1. Introduction

Welcome to the series of E-learning modules on Confidence Intervals for the ratio of variances of two independent normal populations. In this module, we are going to cover the Interval Estimation - procedure to estimate the ratio of two population variances or standard deviations when the population means are known and unknown.

By the end of this session, you will be able to:

- Explain the role of F distribution in an Interval Estimation
- Explain the confidence interval for the ratio of variances or standard deviations of two independent normal populations when the population means are known
- Explain the confidence interval for the ratio of variances or standard deviations of two independent normal populations when the population means are unknown

We saw several situations in which we wanted to compare the means of two different populations. Recall, we did this by looking at the difference of the means of the two samples drawn from those populations.
Like the importance of a single variance differences in variances to an experimenter, it also important to compare the two population variances.

We might need to compare the precision of one measuring device with that of another and also the stability of one manufacturing process with that of another or even the variability in the grading procedure of one college professor with that of another.

However, when we want to compare the variances of two populations rather than looking at the difference of the two sample variances, it turns out to be more convenient if we look at their ratio.
For example, a prominent sociologist at a large Midwestern University wants to estimate ratio of variability in the earnings of college graduates to the earnings of those who did not attend college.
Hence, a way to compare the variances of two normally distributed populations is to use the variance ratio, sigma 1 square by sigma 2 square.

Consider two independent normal populations X and Y having unknown variances sigma 1 square and sigma two square respectively. We wish to construct a confidence interval for the ratio of variances sigma 1 square by sigma two square. This ratio can be initially estimated by the ratio of sample variances (sx) square by (sy) square obtained from independent random samples from X and Y having sizes m and n respectively.

If sx square by sy square is nearly equal to 1, we will find little evidence to indicate that sigma 1 square and sigma two square are unequal. On the other hand, a very large or very small value of sx square by sy square provides evidence of a difference in the population variances. How large or small must sx square by sy square be found by studying the distribution of sx square by sy square in repeated sampling.

Sampling distribution
The sampling distribution of (sx square by sigma x square) by (sy square by sigma y square)

is used. Since the population variances are usually not known, the sample variances are used. The assumptions are that sx square and sy square are computed from independent samples of size m and n, respectively, drawn from two normally distributed populations. If the assumptions are met, (sx square by sigma x square) by (sy square by sigma y square) follows a distribution known as the F distribution with two values used for degrees of freedom.

2. F - Distribution

F-Distribution

The confidence interval for the ratio of two variances requires the use of the probability distribution known as the *F*-distribution. So, let us spend a few minutes learning the definition and characteristics of the *F*-distribution.

Characteristics of the *F*-Distribution

F-distributions are generally skewed. The shape of an *F*-distribution depends on the values of m and n, the numerator and denominator degrees of freedom respectively.

The *F*-Table

One of the primary ways that we will need to interact with an *F*-distribution is by knowing either (1) an *F*-value, or (2) the probabilities associated with an *F*-random variable in order to complete a statistical analysis

Degrees of freedom

The F distribution uses two values for degrees of freedom. The numerator degree of freedom is the value of m minus 1, which is used in calculating sx square. The denominator degree of freedom is the value of n minus 1, which is used in calculating sy square.

Reading F tables

F tables come in denominations based on F (1 minus alpha by 2) which are F point nine nine five, F point nine nine, F point nine seven five etc or with F alpha by 2 must be calculated to give values of F point zero five, F point zero two five and F point zero zero five.

Here are the steps for using the *F*-table to find a *F*-value.

1. Find the column that corresponds to the relevant numerator degrees of freedom m.
2. Find the row that corresponds to the relevant denominator degrees of freedom n and alpha by 2
3. Determine the *F*-value, where the m column and the probability row identified in the row corresponding to n degree of freedom and alpha by 2 or or 1 minus alpha by 2 level of significance intersect

Confidence Interval for variance

Confidence intervals are not just for means or difference of means. You are already familiar with confidence intervals for the difference of means of two populations. The idea of a confidence interval is very general, and you can express the precision of any computed value as a ninety five percent or ninety-nine percent confidence interval (CI).

Very often practical problems lead to estimation of sigma square, the variance of the population and the ratio of variances of two independent populations. A confidence interval for the ratio of two population variances specifies a range of values within which the unknown population parameter, in this case the ratio of variances may lie.

The ninety five percent confidence interval for the ratio of variances

The sample variances are the values that we compute from a samples of data, which gives point estimates of the variances of two populations. It is not done often, but it is certainly

possible to compute a confidence interval for the ratio of two independent population variances.
If you assume that your data were randomly and independently sampled from a Gaussian (Normal) distribution, you can be ninety-five percent sure that the confidence interval computed from the sample variances contains the ratio of two true population variances.

