**Essential Information on Hypothesis Tests**

This document contains basic information on how to perform the following three hypothesis tests: paired t test, two-sample t-test, and rank sum test. The information in this document is a very brief summary of the relevant chapters from the 6th edition of Probability and Statistics for Engineers and Scientists by Jay L. Devore.

For those of you who have a copy of Devore’s book, you should read the sections on these tests (9.2-9.3 and 15.2 in 6th edition). For those who do not have a copy of the book or access through a peer to a copy of the book, you may want to consider adding it to your library if you do not already have a good statistics book. Devore’s book serves as an excellent reference text on statistics for engineers, and does not assume prior knowledge of the material. I also have a copy of the book that I would be willing to loan on reserve for periods of no more than several hours to students who do not have access to it.

**Paired t Test**

Use this test when you are comparing results from two algorithms with identical number of runs and identical initial solutions.

**Assumptions**

The data consist of *n* independently selected pairs (X1,Y1), (X2,Y2), …, (Xn, Yn), with E(Xi) = μx and E(Yi) = μY. Let D1 = X1 – Y1, D2 = X2 – Y2, Dn = Xn – Yn, so the Di’s are the differences within pairs. Then the Di’s are assumed to be normally distributed with mean value μD and variance σD2 (this is usually a consequence of X and Y themselves being normally distributed).

**Null Hypothesis**



**Test Statistic t**



Under the null hypothesis, the test statistic t has a Student t distribution with n-1 degrees of freedom

**Alternative Hypotheses Rejection Region for Level** α **Test**

 (two-tailed test)  or 

 (lower-tailed test) 

 (upper-tailed test) 

where  given that T has a Student t distribution with *n* – 1 degrees of freedom.

**Computing t**α**, n-1**

In Matlab, use the tinv function, which has two arguments: P and V. P is the non exeedance probability and is equal to 1 – α, and V is the degrees of freedom (n-1 for this test). For example, if α = 5%, and n = 10, then you would compute ta,n-1 as follows:

t\_alpha=tinv(0.95,9). To compute ta/2,n-1, use tinv(0.975,9).

**Computing P-values**

The p-value is the smallest possible value that you could assign to α and still reject the null hypothesis. If your t = 2.73 and you have n =10, you would compute your p-value for a one-sided test in Matlab as follows: pval = tcdf(2.73,9). For a two sided test, pval = 2\*tcdf(2.73,9). There are also tables in any stats book where you could look up the pvalue.

**Two-Sample t Test**

Use this test when you are comparing two algorithms with different initial conditions and/or different sample sizes (e.g. number of trials).

**Assumptions**

Both populations are normal, so that X1, X2, …, Xm is a random sample from a normal distribution and so is Y1, Y2, …, Yn (with the X’s and Y’s independent of one another).

**Null Hypothesis**



**Test Statistic t**



Under the null hypothesis, the test statistic t has a Student t distribution with υ degrees of freedom, with υ estimated by the following equation:



Note that if you assume that the variances of X and Y are equal, that υ = m + n – 2. Beware that some statistical packages (including Matlab) will assume this unless you specify otherwise.

**Alternative Hypotheses Rejection Region for Level** α **Test**

 (two-tailed test) t ≥ tα/2, υ or t ≤ - tα/2, υ

 (lower-tailed test) t ≤ - tα, υ

 (upper-tailed test) t ≥ tα, υ

where α = P(T ≤ tα) given that T has a Student t distribution with υ degrees of freedom.

**Computing t**α,υ **and Computing P-values**

See procedures for paired t test.

**Rank Sum Test**

This test procedure is distribution free, which means that it will have the same desired level of statistical significance regardless of the underlying distributions of the data. Use this test when at least one of your two sample sizes is small or if you have doubts that either set of data is at least approximately normal.

**Assumptions**

X1, X2, …, Xm and Y1, Y2, …, Yn are two independent random samples from continuous distributions with means μX and μY, respectively. The X and Y distributions have the same shape and spread, the only possible difference between the two being the values of μX and μY.

**Null Hypothesis**



**Test Statistic t**

where r*i* = rank of (*x*i - Δo) in the combined sample of m + n (*x* - Δo)’s and *y*’s

**Alternative Hypotheses Rejection Region for Level** α **Test**

 (two-tailed test) *w* ≥ *c* or *w* ≤ *m*(*m* + *n* + 1) – *c*

 (lower-tailed test) *w* ≤ *m*(*m* + *n* + *1*) – *c*1

 (upper-tailed test) *t* ≥ *c*1

where P(W ≥ c1 | Ho true) ≈ α, and P(W ≥ c | Ho true) ≈ α/2

Note that W has a discrete distribution. If, however, both *m* and *n* exceed 8, the distribution of *W* can be approximated with a normal distribution. In this case, compute the following test statistic, Z, which has a standard normal distribution under the null hypothesis:



**Alternative Hypotheses Rejection Region for Level** α **Test**

 (two-tailed test) z ≥ zα/2 or z ≤ - zα/2

 (lower-tailed test) z ≤ - zα

 (upper-tailed test) z ≥ zα

where α = P(Z ≤ zα) given that Z has a Standard Normal distribution.

**Computing Critical Values and Computing P-values**

For small values of *m* and *n* you must build the discrete distribution of W and determine the values of c and c1 based on that distribution. You will also use the discrete distribution to determine the p-values.

For the normal approximation, compute *z*α in Matlab with the function norminv. If α = 0.05, z\_alpha = norminv(0.95). To compute *z*α*/2*, use norminv(0.975). The p-value of the test statistic Z should be 2\*[1-Φ(|Z|)] for a two-tailed test or 1-Φ(|Z|) for a one-tailed test. There are also standard normal tables in just about any stats book.