<u>Hypothesis Testing</u>

Hypothesis is a statement, the validity of which is to be tested. In statistics, a hypothesis is a claim or statement about a property of a population.

Null Hypothesis, H_0**:** The null hypothesis is a statement about the value of a population parameter (such as a mean), and it must contain the condition of equality and must be written with one of the following symbols:

$$=, \quad \geq, \quad \leq$$

For the mean, the null hypothesis will be stated in one of the following three possible forms:

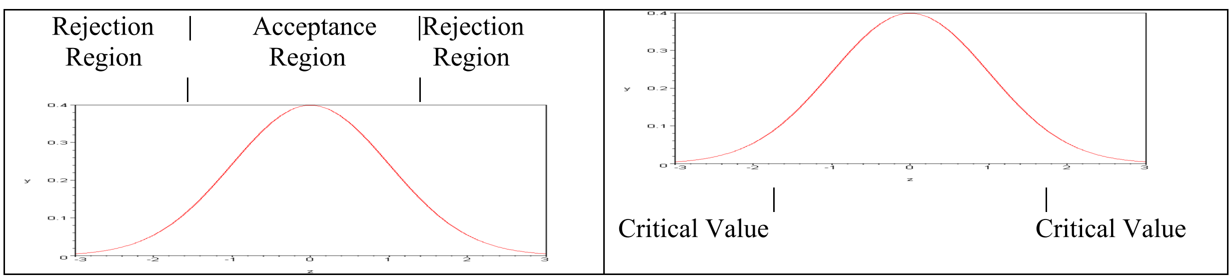
$H_0 : \mu =$ same value $\qquad H_0 : \mu \geq$ same value $\qquad H_0 : \mu \leq$ same value

Alternative Hypothesis, H_A**:** The alternate hypothesis is the statement that must be true if the null hypothesis is false.

For the mean, the alternative hypothesis will be stated in one of the following three possible forms:

$H_A : \mu \neq$ same value $\qquad H_A : \mu <$ same value $\qquad H_A : \mu >$ same value

Critical Region: The critical region is the set of all values of the test statistic that cause us to reject the null hypothesis.



Critical Value: The critical value that separates the critical region (where we reject the null hypothesis) from the values of the test statistics that do not lead to reject the null hypothesis.

Significant Level: The significant level denoted by α, is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true.

Confidence Level	α	**Critical Value,**	
		$Z_{\alpha/2}$	Z_{α}
99%	**0.01**	**2.575**	**2.33**
95%	**0.05**	**1.96**	**1.65**
90%	**0. 1**	**1.646**	**1.28**

Type I Errors: When H_0 is true, we reject H_0.

Type II Errors: When H_0 is false, we do not reject H_0.

The Central Limit Theorem: The central limit theorem describes the distribution of the sample statistic $\bar{X}$ when samples are sufficient large ($n \geq 30$).

The *z*-test: The *z*-test is a statistical test for the mean of a population. It can be used when $n \geq 30$, or when the population is normally distributed and σ is known.
The formula for the *z*-test is

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

where $\bar{X}$ = sample maen

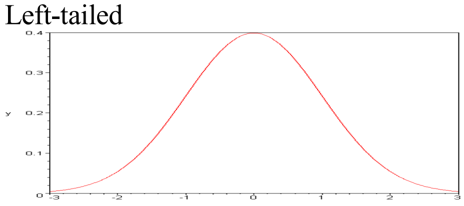
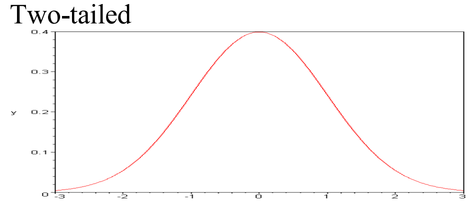
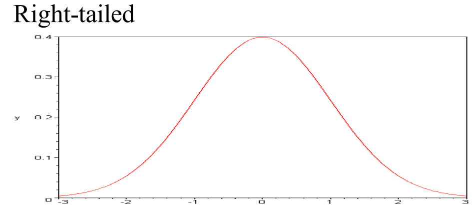
μ = hypothesized populatin mean

σ = population standard deviation

n = sample size

For the *z*-test, the observed value is the value of the sample mean. The expected value is the value of the population mean, assuming the null hypothesis value is true. The denominator $\sigma / \sqrt{n}$ is the standard error of the mean.

Critical Regions for $H_0 : \mu = k$ ($\bar{x}$ Values are under the shaded regions)

$H_A : \mu < k$ Left-tailed  ↑ k Critical region	$H_A : \mu \neq k$ Two-tailed  ↑ k ↑ Critical region Critical region	$H_A : \mu > k$ Right-tailed  k ↑ Critical region
One-Tailed Left Test $H_0 :\ \geq$ $H_A :\ <$	Two-Tailed Test $H_0 :\ =$ $H_A :\ \neq$	One-Tailed Right Test $H_0 :\ \leq$ $H_A :\ >$

The *t*-test: The *t*-test is a statistical test for the mean of a population. It can be used when the population is normally or approximately normally distributed and σ **is unknown** and $n < 30$. The formula for the *t*-test is

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

where $\bar{X}$ = sample maen

μ = hypothesized populatin mean

s = sample standard deviation

n = sample size

The degrees of freedom are $d.f. = n - 1$.

Meaning of Accepting:

In most statistical applications, the level of significance is specified to be $\alpha = 0.05$ or $\alpha = 0.01$, although other values can be used. If $\alpha = 0.05$, then we say we are using a 5% level of significance. This means that in 100 similar situations, null hypothesis will be reject 5 times, on the average, when it should not have been rejected.

When we accept (or fail to reject) the null hypothesis, we should understand that we are not proving the null hypothesis. We are saying only that the sample evidence (data) is not strong enough to justify rejection of the null hypothesis.