An interval estimator is a rule for computing two numbers say A and B. This is to create an interval that contains the parameter of interest that is the ratio of two population variances sigma one square by sigma two square.
Hence, a confidence interval gives an estimated range of values for the ratio of two population variances or standard deviations, which is likely to include a ratio of two unknown population variances or standard deviations, the estimated range being calculated from a given set of sample data.

3. Result 1

Result 1:
Confidence Intervals for the ratio of variances of two independent populations when the population means are known
Assumptions:

- Population Means are known
- Populations are Normally distributed
- If population is not normal, use large sample

Suppose x one, x two, up to xm and yone, y two, up to yn are the random samples of size m and n drawn from the populations of size N1 and N2 with means mu 1 and mu 2 and variances sigma 1 square and sigma 2 square, we know that x bar is equal to summation xi by m and y bar is equal to summation yi by n
Xi follows Normal distribution with mean mu 1 and variance sigma 1 square and Yi follows Normal distribution with mean mu 2 and variance sigma 2 square

X bar follows Normal distribution with mean mu 1 and variance sigma 1 square by m and y bar follows Normal distribution with mean mu 2 and variance sigma 2 square by n
Summation (xi minus mu 1) whole square by sigma 1 square follows chi square with m degrees of freedom and
Summation (yi minus mu 2) whole square by sigma 2 square follows chi square with n degrees of freedom
Then, F of (m, n) is equal to chi square with m degrees of freedom by m whole divided by chi square with n degrees of freedom by n which is equal to Summation (xi minus mu 1) whole square by sigma 1 square into m divided by Summation (yi minus mu 2) whole square by sigma 2 square into n
Which is equal to n into Summation (xi minus mu 1) whole square by sigma 1 square into sigma 2 square by m into Summation (yi minus mu 2) whole square
Which is equal to sigma 2 square into n into summation (xi minus mu 1) whole square by sigma 1 square into m into Summation (yi minus mu2) whole square

From the table of probabilities of Snedekor's F distribution, we can always find two quantities A and B such that
Probability of A less than or equal to F of (m, n) less than or equal to B is equal to 1 minus alpha
Probability of A less than or equal to sigma 2 square into n into summation (xi minus mu 1) whole square by sigma 1 square into m into Summation (yi minus mu2) whole square less than or equal to B is equal to 1 minus alpha

Probability of A into m into Summation (yi minus mu2) whole square by n into summation (xi minus mu 1) whole square less than or equal to sigma 2 square by sigma 1 square is less than or equal to B into m into Summation (yi minus mu2) whole square by n into summation (xi minus mu 1) whole square is equal to 1 minus alpha
Probability of n into summation (xi minus mu 1) whole square by A into m into Summation (yi minus mu2) whole square is greater than sigma 1 square by sigma 2 square is greater than n into summation (xi minus mu 1) whole square by B into m into Summation (yi minus

mu2) whole square is equal to 1 minus alpha
Probability of n into summation (xi minus mu 1) whole square by B into m into Summation (yi minus mu2) whole square is less than or equal to sigma 1 square by sigma 2 square is less than or equal to n into summation (xi minus mu 1) whole square by A into m into Summation (yi minus mu2) whole square is equal to 1 minus alpha

Therefore, hundred into (1 minus alpha) percent Confidence Interval for the ratio of variances of two populations with known means as mu 1 and mu 2   is given by
[n into summation (xi minus mu 1) whole square by B into m into Summation (yi minus mu2) whole square, n into summation (xi minus mu 1) whole square by A into m into Summation (yi minus mu2) whole square]

4. Result 2

Result 2:
Confidence Interval for the ratio of two variances of two populations with unknown means
Assumptions:

- Populations means are unknown
- Population is normally distributed
- If population is not normal, use large sample

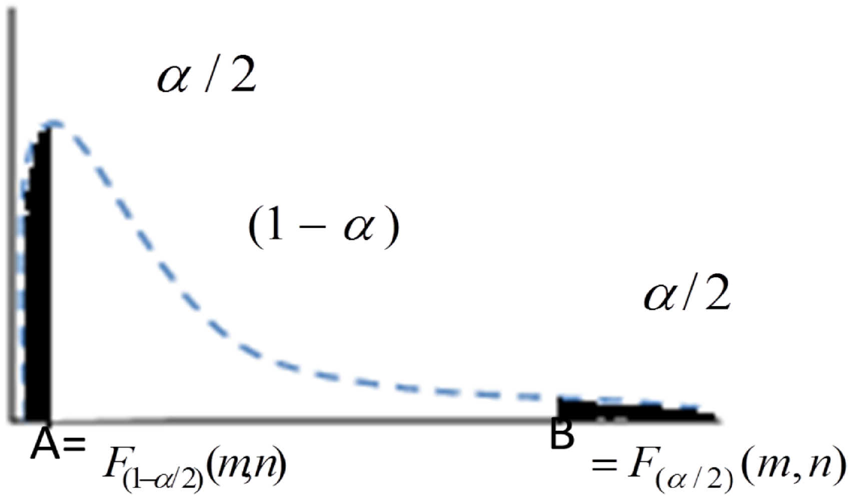
Suppose x one, x two, up to xm and y one, y two, up to yn are the random samples of size m and n drawn from the populations of size N1 and N2 with means mu 1 and mu 2 and variances sigma 1 square and sigma 2 square,
We know that x bar is equal to summation xi by m and y bar is equal to summation yi by n
Xi follows Normal distribution with mean mu 1 and variance sigma 1 square and Yi follows Normal distribution with mean mu 2 and variance sigma 2 square

