

Z-Test for Population Means

by Sophia Tutorial



WHAT'S COVERED

This tutorial will cover how to perform a z-test for population means. Our discussion breaks down as follows:

1. Calculating a Z-Test for Population Means
2. Conducting a Z-Test for Population Means

1. Calculating a Z-Test for Population Means

A **z-test for population means** is a type of hypothesis test that compares a hypothesized mean from the null hypothesis to a sample mean. This can be used with quantitative data and when the population standard deviation is known.

This type of z-test is not done often because it is unlikely that we would know the population standard deviation without knowing the population mean.

When calculating a z-test for population means, you need the following information:

- Population mean (μ)
- Population standard deviation (σ)
- Sample mean (\bar{x})
- Sample size (n)

This information will be plugged into the formula for a z-statistic of population means:



FORMULA

Z-Statistic for Population Means

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

IN CONTEXT

The average weight of newborn babies is 7.2 pounds, with a standard deviation of 1.1 pounds. A

local hospital has recorded the weights of all 285 babies born in a month, and the average weight was 6.9 pounds.

Find the z-test statistic for this data set.

We know the average weight is 7.2 pounds, with a standard deviation of 1.1 pounds. Because we know the population standard deviation and it is quantitative data, we can use the normal distribution and find a z-score. We also know the average weight was 6.9 pounds. We can plug this information into the following formula to calculate our z-test statistic:

$$\bar{x} = \text{sample mean} = 6.9$$

$$\mu = \text{population mean} = 7.2$$

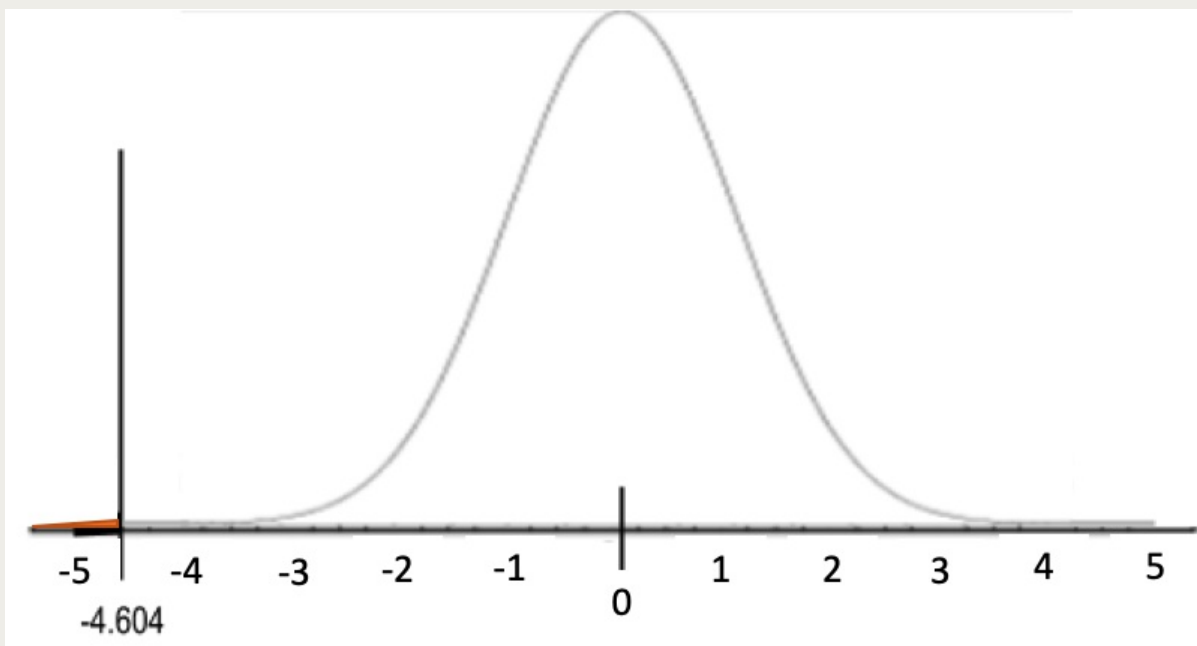
$$\sigma = \text{population standard deviation} = 1.1$$

