

Colorado Procedure 14-12

Standard Practice for

F and t-test Statistical Method

1. SCOPE

1.1 Use this procedure as required by the project specifications to provide a method of comparing two independent data sets of multiple test results (e.g. Contractor’s Quality Control and the Department’s Acceptance test results, Contractor’s Quality Control and CDOT Verification test results, CDOT and Contractor’s Verification test results, etc.) to determine if the materials tested come from the same population. This statistical procedure employs estimation and hypothesis testing using F-test and t-tests to make the comparisons.

1.2 Compare two populations that are assumed normally distributed by calculating and comparing the population means (arithmetic averages) and variances (standard deviation x standard deviation). The *F-test compares the population variances* while the *t-test compares the population means*.

1.3 Select all samples using random or stratified random procedures. Perform all testing and measuring in accordance with standard acceptable practices. All sampling and testing will be in accordance with applicable specifications.

1.4 The following sections provide reference materials, the mathematical equations, combined manual and computer-assisted calculations, and completely automated procedure using computer software to calculate the F-test and t-test statistics.

2. REFERENCED DOCUMENTS

- 2.1 *Colorado Procedures:*
 CP 41 Sampling Hot Mix Asphalt
 CP 55 Reducing Field Samples of Hot Mix Asphalt to Testing Size.
 CP 61 Sampling Freshly Mixed Concrete

- 2.2 *Other References:*
 AASHTO R 9 “Standard Practice for Acceptance Sampling Plans for Highway Construction”.
 Implementation Manual for Quality Assurance, 1996, AASHTO Highway Subcommittee on Construction.
 Statistical Reasoning, 1985, Gary Smith.
 Probability and Statistics, 1975, Murray R. Spiegel.
 Elementary Statistics, 1976, Robert R. Johnson.
 Probability and Statistics for Engineers and Scientists, 1972, Ronald E. Walpole and Raymond H. Myers.

3. DEFINITION OF TERMS, SYMBOLS, AND EQUATIONS

3.1 Equations and Definitions

$$\bar{X} = \sum x_i / n \tag{Eq. 3.1}$$

Where \bar{X} is the sample mean or average.

- Σ = summation symbol
- x_i = any individual test value (i = 1, 2, 3, ...n)
- n = total number of tests (sample size)

$$S = \sqrt{\frac{\Sigma(X_i - \bar{X})^2}{n - 1}} \tag{Eq. 3.2}$$

Where S is the standard deviation

n-1 = degree of freedom

$$V = S^2 \quad \text{Eq. 3.3}$$

Where V is the sample variance

$$F = (V_1/V_2) \text{ or } (V_2/V_1) \quad \text{Eq. 3.4}$$

Where F is the ratio of the variance from each data set (larger variance divided by the smaller variance) depending on which ratio yields a value equal to or greater than 1. This is called the F distribution (aka F-test).

$$S_p = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 + n_2 - 2)}} \quad \text{Eq. 3.5}$$

Where S_p is a weighted average of the sample variances each weighted by the degrees of freedom.

$$t = (\bar{X} - \mu) / (S_p / \sqrt{n_c + n_a})^{1/2} \quad \text{Eq. 3.6}$$

Where t is the statistic used to compare the mean of a sampled population to some fixed, known value (aka t-test).

μ = Mean from the contractor's population

S_p = Variance of the pooled data

n_c = Number of tests in the contractor's population

n_a = Number of tests in CDOT's population

α = level of significance or critical region. This is the probability of incorrectly deciding the data sets are different when they actually come from the same population. In either the construction or the manufacturing industry, α is the risk of rejecting a good material or product. The critical region, α in the F and t probability distribution curves is equivalent to the rejection area. Since the total area bounded by either the F or t distribution curve is equal to 1, the acceptance region is $1 - \alpha$. For example, when $\alpha = 0.05$, there is a probability of 95 percent that the two data sets are from the same population.

$\alpha_{critical}$ = is the maximum value in the F distribution and t statistic for the level of significance and the degree of freedom for the contractor and CDOT at which the comparison between the two sample populations should not be exceeded. If the t statistic is less than the critical value, the hypothesis is that they came from the same population.

The two-tailed test determines if the population parameters (variances or averages) are either equal or not equal. All the values of α obtained from this procedure are based on the two-tails of the distribution curve.

4. SUMMARY OF METHOD

4.1 The method involves calculating sample statistics from three or more representative measurements, test results, or values, for each specified element in a lot or sample. The specimen will be independent samples. The statistical variables to be calculated include the mean, standard deviation, variance, F and t values, and the $\alpha_{critical}$ value. The following sections summarize the F-test and t-test method to be employed in this procedure.

4.2 Determine the appropriate population parameters and sample statistics to be used in estimation and hypothesis testing (F & t-tests). For the F-test calculation, test the assumption that the population variances are equal against the assumption that they are not equal (use a two-tailed F-test). For the t-test calculation, assume the population variances are equal and test the assumption that the population means are equal against the assumption that they are not equal (use a two-tailed t-test).

4.3 Choose a level of significance or critical region (α) for each of the F-test and t-test calculations. AASHTO R 9 provides suggested critical values of $\alpha_{critical}$ used in the highway construction industry. CDOT typically uses $\alpha_{critical}$ values of 0.10, 0.05, 0.01, and 0.005. In this procedure, use $\alpha_{critical}$ values as specified in the project specifications.

