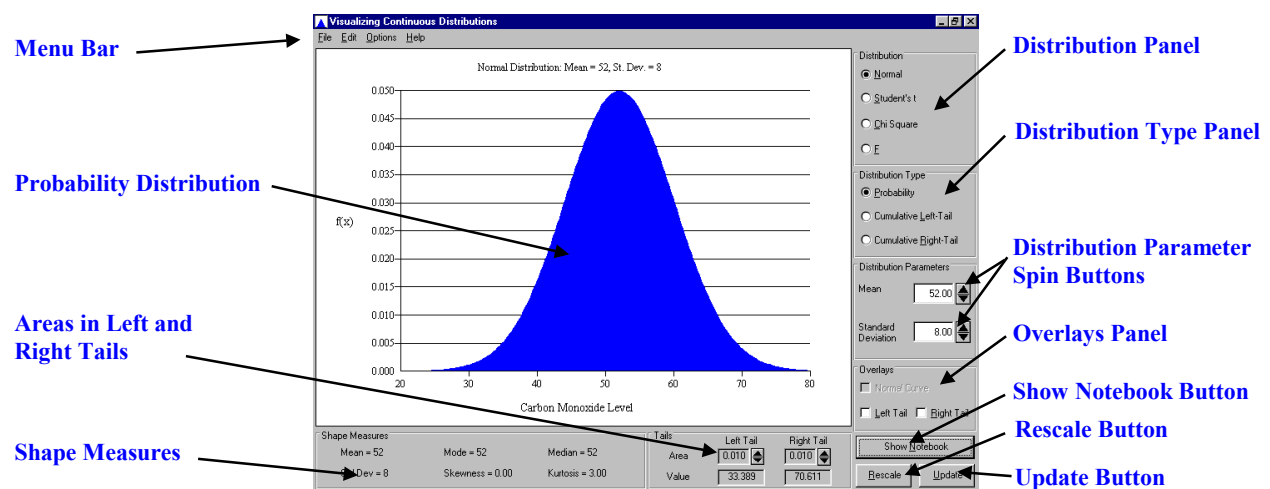
2. **Changing Parameter Values**
   Use the Parameter spin buttons to change the parameters of the normal distribution. Alternatively, you may type in parameter values. Click on Update to view the new distribution. Note how some of the shape measures change and some remain the same.

3. **Rescale**
   If the distribution is not fully visible, click the Rescale button.

4. **Viewing the Tails of the Distribution**
   In the Overlays panel on the Control Panel click on Right Tail. Note the shaded green area on the display. You can change the size of this area with the Right Tail spin button in the Tails panel below the distribution display. You can also view the left tail area.

5. **Type of Display**
   Select either Cumulative Right-Tail or Cumulative Left-Tail on the Distribution to Display panel. Click Update to see the graph selected.

6. **Selecting a Distribution**
    Return to the Notebook by clicking the Show Notebook button. Click on Return to the Scenario contents page. Select either Student's t distribution, chi-square distribution, or F distribution. Select a scenario from the choices given. Note that for each of these distributions the scenarios involve statistical tests. If you do not wish to deal with such scenarios you can alternatively select Student's t, Chi-Square, or F from the Distribution panel and click Update.

7. **Options**
    Two basic options are available from the Options menu on the menu bar on the top of the screen.
    a. Select Auto Update from the Options menu to update the distribution automatically as parameter values are changed.
    b. Select Auto Rescale on Update from the Options menu to rescale the distribution automatically when any parameter is changed. Although it guarantees plenty of graph detail, this option can make it difficult to see changes in the shape of the distribution because the scale on both axes changes each time you change a parameter.

8. **Copying a Display**
    Click on the display you wish to copy. Its window title will be highlighted. Select Copy from the Edit menu (on the menu bar at the top of the screen) or Ctrl-C to copy the display. It can then be pasted into other applications, such as Word or WordPerfect, so it can be printed.

