Z Test and Test on Proportions

What Is a Z-Test?

- A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large.

- The test statistic is assumed to have a normal distribution, and nuisance parameters such as standard deviation should be known in order for an accurate z-test to be performed.

z-Test

- A z-test is a statistical test to determine whether two population means are different when the variances are known and the sample size is large.
- A z-test is a hypothesis test in which the z-statistic follows a normal distribution.
- A z-statistic, or z-score, is a number representing the result from the z-test.
- Z-tests are closely related to t-tests, but t-tests are best performed when an experiment has a small sample size.
- Z-tests assume the standard deviation is known, while t-tests assume it is unknown.

- The z-test is also a hypothesis test in which the z-statistic follows a normal distribution.
- The z-test is best used for greater-than-30 samples because, under the central limit theorem, as the number of samples gets larger, the samples are considered to be approximately normally distributed.
- When conducting a z-test, the null and alternative hypotheses, alpha and z-score should be stated
    - A z-statistic, or z-score, is a number representing how many standard deviations above or below the mean population a score derived from a z-test is.
- Examples of tests that can be conducted as z-tests include a one-sample location test, a two-sample location test, a paired difference test, and a maximum likelihood estimate.
- The z test can be performed on one sample, two samples, or on proportions for hypothesis testing. It checks if the means of two large samples are different or not when the population variance is known
- A z test can further be classified into left-tailed, right-tailed, and two-tailed hypothesis tests depending upon the parameters of the data.

One-Sample Z Test

- A one-sample z test is used to check if there is a difference between the sample mean and the population mean when the population standard deviation is known. The formula for the z test statistic is given as follows:
- $z = \frac{\overline{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$. $\overline{X}$ is the sample mean, μ is the population mean, σ is the population standard deviation and n is the sample size.

Left Tailed Test:

Null Hypothesis: $H_0 : \mu = \mu_0$

Alternate Hypothesis: $H_1 : \mu < \mu_0$

Decision Criteria: If the z statistic < z critical value then reject the null hypothesis.

Right Tailed Test:

Null Hypothesis: $H_0 : \mu = \mu_0$

Alternate Hypothesis: $H_1 : \mu > \mu_0$

Decision Criteria: If the z statistic > z critical value then reject the null hypothesis.

Two Tailed Test:

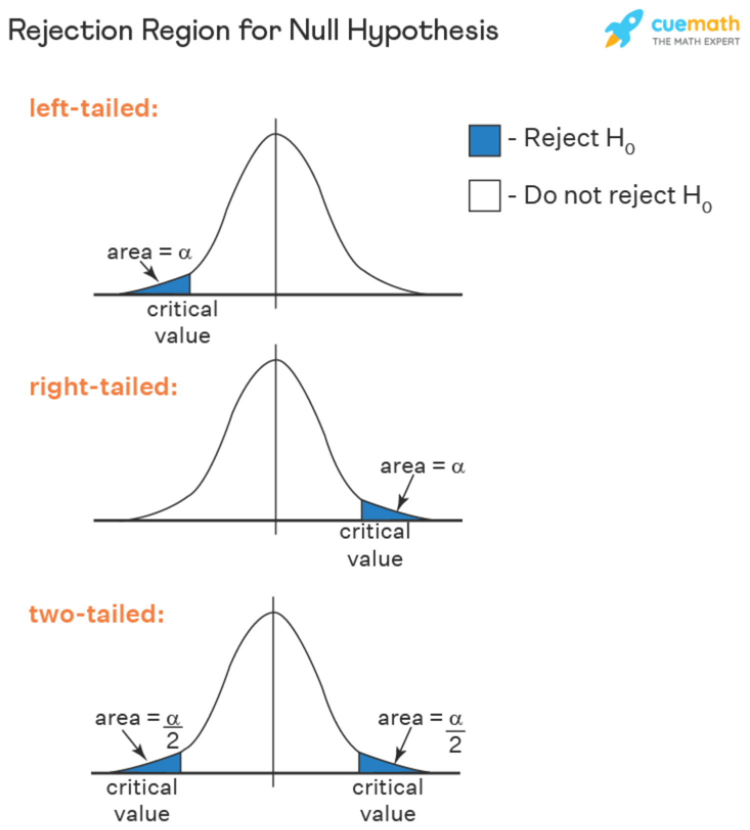
Null Hypothesis: $H_0 : \mu = \mu_0$

Alternate Hypothesis: $H_1 : \mu \neq \mu_0$

Decision Criteria: If the z statistic > z critical value then reject the null hypothesis.

Two Sample Z Test

- A two sample z test is used to check if there is a difference between the means of two samples. The z test statistic formula is given as follows:
- $z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$. $\overline{x_1}$, μ_1, σ_1^2 are the sample mean, population mean and population variance respectively for the first sample. $\overline{x_2}$, μ_2, σ_2^2 are the sample mean, population mean and population variance respectively for the second sample.
- The two-sample z test can be set up in the same way as the one-sample test. However, this test will be used to compare the means of the two samples
- For example, the null hypothesis is given as $H_0 : \mu_1 = \mu_2$



Z Test for Proportions

- A z test for proportions is used to check the difference in proportions. A z test can either be used for one proportion or two proportions. The formulas are given as follows.
- **One Proportion Z Test**
  - A one proportion z test is used when there are two groups and compares the value of an observed proportion to a theoretical one. The z test statistic for a one proportion z test is given as follows:

    $z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$. Here, p is the observed value of the proportion, p_0 is the theoretical proportion value and n is the sample size.

