

**Count the Occurrence of letter ‘I’**

A country preacher was walking the back-road near a church. He became thirsty so decided to stop at a little

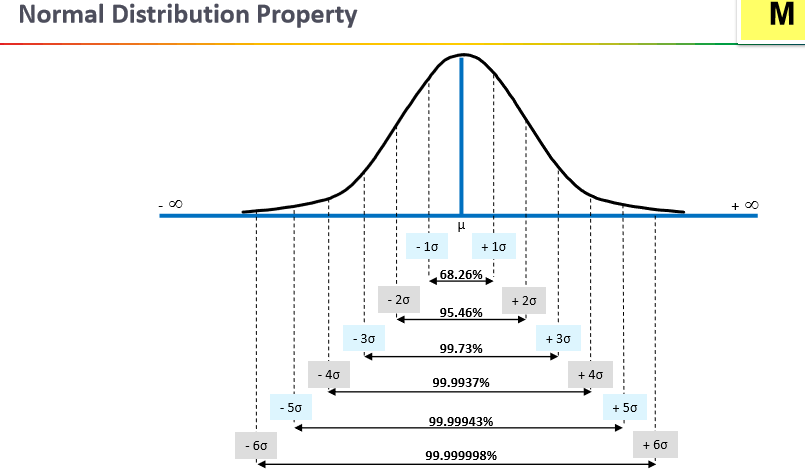
cottage and ask for something to drink. The lady of the house invited him in and in addition to something to

drink, she served him a bowl of soup by the fire. There was a small pig running around the kitchen. The pig was

constantly running up to the visitor and giving him a great deal of attention. The visiting pastor commented that

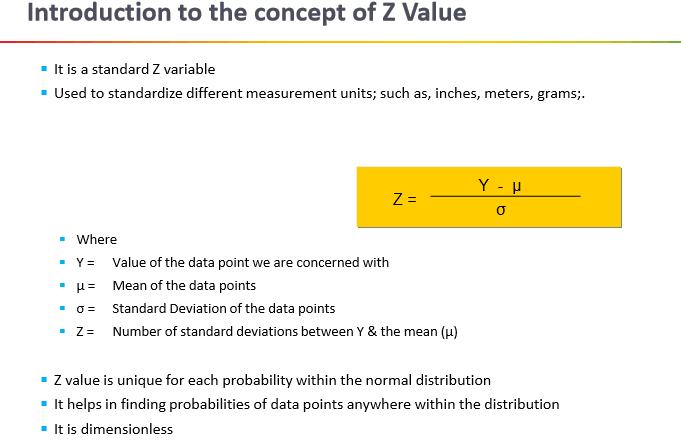
he had never seen a pig this friendly. The housewife replied: "Ah, he's not that friendly. Actually, that's his bowl

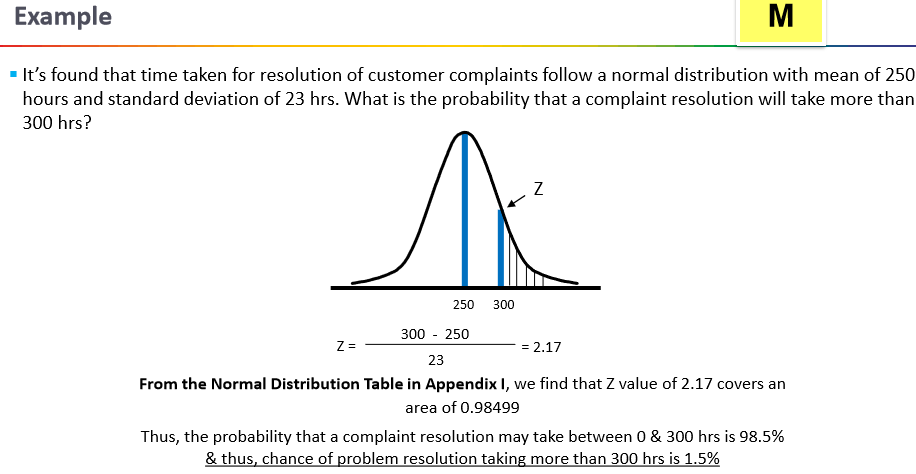
you're using!"