X bar follows Normal distribution with mean mu 1 and variance sigma 1 square by m and y bar follows Normal distribution with mean mu 2 and variance sigma 2 square by n
(m minus 1) into s1 square by sigma 1 square follows Chi square with (m minus 1) degrees of freedom and (n minus 1) into s2 square by sigma 2 square follows Chi square with (n minus 1) degrees of freedom
F of (m minus 1, n minus 1) is equal to Chi square (m minus 1) by (m minus 1) whole divided by Chi square (n minus 1) by (n minus 1)
Which is equal to (m minus 1) into s1 square by sigma 1 square into (m minus 1) whole divided by (n minus 1) into s2 square by sigma 2 square into (n minus 1)
Which is equal to s1 square into sigma 2 square by s2 square into sigma 1 square which is equal to sigma 2 square into s1 square by sigma 1 square into s2 square

From the table of probabilities of Snedekor's F distribution, we can always find two quantities A and B such that
Probability of A less than or equal to F of (m minus 1, n minus 1) less than or equal to B is equal to 1 minus alpha
Probability of A less than or equal to sigma 2 square into s1 square by sigma 1 square into s2 square less than or equal to B is equal to 1 minus alpha
Probability of A into s2 square by s1 square less than or equal to sigma 2 square by sigma 1 square less than or equal to B into s2 square by s1 square is equal to 1 minus alpha
Probability of s1 square by A into s2 square greater than sigma 1 square by sigma 2 square greater than s1 square by B into s2 square is equal to 1 minus alpha

Probability of s1 square by B into s2 square less than or equal to sigma 1 square by sigma 2 square less than or equal to s1 square by A into s2 square is equal to 1 minus alpha
Therefore, hundred into (1 minus alpha) percent confidence Interval for the ratio of variances of two populations with unknown means is given by
[s1 square by B into s2 square, s1 square by A into s2 square]

Figure 1



As shown in the graph, here there are two critical values for each level of confidence. The value of B is equal to F (alpha by two) with (m, n) degrees of freedom represents the right-tail critical value and A is equal to F (1 minus alpha by two) with (m, n) degrees of freedom represents the left-tail critical value. Each area in the F- table represents the region under F curve to the RIGHT of the critical value.

We can use the critical values B is equal to F (alpha by two) with (m, n) degrees of freedom and A is equal to F (1 minus alpha by two) with (m, n) degrees of freedom to construct confidence intervals for the ratio of population variances and standard deviations when means are known.

5. Steps to Obtain Confidence Interval for the Ratio of two Variances

Steps to obtain confidence interval for the ratio of two variances

- First, obtain the point estimates of mu 1 and mu 2 that is, the sample mean x bar and y bar if population means are not known otherwise consider the value of mu 1 and mu 2

- Select a confidence level. The most common level is ninety-five percent or ninety-nine percent

- Obtain the critical values of F probabilities for alpha by two and 1 minus alpha by two level of significance with (m, n) degrees of freedom if mu 1 and mu 2 are known ((m minus 1, n minus 1) degrees of freedom if mu 1 and mu 2 are unknown)

- Let B is equal to F (alpha by 2) with (m, n) degrees of freedom and A is equal to F (1 minus alpha by 2) with (m, n) degrees of freedom

- Then, the confidence interval for the ratio of variances when means are known as mu 1 and mu 2 is computed by making use of the given derived formula in result one

- When the population means are unknown, Let B is equal to F (alpha by 2) with (m minus 1, n minus 1) degrees of freedom and A is equal to F (1 minus alpha by 2) with (m minus 1, n minus 1) degrees of freedom

- Then, the confidence interval for the ratio of variances when means are unknown is computed by making use of the above-derived formula in result two

Here's a summary of our learning in this session, where we understood:

- The derivation for confidence limits for the ratio of the two unknown population variances when means are known
- The derivation for confidence limits for the ratio of two unknown population variances when the means are unknown