  - The null hypothesis is that the two proportions are the same while the alternative hypothesis is that they are not the same

Two Proportion Z Test

- A two proportion z test is conducted on two proportions to check if they are the same or not. The test statistic formula is given as follows:
- $$z = \frac{p_1 - p_2 - 0}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

  where $p = \frac{x_1 + x_2}{n_1 + n_2}$

  p_1 is the proportion of sample 1 with sample size n_1 and x_1 number of trials.

  p_2 is the proportion of sample 2 with sample size n_2 and x_2 number of trials.

How to Calculate Z Test Statistic?

- The most important step in calculating the z test statistic is to interpret the problem correctly.
- It is necessary to determine which tailed test needs to be conducted and what type of test does the z statistic belong to.
- Suppose a teacher claims that his section's students will score higher than his colleague's section. The mean score is 22.1 for 60 students belonging to his section with a standard deviation of 4.8.
- For his colleague's section, the mean score is 18.8 for 40 students and the standard deviation is 8.1.
- Test his claim at $\alpha = 0.05$.
- The steps to calculate the z test statistic are as follows:

- Identify the type of test. In this example, the means of two populations have to be compared in one direction thus, the test is a right-tailed two-sample z test.
- Set up the hypotheses. H_0: $\mu_1=\mu_2$, H_1: $\mu_1>\mu_2$
- Find the critical value at the given alpha level using the z table. The critical value is 1.645.
- Determine the z test statistic using the appropriate formula. This is given by $z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$. Substitute values in this equation. $\overline{x_1}$ = 22.1, $\sigma_1 = 4.8$, $n_1 = 60$, $\overline{x_2} = 18.8$, $\sigma_2 = 8.1$, $n_2 = 40$ and $\mu_1 - \mu_2 = 0$.

  Thus, z = 2.32
- Compare the critical value and test statistic to arrive at a conclusion. As 2.32 > 1.645 thus, the null hypothesis can be rejected. It can be concluded that there is enough evidence to support the teacher's claim that the scores of students are better in his class.

- **Example 1:** A teacher claims that the mean score of students in his class is greater than 82 with a standard deviation of 20. If a sample of 81 students was selected with a mean score of 90 then check if there is enough evidence to support this claim at a 0.05 significance level.

Solution: As the sample size is 81 and population standard deviation is known, this is an example of a right-tailed one-sample z test.

$H_0 : \mu = 82$

$H_1 : \mu > 82$

From the z table the critical value at $\alpha = 1.645$

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

$\overline{x} = 90, \mu = 82, n = 81, \sigma = 20$

$z = 3.6$

As $3.6 > 1.645$ thus, the null hypothesis is rejected and it is concluded that there is enough evidence to support the teacher's claim.

Answer: Reject the null hypothesis

- **Example 2:** An online medicine shop claims that the mean delivery time for medicines is less than 120 minutes with a standard deviation of 30 minutes. Is there enough evidence to support this claim at a 0.05 significance level if 49 orders were examined with a mean of 100 minutes?

Solution: As the sample size is 49 and population standard deviation is known, this is an example of a left-tailed one-sample z test.

$H_0 : \mu = 120$

$H_1 : \mu < 120$

From the z table the critical value at $\alpha = -1.645$. A negative sign is used as this is a left tailed test.

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

$\overline{x} = 100, \mu = 120, n = 49, \sigma = 30$

$z = -4.66$

As $-4.66 < -1.645$ thus, the null hypothesis is rejected and it is concluded that there is enough evidence to support the medicine shop's claim.

Answer: Reject the null hypothesis

- **Example 3:** A company wants to improve the quality of products by reducing defects and monitoring the efficiency of assembly lines. In assembly line A, there were 18 defects reported out of 200 samples while in line B, 25 defects out of 600 samples were noted. Is there a difference in the procedures at a 0.05 alpha level?
-

Solution: This is an example of a two-tailed two proportion z test.

H_0: The two proportions are the same.

H_1: The two proportions are not the same.

As this is a two-tailed test the alpha level needs to be divided by 2 to get 0.025.

Using this, the critical value from the z table is 1.96.

$n_1 = 200$, $n_2 = 600$

$p_1 = 18 / 200 = 0.09$

$p_2 = 25 / 600 = 0.0416$

$p = (18 + 25) / (200 + 600) = 0.0537$

$$z = \frac{p_1 - p_2 - 0}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 2.62$$

As 2.62 > 1.96 thus, the null hypothesis is rejected and it is concluded that there is a significant difference between the two lines.

Answer: Reject the null hypothesis