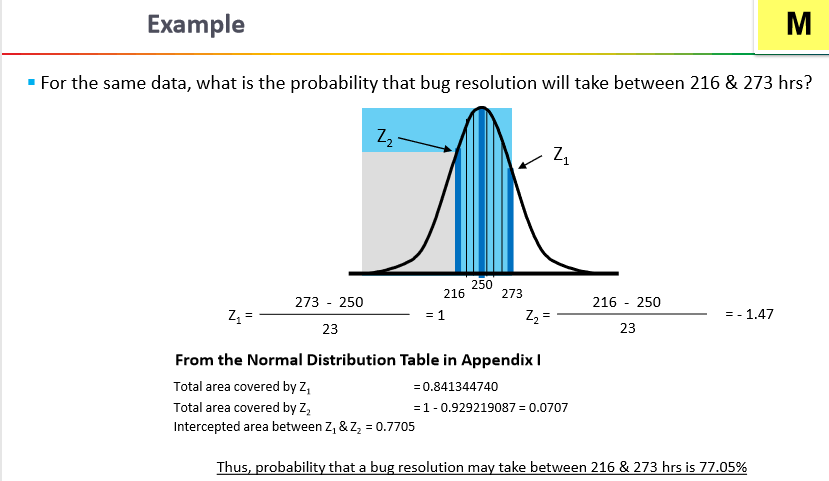


**Z-Test**

A z-test is a test of the standard normal distribution, a bell-shaped curve with a mean of 0 and a standard deviation of 1. These tests arise in many statistical procedures. A P-value is a measure of the statistical significance of a statistical result. Statistical significance addresses the question: "If, in the entire population from which this sample was drawn, the parameter estimate was 0, how likely are results as extreme as this or more extreme?" That is, it provides a basis to determine whether an observation of a sample is merely the result of random chance (that is, to accept the null hypothesis) or whether a study intervention has in fact produced a genuine effect (that is, to reject the null hypothesis).







Z1 = (273 – 250)/23 = 1. From Z Table value is 0.8413

Z2 = (216-250)/23 = -1.47. From Z Table Value is 0.0708

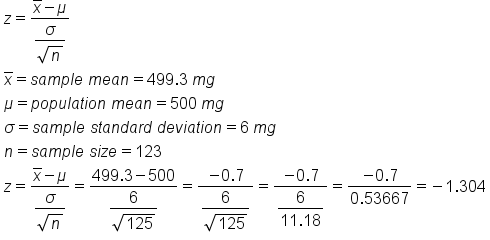
Z1 – Z2 = 0.7705 = 77.05%

**1. Two-Sided Tests**

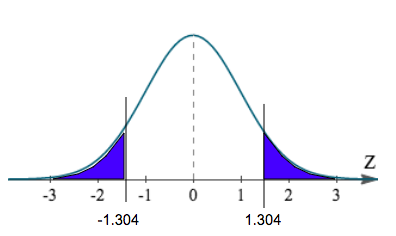
Suppose a pharmaceutical company manufactures ibuprofen pills. They need to perform some quality assurance to ensure they have the correct dosage, which is supposed to be 500 milligrams. This is a two-sided test because if the company's pills are deviating significantly in either direction, meaning there are more than 500 milligrams or less than 500 milligrams, this will indicate a problem.

table attributes columnalign left end attributes row cell H subscript 0 colon space mu equals 500 space m g end cell row cell H subscript A colon space mu not equal to 500 space m g end cell end table

