UNIT 5

SAMPLING DISTRIBUTIONS ASSOCIATED WITH NORMAL POPULATIONS–II

Structure

5.1 INTRODUCTION

In Unit 4, we have discussed the chi-square and t-distributions in detail with their properties and applications. The F-distribution also has numerous real-world applications especially when discussing variance analysis and hypothesis testing of the variances of the two normally distributed populations. For example, in finance, it is used to check whether the variances of stock returns are equal across two or more stocks. In engineering, it is also used to test the effectiveness of different manufacturing processes by comparing the variances of the outcomes. Additionally, the F-distribution is used in biostatistics to compare the variances of health outcomes across different treatments or interventions. In this unit, we discuss the F-distribution in detail and explain the method of reading the tabulated value of t, chi-square and F-distribution tables.

This unit is divided into 11 sections. Section 5.1 is introductive. In Sections 5.2 to 5.5, we discuss the F-distribution with its probability density curve, summary measures, relation to other distributions, properties and applications. As the standard normal distribution has a standard normal table (Z-table), in a similar

Tools You Will Need

The following terms are considered essential background material for this Unit. If you doubt your knowledge of any of these terms, you should review the appropriate Unit or section before proceeding:

- Sampling Distributions for Means and Variance (Units 2 and 3).

Unit Write- Dr. Prabhat Kumar Sangal, School of Sciences, IGNOU, New Delhi

ignou THE PEOPLE'S UNIVERSITY

way, the chi-square, t and F-distributions also have their tables. Therefore, Sections 5.6 to 5.8 are devoted to how to read tabulated values of the t, chi-square, and F-distributions, respectively. The unit ends by providing a summary of what we have discussed in this unit in Section 5.9. The terminal questions and the solution of the SAQs/TQs are given in Sections 5.10 and 5.11, respectively.

In the next unit, we shall discuss the estimation of the unknown parameters.

Expected Learning Outcomes

After studying this unit, you should be able to:

- ❖ explain the F-distribution with its probability curve, summary measures, relations to other distributions, properties and applications; and
- ❖ describe the method of obtaining the tabulated value from the t, chi-square and F-distribution tables.

5.2 F-DISTRIBUTION

(1890 – 1962)
Sir Ronald Aylmer Fisher was a British and worked as a statistician, mathematician, biologist, geneticist, and academic. For his work in statistics, he has been known as father of modern statistical sciences.

(1881– 1974)
George Waddel Snedecor was an American mathematician and statistician who contributed to the foundations of analysis of variance, data analysis, experimental design, and statistical methodology.

The F-distribution is also a continuous probability distribution as the chi-square or t-distribution. It is a sampling distribution that commonly occurs in statistics when we discuss variances. As the chi-square distribution is not very useful in describing real-world data, similarly, the F- distribution is not much used to describe real-world data. However, it is used for estimating and hypothesis testing related to variances of two normal populations.

The F-distribution was first introduced by British statistician Ronald A. Fisher in 1928. So sometimes it is called **Fisher's F-distribution**. Later on, George Waddel Snedecor was an American mathematician and statistician who tabulated the F-distribution and used the letter F in Fisher's honour. The distribution is also known as **Snedecor's F distribution** or the **Fisher-Snedecor distribution**. Prof. Ronald. A. Fisher defined the F-distribution when he was interested in comparing the variances of two normally distributed populations, and he derived the F-distribution as the ratio of two independent chi-square variates when divided by their respective degrees of freedom, that is,

$$F = \frac{\chi^2_{(n_1-1)} / (n_1 - 1)}{\chi^2_{(n_2-1)} / (n_2 - 1)}$$

As we know the chi-square statistic $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$, therefore,

$$F = \frac{(n_1 - 1)S_1^2 / \sigma_1^2 (n_1 - 1)}{(n_2 - 1)S_2^2 / \sigma_2^2 (n_2 - 1)}$$

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \sim F_{(n_1-1,\, n_2-1)}$$

If the variances of both populations are equal i.e. $\sigma_1^2 = \sigma_2^2$ then we can write the F-statistic as follows:

$$F = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1, n_2-1)}$$

Thus, the sampling distribution of the ratio of sample variances follows the F-distribution with $(n_1 - 1, n_2 - 1)$ degrees of freedom. The F-distribution is also a family of distribution, and it has different shapes for each combination of these degrees of freedom. Let us do an example to learn how to calculate the F-statistic.