4.4 Calculate all the required variables in the appropriate F-test and t-test equations and compare the calculated $\alpha_{critical}$ with the level of significance chosen in the previous subsection.

4.5 Conclude that the measurements, test results, or test values come from the same population if the calculated α -value is less than

the α_{critical} at the selected level of significance. Conclude that the measurements, test results, or test values do not come from the same population if the calculated α -value is less than the α_{critical} value at the selected level of significance.

5. COMPUTER-ASSISTED PROCEDURE

5.1 Any applicable computer software with statistical functions may be used to conduct F-test and t-test calculations. The Microsoft Excel statistical function FTEST can be used to calculate the α_{critical} value for the F-test while the Microsoft Excel statistical function TTEST can be used to calculate the α_{critical} value for the t-test. The FTEST function has the command format FTEST [array1, array2]. Array1 is the first data set and array2 is the second data set. The FTEST function directly calculates the two-tailed α_{critical} value. Compare this value with the selected level of significance. Conclude that the test data are from the same population if the result of the FTEST calculation is less than the selected level of significance. Proceed to conducting a t-test assuming equal population variances if the result of the FTEST calculation is less than the selected level of significance.

5.2 The Microsoft Excel TTEST function has the command format TTEST [array1, array2, tails, type]. Array1 is the first data set and

array2 is the second data set. The tails parameter specifies the number of distribution tails and type refers to the kind of t-test to perform. The type can be 1 (paired t-test), 2 (two-sample equal variance) and 3 (two-sample unequal variance). Type 3 is not used in this procedure because the test data sets are automatically concluded to be not from the same population if the sample variances are found to be unequal. The t-test directly calculates the α_{critical} value, given the required values of the variables in the TTEST function. Compare this value with the selected level of significance. Conclude that the test data are from the same population if the result of the TTEST calculation is less than the selected level of significance. Conclude that the test data are not from the same population if the result of TTEST calculation is greater than the selected level of confidence.

5.3 The Department has software to perform F-test and t-test analysis. The software calculates the F-test and t-test values and compares them with the selected level of significance. The software automatically indicates if the test data are either from the same population or not using appropriate label or designation.

6. SAMPLE CALCULATIONS F-test and t-test

6.1 Independent Samples (Non-paired Observations)

This example will demonstrate the procedures to conduct F-test and t-test calculations for independent samples.

Problem Statement:

Using the ignition furnace method to determine the asphalt content of a mix, the following test results were obtained for independent sample populations A and B:

Test Number	Population A	Population B
1	4.65	4.75
2	4.84	4.79
3	4.59	4.74
4	4.75	4.41
5	4.63	4.77
6	4.75	4.58
7	4.58	4.81
8	4.82	
9	4.86	
10	4.60	
11	4.77	
12	4.65	
13	4.80	

Using F-test and t-test, determine if sample A and sample B are from the same population.

Solution:

- a) Select the level of significance (α) at which to evaluate the F-test and t-test. Use the level specified in the project special provisions. Assuming that $\alpha = 0.01$ is specified, determine the F-value using Eq. 3.4 which comes from Eq. 3.3 (variance), Eq. 3.2 (standard deviation), and Eq. 3.1 (mean) in each data set.

	<u>Sample A</u>	<u>Sample B</u>
Arithmetic Average	4.71	4.69
Standard Deviation	0.1013	0.1457
Variance	0.010260	0.021224

F-value (larger variance is divided by smaller variance, $0.021224 / 0.010260$) 2.07
 Degrees of freedom, n-1, (numerator, 7-1) 6
 Degrees of freedom, n-1, (denominator, 13-1) 12

- b) From AASHTO R 9, Table X2.1, the $\alpha_{critical}$ value for the F-test using 6 degrees of freedom in the numerator and 13 degrees of freedom in the denominator translates into $\alpha_{critical}$ of 5.76.

- c) Compare this calculated $\alpha_{critical}$ value with the F-value. Since the $\alpha_{critical}$ is greater than the calculated value ($5.76 > 2.07$), **conclude that the sample variances are equal and proceed to conducting a t-test.**

- d) Calculate the arithmetic averages (\bar{X}_1 and \bar{X}_2) and variances (S_1^2 and S_2^2) for each data set. Calculate the pooled variance, S_p for both data sets using Eq. 3.5. Calculate the absolute t-value using Eq. 3.6. The sample size for sample A is $n = 13$ and the sample size for sample B is $n = 7$.

Arithmetic Average (\bar{X}_1 or \bar{X}_2)	4.71	4.69
Variance (S_1^2 or S_2^2)	0.01026	0.02122
Pooled Variance (S_p)	0.11796	

Calculating the absolute value of t yields: $(4.69 - 4.71) / ((0.11796/14) + (0.11796/7))^{1/2} = 0.126$

- e) From AASHTO R 9, Table X2.2, the $\alpha_{critical}$ value is $= 2.878$. Since the $\alpha_{critical}$ is greater than the calculated value ($2.878 > 0.126$), **conclude that the two data sets are from the same population.**