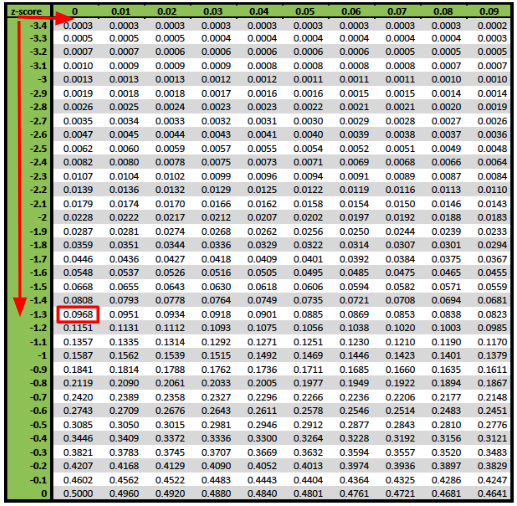
In a random sample of 125 pills, there is an average dose of 499.3 milligrams with a standard deviation of 6 milligrams. Because this is quantitative data, 500 mg is the population mean. We can use the following formula to calculate the z-score:



We get a z-score of negative 1.304. Because this is a two-sided test, it is not enough to just look at the left tail. We also have to look at the equivalent of the right tail, or a positive 1.304.



The first way to find the p-value is to use the z-table. In the z-table, the left column will show values to the tenths place, while the top row will show values to the hundredths place. If we have a z-score of -1.304, we need to round this to the hundredths place, or -1.30. In the left column, we will first find the tenths place, or -1.3. In the top row, we will find the hundredths place, or 0.



This results in a p-value of 0.0968, or 9.68%, for a z-score of negative 1.304. We also need to take the positive 1.304 into account, which is the upper right tail. To calculate the true p-value, we just need to multiply 0.0968 by two, or 0.1936. This would be a p-value of 19.36%.

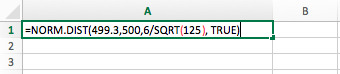
The second method is using a graphing calculator. This can give us a more exact number because we will not have to cut off the z-score at the hundredths place. On the calculator, click "2nd", then "DISTR" for distribution. We will use "normalcdf", which stands for normal cumulative density function. When inserting the values into the calculator, we always go lower boundary to upper boundary.

In this case, the lower boundary was shaded all the way to the left of the curve, which would be negative infinity. We cannot enter negative infinity in our calculator, so instead, we can just enter negative 99. The shading stops at -1.304, so this is the upper boundary.



We get a value of 0.0961, which is about the same value as we got in the table. Again, we need to take both tails into account, so we can simply multiply this value by two to get a p-value of 0.1922, or 19.22%.

The third method to find the p-value is to use Excel. First, select "Formulas", choose the "Statistical" option, and pick "NORM.DIST". The first value we are going to input is the mean of the sample, which was 499.3, then the population mean which we are testing against, or 500, and finally the standard deviation, which was 6, divided by the square root of the sample of n. We can find the square root under the "Math and Trigonometry" option in "Formulas". The last value that we need to enter is "TRUE".

File:5230-pvalue5.png

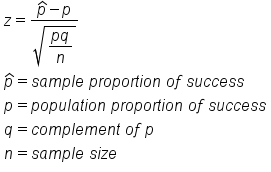
We get about the same value as we did with the table and the calculator. Since this is a two-sided test, we need to multiply the value by two, or 0.096 times two equals 0.1812.

**2. Left-Tailed Test**

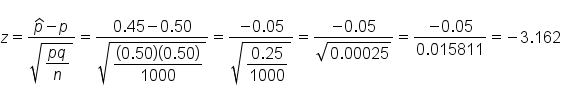
In this next example, we'll look at the proportion of students who suffer from test anxiety. We want to test the claim that fewer than half of students suffer from test anxiety.

table attributes columnalign left end attributes row cell H subscript 0 colon space p equals 0.50 end cell row cell H subscript A colon space p less than 0.50 end cell end table

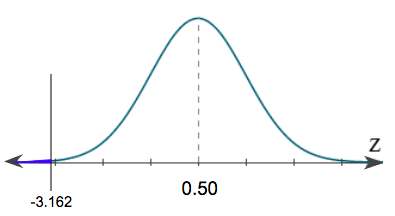
In this case, we will have a left-tailed test. Because this is qualitative data, meaning the students answer yes or no to suffering from test anxiety, this is a population proportion and we can use the following formula to calculate the z-test statistic:



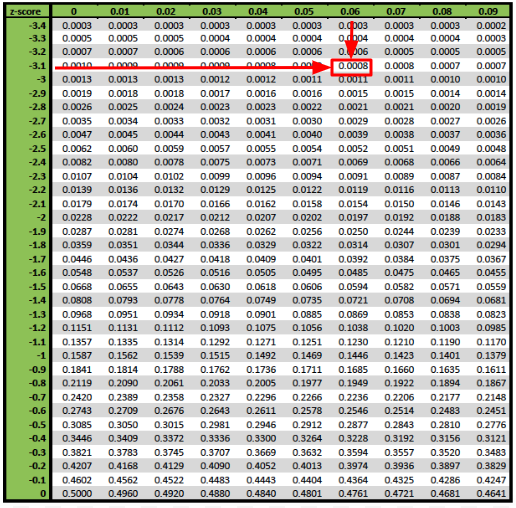
In a random sample of 1000 students, 450 students claimed to have test anxiety. This will be p-hat, or the sample proportion. We can calculate this by dividing 450 by 1000, or 0.45. The population proportion, p, is 50%, or 0.50. The complement of p, or q, can be found by calculating 1 minus 0.50, or 0.50. The sample size is 1000.



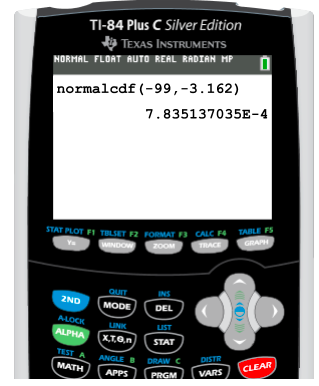
The corresponding z-score is negative 3.162. Testing against that half, or 50%, of students suffer from test anxiety, we get the following shaded region all the way to the left of our curve:



The first way to find the p-value is with the z-table. Remember, we can only go up to the hundredths place, so we will need to round -3.162 to -3.16. In the left column, we will first find the tenths place, or -3.1. In the top row, we will find the hundredths place, or 0.06. This gives us a p-value of 0.0008, or 0.08%.



To find the p-value on the graphing calculator, click "2nd", then "DISTR" for distribution. Again, we will use "normalcdf". When inserting the values into the calculator, remember we always go lower boundary to upper boundary. In this case, the lower boundary was shaded all the way to the left of the curve, which would be negative infinity. We cannot enter negative infinity in our calculator, so instead, we can just enter negative 99. The shading stops at -3.162, so this is the upper boundary.



This answer shows a p-value of 0.00078, or 0.078%.

In Excel, select "Formulas", choose the "Statistical" option, and pick "NORM.DIST". The first value we are going to input is the sample proportion, "0.45", then the population proportion, "0.50", and finally the standard deviation, which was the square root of pq divided by n, or 0.50 times 0.50 divided by 1000. We can find the square root under the "Math and Trigonometry" option in "Formulas". The standard deviation should be input as "SQRT((0.50\*0.50)/1000)". The last value that we need to enter is "TRUE".

File:5234-pvalue9.pngFile:5235-pvalue10.png

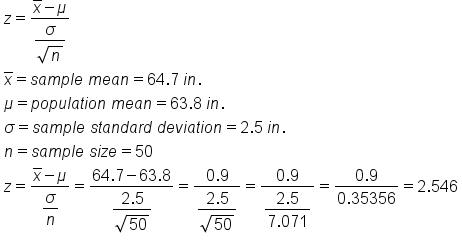
We get about the same p-value as we did with the z-table and the calculator.

**3. Right-Tailed Test**

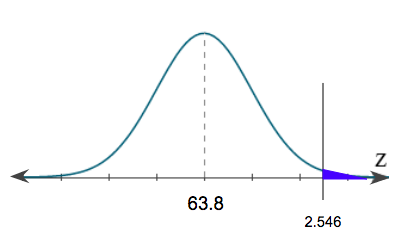
In this final example, we will be testing the claim that women in a certain town are taller than the average U.S. height, which is 63.8 inches.

table attributes columnalign left end attributes row cell H subscript 0 colon space mu equals 63.8 space i n. end cell row cell H subscript A colon space mu greater than 63.8 space i n. end cell end table