9. **Help**
    Click on Help on the menu bar at the top of the screen. Search for Help lets you search a topic index, Contents shows a table of contents for this module, Using Help gives instructions on how to use Help, and About gives licensing and copyright information about this *Visual Statistics* module.

10. **Exit**
    Close the module by selecting Exit in the File menu (or click ☒ in the upper right-hand corner of the window). You will be returned to the *Visual Statistics* main menu.

Orientation to Additional Features

1. **Normal Curve**

    If you select any distribution but the normal you can select Normal Curve from the Overlays panel. This superimposes a normal curve on the distribution for a visual evaluation. For the Student's t distribution, the standard normal is shown (unless you have selected the option discussed below). For all other distributions, the normal is drawn using the mean and standard deviation that are indicated in the shape measures. In the figure below, the normal curve is shown superimposed on a chi-square distribution with 30 degrees of freedom.
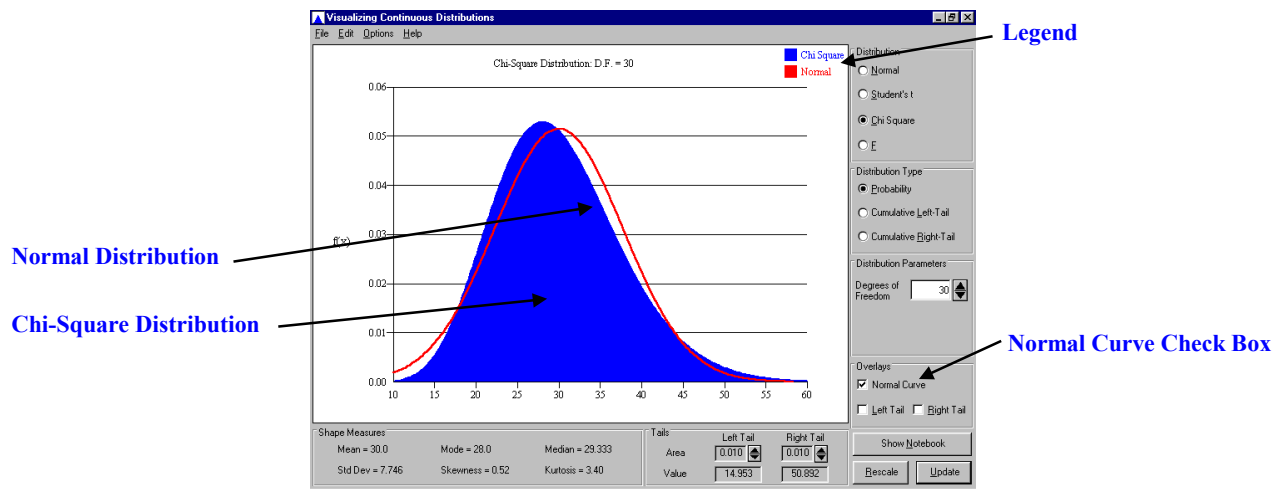
2. **Options**

    From the Options menu on the menu bar on the top of the screen you can select Standard Normal Overlay for t. This option is selected by default. It applies only if you display a Student's t distribution and select the Normal Curve overlay. The standard normal curve has mean 0 and standard deviation 1, as shown in most textbooks. However, if you look at the Shape Measures, you will notice that the standard deviation of the Student's t distribution is *not* exactly 1. In fact, the standard deviation varies with the degrees of freedom (d.f.) and approaches 1 only for large d.f. If you wish to see a normal distribution overlay with the true standard deviation, under Options you can deselect Standard Normal Overlay for t

Basic Learning Exercises   Name _____

Normal Distribution

Select a scenario under Normal Distribution in the Notebook. Make sure all of your overlays are turned off and select Auto Update, but do *not* select Auto Rescale under Options on the menu bar.

1. Change the mean using the Mean spin button. What happens to the shape and location of the distribution? If it slides off the screen, click Rescale.

2. Experiment with the Standard Deviation spin button. What happens to the shape of the distribution as the standard deviation increases? As it decreases? Click the Rescale button. Explain why the shape returns to the bell-shaped curve.

