Remember, hypothesis testing is to determine if there is a significant difference in
the data: “Are two distributions significantly different?” In this case, you are interested in
the distribution of the data – how it is spread out. Where might this apply? In 2000 the
U.S. Census Bureau indicated that 19% of the population was in the Northeast, 35.6% was
in the South, 22.9% in the Midwest, and 22.55 in the West. The Census Bureau might want
to track samples of each area to see if this distribution has changed significantly. The
government might want to know the population distribution in areas of the country to help
determine the number of representatives in Congress or help craft new laws. A business might be interested in this
data for retail outlets (stores) and locations of warehouses. Think of a local cereal company that tracks the
distribution of sales of different kinds of its cereals; a change in data would require a change in production and
distribution of the types of cereal.

In the past, standard deviation was used to examine the distribution of the data; as you know, it is a measure of
the spread of the data. You also know that distributions can be normal, skewed, bimodal, J-shaped or other. Since
standard deviation does not lend itself to a nice analysis, the variance will be used in this type of hypothesis testing.
As you may or may not recall, the variance is the square of the standard deviation. If a number is squared, it can
only be positive, so this distribution, called the Chi-squared distribution, is only in the first quadrant; if the sample
size is small, it is a positively-skewed distribution. The Greek letter chi (… pronounced “Ki”) that looks like a big X,
will be used. The \( X^2 \) (Chi-squared) distribution changes relative to the sample size; if you think about it, with a
small set of data, the standard deviation will not be very big; as “n” gets larger, the standard deviation and the
variance might get larger. The \( X^2 \) distribution works like the Student’s t-distribution using degrees of freedom; d.f. =
n – 1, each row of the table is, essentially, a different distribution.

For teaching purposes, a common example that’s used is the distribution of colors of M&M candies because,
in the past, the percent of each color could be found at the M&M Mars web site; this is no longer available at the web
site, but, since most people are familiar with the candy, it will be used as an example. M&M candies last reported the
distribution of colors to be the following: 24% blue, 20% orange, 16% green, 14% yellow, and 13% red and brown.
For the purpose of learning the process, let’s say you buy a case of 24 bags of candy and count the colors; that will be
about 2,688 pieces. (In a real-life situation, I usually had two math classes actually gather the data insisting they
record it before “destroying” it.) If the results were as follows: 605 blue, 494 orange, 460 green, 403 yellow, 369 red
and 357 brown, do the data indicate the distribution of the colors of M&Ms are different than what they were in the
past? These numbers of the colors from the sample are known as the observed values. You might suspect the
distribution has change since the expected values, the colors you would expect to get from a sample of 2688 pieces of
M&M candies, if the original distribution did not changed, would be 645 blue (24% of 2,688), 538 orange, 430 green, 376 yellow, 349 red and 349 brown. Here are side-by-side comparisons:

| Observed values: | 605 blue, 494 orange, 460 green, 403 yellow, 369 red and 357 brown |
| Expected values:  | 645 blue, 538 orange, 430 green, 376 yellow, 349 red and 349 brown |

Let’s set up a hypothesis test and test \( H_0 \): at the \( \alpha = 5\% = 0.05 \) level of significance.

\( H_0: \quad \) The distribution of colors is the same as previously reported. (There is no change in the color distribution.)

\( H_1 = H_a: \) The distribution of colors is different. (The distribution of colors has changed.)

The \( X^2 \) is determined by this formula: \( X^2 = \) the sum of \( (\text{Observed} - \text{Expected})^2 / \text{Expected} \) =
\( (605 - 645)^2 / 645 + (494 - 538)^2 / 538 + (460 - 430)^2 / 430 + \ldots + (357 - 349)^2 / 349 = 11.44 \)

Go to the \( X^2 \) distribution table and choose row 5; since \( n = 6 \), degrees of freedom = 5. Where does 11.44 show up in
that row? Isn’t it between the 0.05 and 0.025 columns? That means the p-value is between 0.05 and 0.025 but
close to 0.05. Is the p-value less than $\alpha = 0.05$? Yes, it is, so normally you would reject $H_0$; the distribution of colors of M&M candies in our recent sample is significantly different than what it was in the past. Read further!

