**HYPOTHESIS TESTING**

What is a hypothesis? An assertive statement which can be tested with data

Types of hypotheses – Equality versus Inequality

Null vs alternate hypothesis

Test of single sample against a benchmark, test of difference between two samples

How to frame hypotheses

Ho - Null hypothesis

Ha -  Alternate hypothesis

Small samples, large samples:

T-tests used for small samples, Z test for large samples (n>30)

Use 95% confidence level as a rule of thumb in making inferences in hypotheses tests. Sometimes 90% confidence level is acceptable.

t = (x-μ)/standard error with degrees of freedom (d.f.) = n-1

where n = number of observations

If the calculated t > tc (critical value of t from the t-tables at 95% confidence level) then we reject Null hypothesis.

Which test to use if the samples are correlated (i.e., not independent)?

If the same person answers two questions about different products, or answers the same question at two different times then the samples may not be correlated.

Use paired t-test.

**HYPOTHESIS TESTS**

EXAMPLE

*T-test for test of mean against a benchmark for a small sample.*

A supermarket chain (e.g., Albertson’s) is investigating the desirability of adding a new product. The management decides that it will keep the product on its shelf, if it can be shown that at least 100 units will be sold per week. A study was undertaken in which a random sample of ten stores put the product on their shelves and observed the sales for a week.

They obtained the following sales: 86, 97, 114, 108, 123, 93, 132, 116, 105, 120

Can you conclude with 95% confidence that the management should introduce the product?

Null Hypothesis: Ho: Average sales < 100

Alternate hypothesis: Ha: Average sales >= 100

Average sales = Σxi /n = 109.4

Standard deviation = = 14.39

Standard Error = sx = standard deviation/ = 14.39 / = 4.55

t-value = (mean-100)/Standard Error = (109.4-100)/4.55 = 2.07

Degrees of freedom = n-1 = 9

From the t-tables in stats books, we find the critical value of t at 95% confidence level and 9 degrees of freedom = 1.833

Since calculated t > tc, we reject the null hypothesis. This means that there is 5% chance that the calculated value of t would have occurred by chance if indeed sales in the store were less than or equal to 100 units per week. We are 95% confident that the sales are greater than 100 and so the store should introduce the product.

95% confidence level = mean  t\*sx = 109.41.833\*4.55 = {101, 117}

NOTE: For large samples (that is n>30) we use the z-test (z-tables) instead of t- tables.

NOTE: You can use the above test to check if the means from two different populations are equal or not.

Ho : μ1 = μ2

NOTE: If you are testing difference between two proportions, then the formula for computing t will be different.