3. The normal distribution has two parameters: μ and σ. Since the mean of the normal distribution is μ and its standard deviation is σ, we refer to them as the mean and standard deviation. Vary these two parameters and observe the three measures of central tendency (mean, median, and mode). How do they change? Do the shape measures (skewness and kurtosis) change as you vary μ and σ?

4. Click the Show Notebook button. Select Carbon Monoxide Level. Answer the two questions in the scenario. On the Overlays panel select both Left Tail and Right Tail. Use the Left Tail and Right Tail spin buttons to find the approximate probabilities for the two questions.

5. Type 0 for Mean and 1 for Standard Deviation (this is the standard normal distribution). In the Options, click Auto Rescale on Update. Use the Right Tail spin button to find the X axis values that correspond to areas of 0.01 and 0.025. What are these values if you set Standard Deviation = 2? If you set Standard Deviation = 3?

	$\sigma = 1$	$\sigma = 2$	$\sigma = 3$
Area = 0.01	Z Value = _____	Value = _____	Value = _____
Area = 0.025	Z Value = _____	Value = _____	Value = _____

6.   What is the relationship between these X axis values and the standard deviation?

7.   Set Mean = 4 and Standard Deviation = 3. Show algebraically that if the right-tail area is 0.05 the value of the normal distribution is 8.935. Can you write an equation that shows how to get the X value for a right-tail area of 0.05 for any values of μ and σ?

8.   The standard normal variate is defined as $Z = (X - \mu)/\sigma$. In the Carbon Monoxide scenario, what would $Z = 0$ represent? What would $Z = 1$ represent?

9.   Create a standard normal distribution ($\mu = 0$, $\sigma = 1$). a) From the graph, how likely is it that a standard normal random variable is within the range –2 to +2? b) Within the range –3 to +3? c) For any normal distribution, why is a value beyond the range $\mu \pm 3\sigma$ called an "outlier"?

10.   In quality control, we sometimes hear that defects are controlled at the "6-sigma" level. This means that manufacturing capability is so refined that an extremely high proportion of the items produced will meet the desired specifications. How likely is it that a normal variate will fall outside the interval $\mu \pm 6\sigma$? Why might it be difficult to attain a "6-sigma" defect rate?

11.   Keep the standard normal distribution (mean 0, standard deviation 1). In the Distribution Type control panel, select Cumulative Left-Tail. Click Update. a) From the graph, what is the probability that z will be below 0? b) Below 1? Below 2? Below 3? c) If you took a test and scored two standard deviations above the mean, how unusual would that be (assuming that exam scores are normally distributed)? d) If you scored three standard deviations above the mean, what would that tell you?

Intermediate Learning Exercises   Name _____

Student's t Distribution
Select Student's t Distribution and Probability Distribution. Set Degrees of Freedom = 1, display the Right Tail overlay (turn all others off).   Deselect Auto Rescale under Options.

12. Set Right Tail area to 0.05. What t-value corresponds to this area? Find the t-value if DF = 6? DF = 11? DF = 16? DF = 21? What happens to the shape and location of the distribution? Click the Rescale button when necessary.

DF = 1 _____   DF = 6 _____   DF = 11 _____   DF = 16 _____   DF = 21 _____

13. Change the DF and observe the shape measures. Find the mean, mode, median, and skewness. Do the standard deviation and kurtosis approach some value as the DF becomes very large?

14. Select Normal Curve on the Control Panel. Set the DF to 10. Compare the distributions. How are they different? Change the DF to 20, and then to 30. Has the difference increased or decreased? Why? **Hint:** The normal curve displayed has a mean of 0 and a variance of 1.

15. Select Cumulative Left-Tail. Change the DF to 10. Compare the two distributions. How are they different? Change the DF to 20. Has the difference increased or decreased? Why?

16.   Is the difference between the normal and t distribution more noticeable in the probability distribution or the cumulative distribution function? Why?

Chi-Square Distribution

Select Chi-Square and Probability Distribution. Set Degrees of Freedom = 10. Turn all the overlays off, and do *not* select Auto Rescale under Options on the menu bar.