A closer value for the p-value can be found in the calculator; go to the distributions (2nd, VARS) and choose $X^2$ cdf(11.44, some large number like 1000 or larger, 5), which represents the left and right sides of the area under the distribution you want and the degrees of freedom. What did you get? I got p-value = 0.0433, which is close to $\alpha = 0.05$. In this case, since the p-value is so close to $\alpha = 0.05$, I would say the test is inconclusive; you will need more data to make a better decision.

Can the whole thing be done on the calculator? This one cannot, but other types can be done on it. You can improve on the process by doing the following. Put the “observed” in L₁ and the “expected” in L₂; at the very top of L₃, with your curser on L₃, type in $(L₁ - L₂)^2 / L₂$ and move the curser down with your arrow key. The calculator will calculate the individual “addends” for your $X^2$. Now go to the home screen (2nd MODE = QUIT). On the home screen type 2nd, STAT, scroll over to MATH and choose #5: “sum(“; type in L₃ and press ENTER. That will give you the $X^2$ value, 11.44. As long as you typed the data into the lists properly, this will reduce errors and the time it takes to get your $X^2$ value. Now use the distribution $X^2$ cdf( left, right, d.f.) to get the p-value for making your decision to reject or not reject the null hypothesis.

Let’s try another. At a local hospital emergency room, ER, patients occur uniformly throughout the evenings between 5 and 11 pm? Here are summation data for the last four weeks: Sunday 43; Monday 48; Tuesday 51; Wednesday 57; Thursday 56; Friday 83; Saturday 68. What is the sum of the patients? If ER patients came in a uniform way throughout the week, what are the expected values? What are your hypotheses? Test your hypothesis at the $\alpha = 0.05$ level of significance. Find $X^2$ and the p-value. Write your conclusion in the context of the problem.

$H₀$: Patients come uniformly throughout the week to the emergency room between 5 to 11 pm.

$H₁ = Hₐ$: Patients do not come uniformly throughout the week to the emergency room between 5 to 11 pm.

I put the data in the list menus: L₁ is 43; 48; 51; 57; 56: 83: 68. Each entry in L₂ is the sum of all days divided by seven = 406 / 7 = 58, in other words, if patients came uniformly throughout the week, the number of patients each evening would be about 58. The $X^2 = \text{sum of the values in the third column after you type in } (L₁ - L₂)^2 / L₂ \text{ at the top of L₃}$. I got $X^2 = 19.03$. What is the p-value? Go to row 6 of the $X^2$ distribution and find 19.03; it is between which two columns? Since it’s larger than the 18.548 in the last column, then the p-value is less than 0.005, which is $< \alpha = 0.05$, so reject $H₀$; Patients do not come uniformly throughout the week to the emergency room from 5 to 11 pm. The more accurate p-value is $X^2$ cdf( left, right, d.f.) = $X^2$ cdf( 19.03, 1000, 6) $≈ 0.0041$.

There are other uses for the $X^2$ distribution. One is to determine if two variables are independent of each other in the setting of a “two-way contingency” table. Here is an example of some factious data. “Are you generally satisfied with your job / career?” Does job satisfaction depend on the level of education?

<table>
<thead>
<tr>
<th>Level of Education</th>
<th>Less than High School</th>
<th>High School</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior College</td>
<td>103</td>
<td>45</td>
</tr>
<tr>
<td>Bachelor</td>
<td>87</td>
<td>35</td>
</tr>
<tr>
<td>graduate</td>
<td>89</td>
<td>18</td>
</tr>
</tbody>
</table>

This situation will be discussed in the next calculator-use handout Test for Independence.