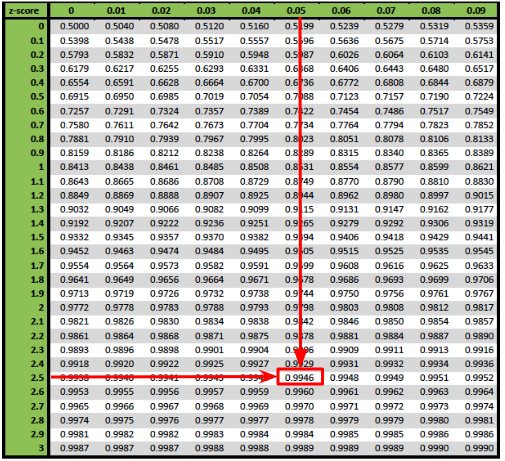
From a random sample of 50 women, we get an average height of 64.7 inches with a standard deviation of 2.5 inches. Inches is a quantitative variable, therefore the 63.8 inches is a population mean. We will then use the following formula to calculate the z-score:



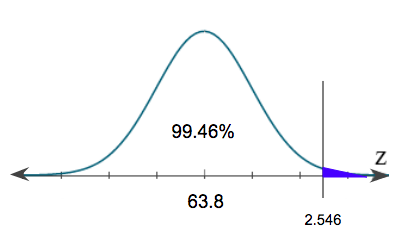
We get a z-score of 2.546, which is labeled on the following distribution:



The first way to find the p-value is to use the z-table. In the z-table, the left column will show values to the tenths place, while the top row will show values to the hundredths place. If we have a z-score of 2.546, we need to round this to the hundredths place, or 2.55. In the left column, we will first find the tenths place, or 2.5. In the top row, we will find the hundredths place, or 0.05. This gives us a p-value of 0.9946, or 99.46%.

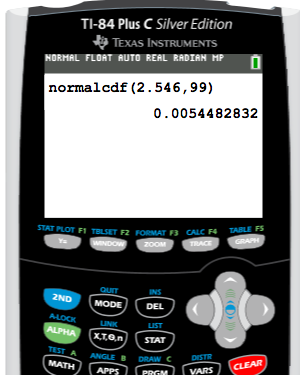


However, when we are performing an upper-tail test, or right-tailed test, that p-value from the table always reads left to right for our distribution. The p-value of 99.46% is associated with the 99.46% percent that is unshaded.



To get the percent that is shaded under the curve, we just need to calculate 100% minus 99.46%. This gives us the p-value of 0.54%, or 0.0054.

On the graphing calculator, again we are going to click "2nd", then "DISTR", and use "normalcdf". When inserting the values into the calculator, remember we always go lower boundary to upper boundary of the shaded region. In this case, the lower boundary of the shaded region is our z-score, 2.546. The upper boundary goes all the way up to positive, but we cannot type positive infinity in our calculator. Instead, we can just enter positive 99.



This answer shows a p-value of 0.0054, or 0.54%.

In Excel, first, select "Formulas", choose the "Statistical" option, and again pick "NORM.DIST". The first value we are going to input is the mean of the sample, which was 64.7, then the population mean which we are testing against, or 63.8, and finally the standard deviation, which was 2.5, divided by the square root of the sample size of 50. We can find the square root under the "Math and Trigonometry" option in "Formulas". The last value that we need to enter is "TRUE".

File:5240-pvalue15.pngFile:5241-pvalue16.png

Notice that we do not get the same p-value as the graphing calculator. In this case, since it is a right-tailed test, Excel always goes from the first part of the distribution and reads left to right. We know that the distribution is 100%, so to get that upper portion of the distribution, we have to do 100%, or 1, minus this value.

File:5242-pvalue17.pngFile:5243-pvalue18.png

We get the same p-value, which is about 0.0054, or 0.54%.

**Z-table**

of *2*

TERMS TO KNOW

P-value

The probability that the test statistic is that value or more extreme in the direction of the alternative hypothesis

Test Statistic

A measurement, in standardized units, of how far a sample statistic is from the assumed parameter if the null hypothesis is true