17. What happens to the location and shape of the distribution as you increase and then decrease Degrees of Freedom (DF) using the spin button? If it slides off the screen, click Rescale.

18. Observe the shape measures as you vary the degrees of freedom. What relationship exists between the DF and the mean, mode, median, and standard deviation? What general observation can you make about skewness and kurtosis?

19. Change DF to 5 and select Normal Curve and Cumulative Left-Tail on the Control Panel. Click Rescale if necessary. Why are the cumulatives so different from one another? **Hint:** Look at their probability distributions.

F Distribution

Select F Distribution and Probability Distribution. Set Numerator DF = 3 and Denominator DF = 3. Select Auto Update and Auto Rescale on Update under Options on the menu bar.

20. Change Numerator DF to 10, then to 20, then to 200. What happens to the distribution's shape? Did the axis scale change?

21. Select the Normal Curve overlay. Keeping Numerator DF at 200, increase Denominator DF to 10, then to 20, then to 200. (a) What happens to the distribution's shape? Does it resemble the normal overlay? (b) What happens to the length of the right tail? (c) What happens to the mean, mode, skewness, and kurtosis as both degrees of freedom increase?

Advanced Learning Exercises   Name _____

Investigate the Chi-Square Distribution
Choose Normal Distribution and Cumulative Left-Tail. Set Mean = 0 and Standard Deviation = 1.
Select Auto Update and Auto Rescale under Options on the menu bar.

22. Using the standard normal cumulative left-tail distribution below, label the ordinate (vertical axis) when X equals -1.5, -1, -0.5, 0, 0.5, 1, and 1.5. As a guide, 1.5 has been done for you.

Normal Distribution: Mu = 0, Sigma = 1
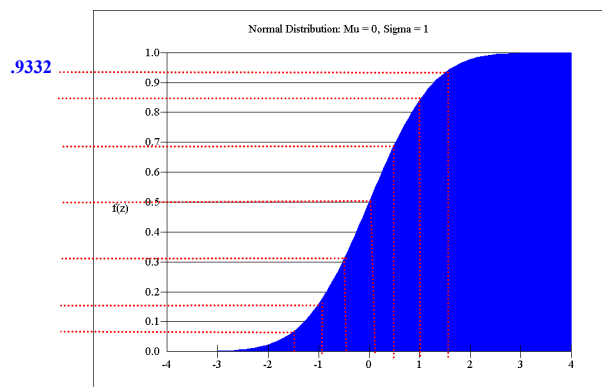
23. Click on Chi-Square Distribution and change Degrees of Freedom to 1. Using the chi-square cumulative left-tail distribution below, label the ordinate (vertical axis) when X equals 0, .25, 1, and 2.25.

Chi-Square Distribution: D.F. = 1
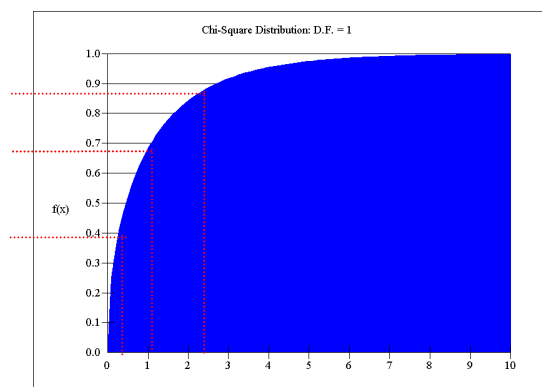
24. If a standard normal random variable is squared, the new random variable is distributed as a chi-square with 1 degree of freedom. Using the cumulatives above, show that the probability between –0.5 and 0.5 on the cumulative normal equals the probability that a chi-square variable will be less than 0.25. Also show that the probability between -1 and 1 on the normal cumulative is equal to the probability that a chi-square random variable will be less than 1.0. Finally, show that the probability between –1.5 and 1.5 on the normal cumulative equals the probability that a chi-square random variable will be less than 2.25.