Example 1: Suppose a student of the MSCAST programme of Jammu Kasmir region wants to test whether the variance of the weight of apples produced by two different orchards in Kasmir is the same. He collects a random sample of 25 apples from Orchard I and 15 apples from Orchard II and obtains the sample variances of 62 grams and 45 grams, respectively. Compute the value of the F-statistic and also find the degrees of freedom associated with the F-statistic.

Solution: Here, we are given that

$$n_1 = 25, n_2 = 15,$$

$$S_1^2 = 62, S_2^2 = 45$$

We can calculate the value of the F-statistic as

$$F = \frac{S_1^2}{S_2^2} = \frac{62}{45} = 1.38$$

Since the F-statistic is associated with S_1^2 and S_2^2, therefore, the degrees of freedom of it depends on the degrees of freedoms of S_1^2 and S_2^2. Since S_1^2 has $n_1 - 1 = 25 - 1 = 24$ degrees of freedom and, S_2^2 has $n_2 - 1 = 15 - 1 = 14$ degrees of freedom, therefore, the F-statistic has (24, 14) degrees of freedom.

After understanding the F-statistic, let us discuss the probability density function (pdf) of the F-distribution. As you seen, it has two degrees of freedom, one for the numerator (n_1) and the other for the denominator (n_2), therefore, the pdf of the F-distribution is more complex than the chi-square and the t-distributions. We can define the F-distribution as

A continuous random variable X follows an F-distribution with (n_1, n_2) degrees of freedom if and only if it has the following probability density function:

$$f(x) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{x^{\frac{n_1}{2}-1}}{\left(1 + \frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}; \quad 0 < x < \infty$$

Symbolically, we denote that X follows an F-distribution with (n_1, n_2) degrees of freedom as $X \sim F_{(n_1, n_2)}$.

Since the F-distribution is a sampling distribution, therefore, to distinguish it from the well-known distributions such as normal, exponential, binomial, etc. We may use **F** symbol instead of **X** and define it as follows:

If a random sample $X_1, X_2, ..., X_{n_1}$ of size n_1 is taken from a normal population with mean μ_1 and variance σ_1^2 and another independent random sample

$Y_1, Y_2, ..., Y_{n_2}$ of size n_2 is taken from another normal population with mean μ_2 and variance σ_2^2, respectively, then the probability density function of the F- distribution with (n_1, n_2) degrees of freedom is given by

$$f(F) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{F^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}F\right)^{\frac{n_1+n_2}{2}}}; \quad 0 < F < \infty$$

The distribution is bounded on the left by zero and has no upper limit, that is, extending indefinitely to the right, which reflects the fact that variances are always non-negative. The F-distribution has two parameters that specify the number of degrees of freedom. It means that an F-distribution is determined by its degrees of freedom and there is a different F-distribution for each pair of degrees of freedom. Therefore, it is also a family of continuous probability distributions.

Let us take an example to understand how to find degrees of freedom when the pdf of an F-distribution is given.

Example 2: If a random variable X follows an F-distribution whose pdf is given by

$$f(x) = \frac{1}{(1+x)^2}; \quad 0 < x < \infty$$

then obtain the degrees of freedom of this distribution.

Solution: If random variable X follows the F-distribution with (n_1, n_2) degrees of freedom, then the probability density function of X is given as

$$f(x) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{x^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}; \quad 0 < x < \infty$$

We now try to convert the given pdf in the form of the standard form of the F-distribution so that we can compare and find the degrees of freedom. Therefore,

