Test of hypotheses or Test of significance

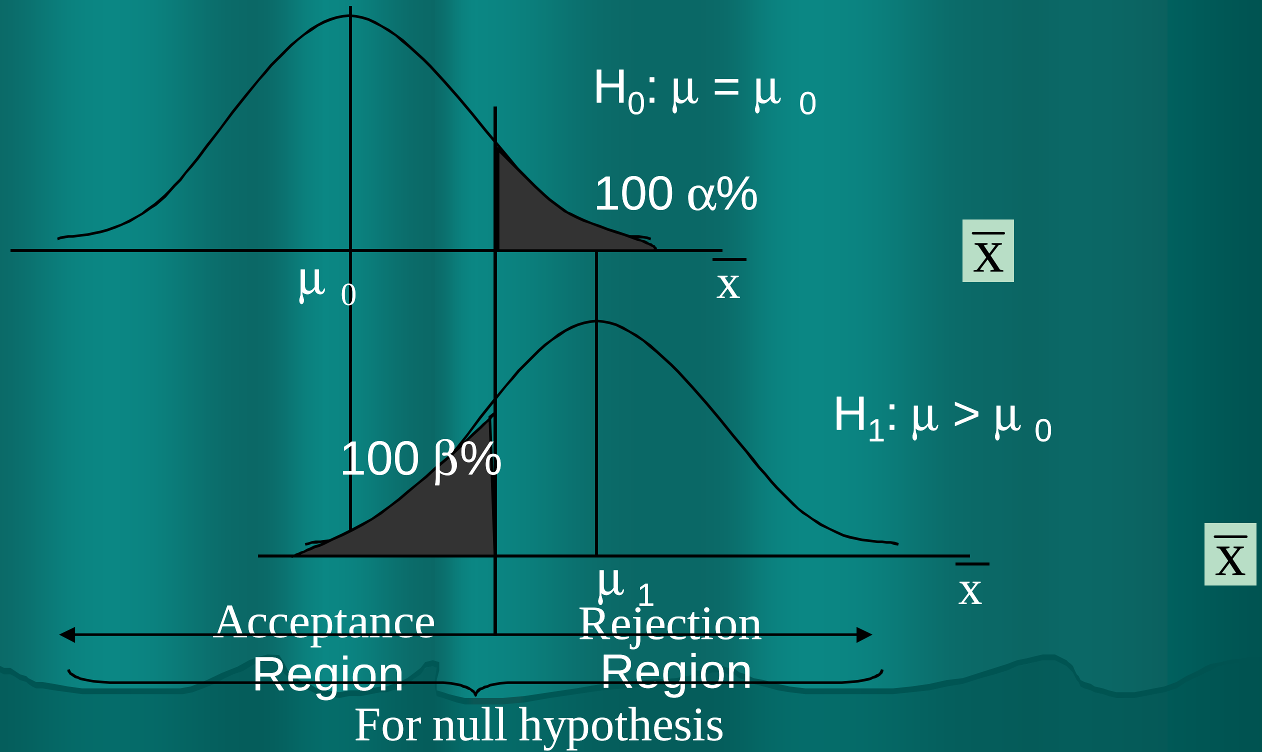
- The testing of statistical hypotheses is perhaps the most important area of decision theory. The purpose of this test is to aid the clinical, or researcher in reaching a conclusion concerning a population by examining a sample from the population. For example we may wish to decide one the basis of a sample data whether a new serum is really effective curing a disease.
- <u>Statistical hypothesis</u>: is an assumption or statement, which may or may not be true, concerning one or more populations. There are two types of this hypothesis:
- A). *Null Hypothesis* (H_o): is the hypothesis to be test and sometimes is referred to as a hypothesis of no difference.
- B). *Alternative hypothesis* (H_1 *or* H_A)*:* is a statement of what we will believe is true if our sample data cause us to reject the null hypothesis.

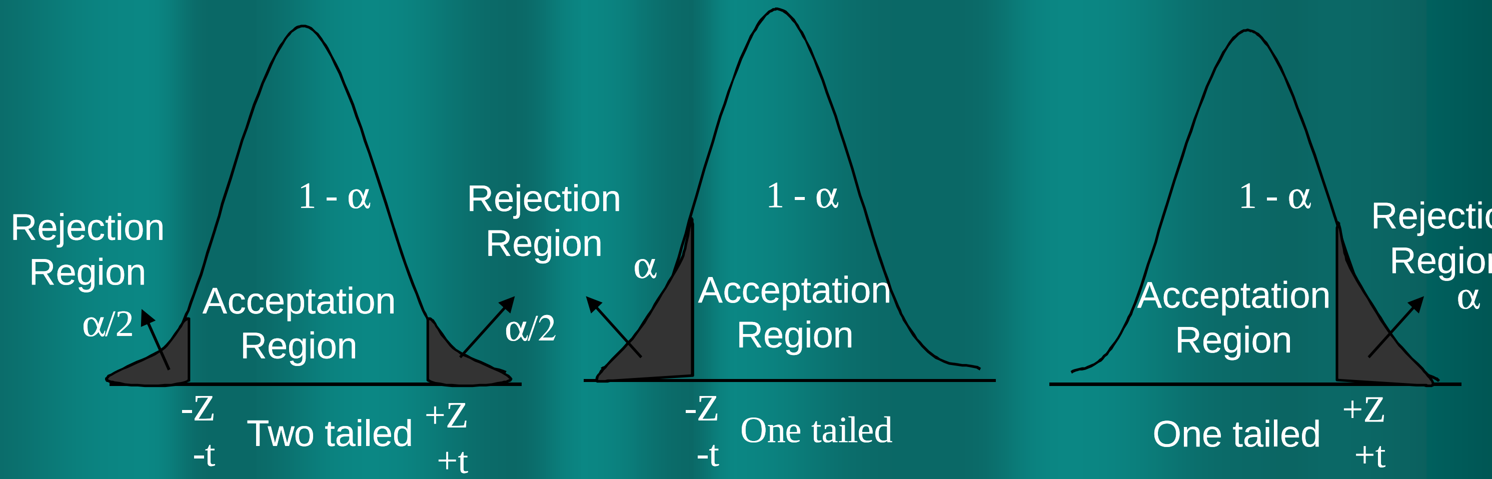
Types of Error:

- 1). Type I error *or level of significance* (ɑ)*:* is the probability of rejecting a true null hypothesis. The more frequently encountered values of ɑ are 0.10, *0.05, and 0.01*.

- 2). Type II error (ß): is the probability of accepting a null hypothesis when it should be rejected.

- Power test (1 - ß): is the probability of rejecting a null hypothesis when it should be false.





- *The steps for testing a hypotheses concerning a population parameters may be summarized as follows:*
- 1- Select the appropriate hypotheses.
- 2- Select the appropriate test statistics and establish the critical region.
- 3- Conclusion and Decision;
- A- Statistical,
- B- Clinical.

Z Critical Values

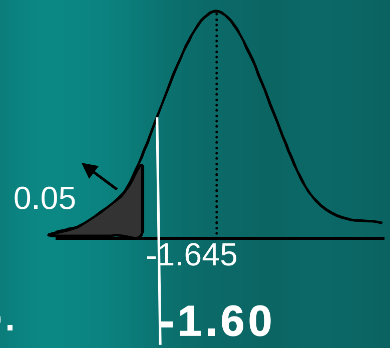
α	0.1	0.05	0.01
Two tailed test	±1.645	±1.96	±2.58
One tailed test	±1.28	±1.645	±2.33

Two-sided test: No a priori reason 1 group should have stronger effect. Used for most tests.
One-sided test: Specific interest in only one direction. Not scientifically relevant/interesting if reverse situation true.

- Sample from Normally Distributed Population, when σ^2 is Known:
- Example: A nutrition survey was conducted in developing country. A sample of 500 rural adults reported a mean minimum daily caloric intake of 1985 and a standard deviation of population is 210. Can one conclude from these data that the population mean at 0.05 level of significance ($\alpha = 0.05$) is;
- 1- Less than 2000?
- 2- Not equal 2000?
  - 1- H_0: $\mu = 2000$, H_1: $\mu < 2000$

$$Z = \frac{\bar{x} - m_0}{s / \sqrt{n}}$$

$$Z = \frac{1985 - 2000}{210 / \sqrt{500}} = -1.60$$



- The critical value of test statistics is ± 1.645
- Statistical decision: Accept H_0 because $Z > -1.645$.
- Clinical decision: Conclude that, on the bases of these data, there is an indication that the mean of daily caloric intake is equal 2000. ($P>0.05$)
- or exactly $P = 0.5 - 0.4452 = 0.0548$.

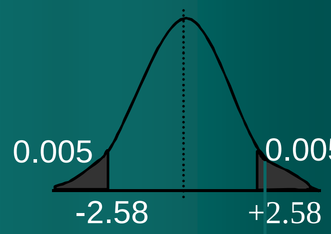
Common misinterpretation of a P value

Many people misunderstand P values. If the P value is 0.05, that means that there is a 5% chance of observing a difference as large as you observed even if the two population means are identical (the null hypothesis is true). It is tempting to conclude, therefore, that there is a 95% chance that the difference you observed reflects a real difference between populations and a 5% chance that the difference is due to chance. However, this would be an incorrect conclusion. *What you can say is that random sampling from identical populations would lead to a difference smaller than you observed in 95% of experiments and larger than you observed in 5% of experiments. This distinction may be more clear after you read*

- Sampling from Normally Distributed Populations: Population Variances, (σ_1 and σ_2) are Known.
- Example: In a large hospital for the treatment of the mentally retarded, a sample of 12 individuals with mongolism yielded a mean serum uric acid value of 4.5 mg/100ml. Another sample of size 15 normal individuals from other hospital of the same age and sex were found to have a mean value of 3.4 mg/100ml.If it is reasonable to assume that the two populations of values are normally distributed with variances equal to 1. Do these data provide sufficient evidence to indicate a difference in mean serum uric acid levels between normal individuals and individuals with mongolism? Let α = 0.01.
  - H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\left(\frac{s_1^2}{n_1}\right) + \left(\frac{s_2^2}{n_2}\right)}}$$

$$Z = \frac{(4.5 - 3.4) - 0}{\sqrt{\frac{1}{12} + \frac{1}{15}}} = 2.82$$



2.82

- The critical value of test statistics is ± 2.58
- *Statistical decision*: Reject H_0, since Z > 2.58.
- *Clinical decision*: Conclude that, on the bases of these data, there is an indication that the means of uric acid are not equal. (P<0.01), or exactly 0.5-0.4976=0.0024.