25. Why are the answers in exercise 24 consistent with the fact that a standard normal random variable squared becomes a chi-square random variable?

Investigate the F Distribution

Choose F Distribution and Cumulative Left-Tail. Set Numerator DF = 1, Denominator DF = 10, and select Auto Update, and Auto Rescale under Options on the menu bar.

26. Student's t distribution is the ratio of a standard normal and the square root of an independent chi-square random variable divided by its DF. An F distribution is the ratio of two independent chi-square random variables divided by their DF. Therefore, if a variable that is distributed as t with k DF is squared, the new random variable is distributed as F with 1 and k DF. Design an experiment similar to the one in exercises 23–25 to illustrate this relationship.

27. An F distribution is the ratio of two independent chi-squares divided by their degrees of freedom. As a chi-square's DF increases, it approaches a normal distribution. Therefore, it is reasonable to assume that as Denominator DF increase, an F distribution approaches the ratio of the chi-square distribution (whose DF equals the F's Numerator DF) to its DF. Fill in the table below with the values along the horizontal axis that correspond to an area of 0.05 in the right tail of the F and the chi-square distributions. How do these results illustrate that an F with n and d DF approach a chi-square with n DF as d becomes large? Why does the F approach a chi-square divided by its DF rather than just a chi-square?

Values Along Horizontal Axis That Correspond to an Area of 0.05 in the Right Tail

	$F_{1,d}$	$F_{10,d}$	$F_{200,d}$
d=1	_____	_____	_____
d=50	_____	_____	_____
d=200	_____	_____	_____
	$\chi^2_1/1=$ _____	$\chi^2_{10}/10 =$ _____	$\chi^2_{200}/200 =$ _____

Individual Learning Projects

Write a report on one of the three topics listed below. Use the cut-and-paste facilities of the module to place the appropriate graphs in your report.

1. Investigate the normal approximation to the Student's t, chi-square, and F distributions. Select a value of skewness and a value of kurtosis as criteria for when any of the three distributions can be approximated by a normal distribution. The criteria should provide a *Very Good* approximation. How did you select the values for skewness and kurtosis? Suggest a procedure to evaluate your criteria. Using this procedure, do the criteria work equally well for all three distributions? Repeat the process to establish and evaluate a *Fair* approximation criterion (not as stringent a criterion). **Hint:** A normal distribution has skewness of 0 and kurtosis of 3.

2. Investigate the shape of the chi-square distribution for a number of parameter values. Show clearly how degrees of freedom affect the appearance of this distribution and its shape measures. Over what range of degrees of freedom (if any) does it closely resemble a normal distribution, in your judgment? How would you determine the mean and variance of the normal distribution that approximates a chi-square distribution with k degrees of freedom? **Hint:** Use the Help files to obtain formulas for the various shape measures.

3. Investigate the Student's t distribution with 1 to 6 degrees of freedom. Compare the shape measures as well as the shape and size of the tails of the distribution. A Student's t distribution with 1 degree of freedom has an unusual appearance. How is it unusual? Why does the graph of its cumulative in this module not show the values of t whose probability appears to equal 0 or 1? **Hint:** Use the Tails spin buttons to find the t-value that corresponds to an area of 0.005 in each tail.

Team Learning Projects

Select one of the three projects listed below. In each case produce a team project that is suitable for an oral presentation. Use presentation software or large poster boards to display your results. Graphs should be large enough for your audience to see. Each team member should be responsible for producing some of the graphs. Ask your instructor if a written report is also expected.

1. This project is for a team of three. The team is to investigate the shape of the F distribution and how it depends on its two parameter values. The team should select five values for the numerator degrees of freedom (being sure that the range of values from 1 to 200 is covered). Each team member should select two values of the denominator degrees of freedom (the team should cover the range of degrees of freedom from 1 to 200). Each team member will produce ten distributions (one for each combination of the numerator and denominator degrees of freedom). The team should be able to show how increasing the numerator degrees of freedom with constant denominator degrees of freedom affects the shape of the F distribution. The team should also show how increasing the denominator degrees of freedom with constant numerator degrees of freedom affects the shape of the F distribution. Be sure to note the values of the shape measures and the height and width of each F distribution.