$$n = \text{sample size} = 285$$

$$z = \frac{\frac{\bar{x} - \mu}{\sigma}}{\sqrt{n}} = \frac{6.9 - 7.2}{\frac{1.1}{\sqrt{285}}} = \frac{-0.3}{\frac{1.1}{\sqrt{285}}} = \frac{-0.3}{\frac{1.1}{16.88}} = \frac{-0.3}{0.6516} = -4.604$$

We have 6.9, which is our sample mean, minus the population mean of 7.2, divided by the population standard deviation, 1.1, divided by the square root of our sample size, 285. This gives us a z-test statistic of negative 4.604. We should expect to get a negative z-score because our sample mean was less than the population mean.

If we were to put this on a normal distribution, it's centered at the population mean, which is 7.2 pounds. The average weight of the babies at the hospital was 6.9, which is less than 7.2 pounds, so it should fall in the lower part of our distribution.



The corresponding z-score is all the way down at the negative 4.604. At this hospital, the average weight of the babies was definitely far below the average weights of the babies of the population.



Technology is often used when conducting a z-test for population means.



Z-Test for Population Means

A hypothesis test that compares a hypothesized mean from the null hypothesis to a sample mean, when the population standard deviation is known.

2. Conducting a Z-Test for Population Means

There are four parts to running any hypothesis test. This is regardless of the type of tests that you use.



Step 1: State the null and alternative hypotheses.

Step 2: Check the conditions necessary in order to actually perform the inference that you're trying to do.

Step 3: Calculate the test statistic--in this case, a z-statistic--and calculate the p-value based on the normal sampling distribution.

Step 4: Compare your test statistic to your chosen critical value or your p-value to our chosen significance level. Those are both acceptable approaches. Based on how they compare, state a decision regarding the null hypothesis. Circle it back around to the null hypothesis and decide if it supports the null hypothesis or refutes the hypothesis. Make a decision to either reject or fail to reject it based on your evidence. It should also be in the context of the problem.

 **EXAMPLE** Consider the following problem:

According to their packaging, the standard bag of M&M's candies contains 47.9 grams. Suppose you take 14 bags at random and weigh them. The sample mean is 48.07 grams.

48.2	48.4	47.0	47.3	47.9	48.5	49.0
48.3	48.0	47.9	48.7	48.8	47.4	47.6

Assuming the distribution of bag weights is approximately normal, and the standard deviation of all M&M's bags is 0.22 grams, is this evidence that the bags do not contain the claimed amount of 47.9 grams in each bag?

This could mean that it's either higher than 47.9 grams or lower than 47.9 grams. If you take a look, some of the weights in the sample are fairly off, some by almost a full gram.

You are also assuming that you know the standard deviation of all M&M's bags, which is not always a reasonable assumption, but is for this example.

Let's walk through each of the steps of running a hypothesis test with our M&M's example.

STEP BY STEP

Step 1: State the null and alternative hypotheses.

For this problem, the null hypothesis is that the M&M's bags are doing exactly what you thought they would do. The mean of all M&M's bags is the 47.9 grams that was claimed.

The alternative hypothesis is that they're not that number. This is going to be a two-sided test based on this "not equal to" symbol.

You should also state what your alpha level or significance level is going to be. By stating that alpha equals 0.05, which is the most common significance level, you are saying if the p-value is less than 0.05, reject the null hypothesis. If this is above 0.05, you should fail to reject it.

$H_0: \mu = 47.9$; The mean of all M&M's bags is 47.9 g

$H_a: \mu \neq 47.9$; The mean of all M&M's bags is not 47.9 g

$\alpha = 0.05$

Step 2: Check the conditions necessary for inference.

Look at the conditions necessary for inference on a population mean.