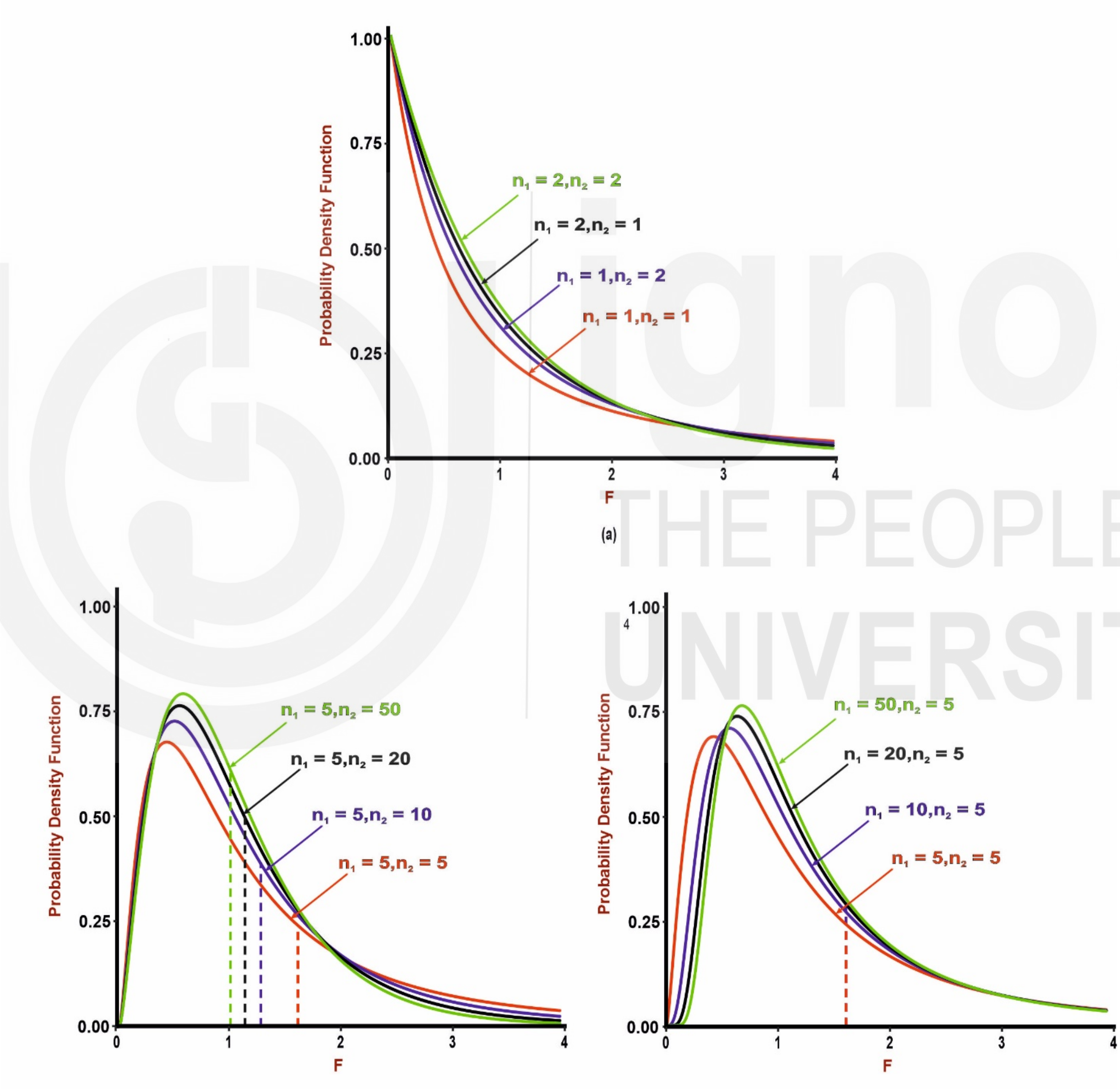
$$f(x) = \frac{\left(\frac{2}{2}\right)^{\frac{2}{2}}}{B\left(\frac{2}{2}, \frac{2}{2}\right)} \frac{x^{\frac{2}{2}-1}}{\left(1+\frac{2}{2}x\right)^{\frac{2+2}{2}}}; \quad 0 < x < \infty \; \left[\because B(1,1) = 1\right]$$

By comparing the above form with the standard form, we get degrees of freedom $n_1 = 2$ and $n_2 = 2$.