2. This project is for a team of two. The team is to investigate the normal approximation to the Student's t, χ^2, and F distributions. Many statisticians would say that a Student's t with 30 degrees of freedom can be approximated with a standard normal distribution. Compare the areas in the tails of the Student's t and standard normal distributions. Make this comparison at four different points. Many statisticians would say that a chi-square distribution with 200 degrees of freedom can be approximated with a normal distribution. Find its mean and standard deviation. Compare the areas in the upper and lower tails of the chi-square and normal distributions. Make these comparisons at four different points. Many statisticians would say that an F distribution with 200 and 200 degrees of freedom can be approximated with a normal distribution. Find its mean and standard deviation. Compare the areas in the upper and lower tails of the F and normal distributions. Make these comparisons at four different points. Based upon these three evaluations, characterize the quality of each approximation. **Hint:** Use the Help files to obtain relevant formulas.

3. Investigate the F distribution when both degrees of freedom are less than or equal to 9. Each team member should take a value for the denominator's degrees of freedom (the team should make sure the range from 1 through 9 is covered) and examine the F distribution for the case of 1 through 9 numerator degrees of freedom. Compare the distribution's shape, the values for both the skewness and kurtosis shape measures, and the area in the tails of the distribution.

Self-Evaluation Quiz

1.   Which is *not* a characteristic of a continuous distribution?
     a. It always has an area of 1.
     b. It always has a height (probability) between 0 and 1.
     c. Its domain is defined as an interval rather than a series of points.
     d. It is characterized by one or more parameters.
     e. Its cumulative distribution approaches 1 as we move to the right.

2.   In the normal distribution with $\mu = 100$, $\sigma = 20$, which X defines the smallest upper-tail area?
     a. 80
     b. 110
     c. 95
     d. 130
     e. Insufficient information is given.

3.   A right-tail area of 0.10 in a normal distribution with $\mu = 80$, $\sigma = 15$ corresponds to an X value of
     a. 114.89
     b. 104.68.
     c. 99.23.
     d. 109.40.
     e. Insufficient information is given.

4.   If the upper 1% point in a normal distribution with $\mu = 500$, $\sigma = 20$ is 546.52, then the upper 1% point in a normal distribution with $\mu = 500$, $\sigma = 40$ is
     a. 546.52
     b. 580
     c. 612.24
     d. 593.04
     e. Insufficient information is given.

5.   Which X value defines the largest right-tail area in a standard normal distribution?
     a. 0.000
     b. 1.645
     c. 1.960
     d. 2.326
     e. 2.576

6.   Which is *not* true of the Student's t distribution?
     a. It closely resembles a normal distribution for degrees of freedom above 10.
     b. It was discovered about a century ago.
     c. It is more symmetric for large degrees of freedom.
     d. It is always centered at 0.
     e. It was discovered by a beer brewer.

7.   Which is *not* the name of a parameter of a common continuous distribution?
    a. Mean.
    b. Kurtosis.
    c. Variance.
    d. Numerator degrees of freedom.
    e. Denominator degrees of freedom.

8.   The chi-square distribution
    a. is almost indistinguishable from a normal distribution for degrees of freedom above 30.
    b. is always skewed to the right.
    c. has two parameters (numerator and denominator degrees of freedom).
    d. has the same left-tail and right-tail 1% critical values except for the sign.
    e. fulfills more than one of the above.

9.   The chi-square distribution
    a. has a mean equal to its degrees of freedom.
    b. has a mode equal to its degrees of freedom.
    c. is leptokurtic (more peaked than normal).
    d. has more than one of the above characteristics.
    e. has none of the above characteristics.

10.   The F distribution looks quite like a normal distribution
    a. when $df_1 = 1$ and $df_2 = 10$.
    b. when $df_1 = 2$ and $df_2 = 30$.
    c. when $df_1 = 5$ and $df_2 = 100$.
    d. when $df_1 = 10$ and $df_2 = 200$.
    e. in none of the above cases.