Criteria	Description
Randomness	How were the data collected? The randomness should be stated somewhere in the problem. Think about the way the data was collected.
Independence	Population $\geq 10n$ You want to make sure that the population is at least 10 times as large as the sample size. This was your workaround for independence. If the population is sufficiently large, then taking out the number of bags that you took doesn't make a huge difference.
Normality	$n \geq 30$ or normal parent distribution There are two ways to verify normality. Either the parent distribution has to be normal or the central limit theorem is going to have to apply. The central limit theorem says that for most distributions, when the sample size is greater than 30, the sampling distribution will be approximately normal.

Going back to the M&M's example:

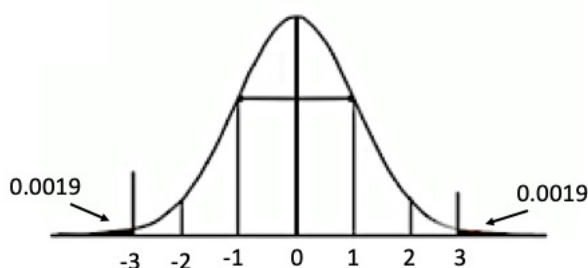
- Randomness: In the problem, it does say the bags were randomly selected. So, thinking about the way that the data was collected in the problem is important.
- Independence: We can also assume there are at least 140 bags of M&M's, which is a reasonable assumption. Why 140? Because there were 14 bags in our sample. So you're going to assume that the population of all bags of M&M's is at least 10 times that size.
- Normality: Finally, the distribution of bag weights is in fact approximately normal as stated in the problem.

Step 3: Calculate the test statistic and calculate the p-value based on the normal sampling distribution.

In this problem, your test statistic is going to be a z-statistic. How is this done? Take the sample mean minus the hypothesized population mean of 47.9 from the null hypothesis, and divide by the standard error, which is the standard deviation of the population divided by the square root of sample size.

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{48.07 - 47.9}{\frac{0.22}{\sqrt{14}}} = +2.89$$

When you do all of that and input the numbers, you get a z-statistic of positive 2.89. Look at where that lies on the normal distribution that you're using. A z-statistic of 2.89 on the standard normal distribution centered at 0 is between two and three standard deviations above the mean. Because this is a two-sided test, find the probability that your z-statistic is above positive 2.89 and the probability that it's below negative 2.89.



Use a z-table to find the probability of 0.0019 for a z-score of -2.89.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048

That probability, when doubled, gives us 0.0038.

$$\begin{aligned} 2[P(z < -2.89)] &= 2(0.0019) \\ &= 0.0038 \end{aligned}$$

Step 4: Compare your test statistic to your chosen critical value or your p-value to our chosen significance level.

This actually contains three parts: the comparison, the decision, and the conclusion. Since your p-value of 0.0038 is less than the significance level of 0.05, your decision is to reject the null hypothesis. There is evidence to conclude that the M&M's bags are not filled to a mean of 47.9 grams.

- Comparison: $0.0038 < 0.05$
- Decision: Reject the null hypothesis
- Conclusion: There is evidence to conclude that the m&m bags are not filled to the mean of 47.9 grams.



SUMMARY

The steps in any hypothesis test--not just a z-test for population means--are the same. First, state the null and alternative hypotheses both in symbols and in words. Second, state and verify the conditions necessary for inference. Third, calculate the test statistic from your statistics that you have and calculate its p-value. Finally, compare your p-value to the alpha level that you've chosen, or your test statistic to the critical value, and make a decision about the null hypothesis. State your conclusion in the context of the problem. In a z-test for population means, the population standard deviation must be known. That is not very common. You'll have other ways to address it when you don't know the population standard deviation. Since the test statistic is a z, this focuses on a z-test for population means.

Good luck!

Source: Adapted from Sophia tutorial by Jonathan Osters.



TERMS TO KNOW

Z-Test for Population Means

A hypothesis test that compares a hypothesized mean from the null hypothesis to a sample mean, when the population standard deviation is known.



FORMULAS TO KNOW

Z-Statistic for Population Means

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$