11.   Which is *not* true of the F distribution?
    a. It was discovered by Sir Ronald Fisher in the 20th century.
    b. It is nearly normal when $df_1 = 25$ and $df_2 = 25$.
    c. It is always unimodal.
    d. It is always a continuous distribution.
    e. Its mean decreases as we increase its denominator degrees of freedom.

12.   The F distribution
    a. is symmetric if $df_1 = df_2$
    b. is always skewed left.
    c. is always platykurtic (flatter than normal).
    d. fulfills both b and c.
    e. fulfills none of the above.

Glossary of Terms

Chi-square distribution Right-skewed continuous distribution that describes the sum of squares of unit normal deviates. Its one parameter is called degrees of freedom.

Continuous distribution A distribution function where the random variable is defined over a continuous X domain.

Cumulative distribution A function that maps each value of a random variable to the probability of being less than or equal to that-value. The function begins at 0 and rises to 1 as you move to the right (or, less commonly from 1 to 0 as you move to the left). See **Probability distribution**.

Degrees of freedom Name given to parameters of certain distributions. Degrees of freedom will always be an integer. It is sometimes abbreviated DF or d.f. See **Chi-square distribution, F distribution, and Student's t distribution**.

F distribution Right-skewed continuous distribution used to describe the ratio of two independent sample variances. Its two parameters are called* the numerator and denominator degrees of freedom.

Kurtosis Measure of the relative peakedness of a distribution. $K = 3$ indicates a normal "bell-shaped" distribution (mesokurtic). $K < 3$ indicates a platykurtic distribution (flatter than a normal distribution with shorter tails). $K > 3$ indicates a leptokurtic distribution (more peaked than a normal distribution with longer tails).

Mean Expected value of a random variable. It may be interpreted as the fulcrum (balancing point) of the distribution along the X-axis. It is commonly denoted μ.

Median Point along the X-axis that defines the upper and lower 50 percent of the distribution. In a symmetric distribution, it is equal to the mean.

Mode X value that corresponds to the peak of the probability distribution function.

Normal approximation Using a normal distribution to approximate another distribution.

Normal distribution Standard "bell-shaped" or Gaussian distribution. It has two parameters called the mean and variance.

Parameter Numerical characteristic of a population that determines its distribution. Some distributions have several parameters.

Probability distribution Distribution that maps each value of a random variable to a probability. The area (integral) under the entire probability distribution must be 1. Also called a probability density function. See **Cumulative distribution**.

Shape measures Descriptive statistics such as the distribution's mean, median, mode, standard deviation, skewness, and kurtosis.

Skewness Measure of relative symmetry. Zero indicates symmetry. The larger its absolute value the more asymmetric the distribution. Positive values indicate a long right tail, and negative values indicate a long left tail.

Standard deviation The square root of the variance. It is commonly denoted σ.

Student's t distribution Symmetric continuous distribution used to describe the ratio of a normally distributed variable with zero mean divided by its standard deviation. Its one parameter is called degrees of freedom.

Variance A measure of dispersion equal to the expected value of $(X - \mu)^2$. The larger the variance, the greater the dispersion or "spread" around the mean. See **Standard Deviation**.

Solutions to Self-Evaluation Quiz

1. b   Read the Overview of Concepts.
2. d   Do Exercises 5–7.
3. c   Do Exercise 5.
4. d   Do Exercises 6 and 7.
5. a   Do Exercise 5.
6. c   Do Exercise 13. Read the Overview of Concepts.
7. b   Do Exercises 3, 12, 17, and 20–21. Read the Overview of Concepts.
8. b   Do Exercise 18. Read the Overview of Concepts.
9. d   Do Exercises 17–19. Read the Overview of Concepts.
10. e   Do Exercises 20–21. Read the Overview of Concepts.
11. b   Do Exercises 20–21. Read the Illustration of Concepts.
12. e   Do Exercises 20–21. Read the Overview of Concepts.