After understanding the form of the pdf of the F-distribution, you may be interested to know the shape of the F-distribution and the impact of both degrees of freedom on the shape of it. Let us discuss the probability density curve of the F-distribution in the next section.

5.2.1 Probability Density Curve of F-distribution

The F-distribution is a positively skewed distribution that has a minimum value of 0, but no maximum value. The shape of the F-distribution depends on two parameters: degrees of freedom for the numerator (n_1) and degrees of freedom for the denominator (n_2). The shape of the F-distribution for n_1 = 1, 2 and n_2 = 1, 2 is inverse J alphabet and the curve starts out high and then subsequently declines as shown in Fig. 5.1(a). However, for n_1 and n_2 > 2, the shape of the F-distribution is as shown in Fig 5.1(b and c). The probability density curve of the F-distribution reaches a peak (not far to the right of 0), and then gradually approaches the horizontal axis with the larger value of F. The F-distribution approaches the horizontal axis but never touches it.



(b): n_1 is fixed and n_2 is very **(c): n_2 is fixed and n_1 is very**

Fig. 5.1: Probability density curves of F-distribution for various degrees of freedom.

In Fig. 5.1 (b), we plot different probability curves and put n_1 as fixed at n_1 = 5 and increase n_2 = 5 to 10 to 20 to 50. By increasing the second parameter n_2 from 5 to 50, the mean of the distribution (shown by the vertical line)

ignou THE PEOPLE'S UNIVERSITY

decreased, and the probability curve shifted from the tail to the centre of the distribution. Similarly, in Fig. 5.1 (c), we plot different probability curves and put n_2 as fixed at $n_2 = 5$ and increase $n_1 = 5$ to 10 to 20 to 50. By increasing the first parameter n_1 from 5 to 50, the mean of the distribution (shown by the vertical line) does not change but the probability curve shifted from the tail to the centre of the distribution. After looking at the probability curves of the F-distribution, we can observe that the probability curve of the F-distribution is a uni-model curve for $n_1, n_2 > 2$.

After understanding the probability curve of the F-distribution with some of its properties. In this sub-section, we will discuss some summary measures of it.

5.2.2 Summary Measures of F-distribution

The mean of the F-distribution depends only on the degrees of freedom of the denominator, that is, n_2 and is given as follows:

$$\text{Mean} = \frac{n_2}{n_2 - 2} \text{ for } n_2 > 2.$$

Since the F-distribution is asymmetrical and right-skewed, therefore, the mean is greater than the median and mode.

The variance of the F-distribution depends on both (n_1, n_2) dfs and is given by

$$\text{Variance} = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)} \text{ for } n_2 > 4.$$

This programme is applied in nature, therefore, we do not give proof of the mean and variance of the F-distribution.

If someone is interested in that he/she can derive these summary measures as discussed in MST-012.

Let us take some simple examples based on the probability density function and summary measures.

Example 3: For the probability density function of the distribution given in Example 2, find the mean and variance of this distribution.

Solution: In Example 2, we obtained the degrees of freedom of the given F-distribution as $n_1 = 2$, and $n_2 = 2$. We know that the mean and variance of the F-distribution with (n_1, n_2) degrees of freedom are:

$$\text{Mean} = \frac{n_2}{n_2 - 2} \text{ for } n_2 > 2 \text{ and}$$

$$\text{Variance} = \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)} \text{ for } n_2 > 4.$$

Since both mean and variance of the F-distribution exist for $n_2 > 2$ and $n_2 > 4$, respectively, therefore, for $n_1 = 2$, and $n_2 = 2$, the mean and variance of the F-distribution do not exist.

Now, try to answer the following Self Assessment Question to see how much you learn about the F-distribution.

SAQ 1

If a random variable X follows the F-distribution whose pdf is given by

$$f(x) = \frac{3}{8\sqrt{x}} \frac{1}{\left(1+\frac{1}{4}x\right)^{\frac{5}{2}}}; \quad 0 < x < \infty$$

then obtain the degrees of freedom of this distribution. Also, find its mean and variance.

After understanding the F-distribution with its pdf, probability density curve and summary measures, let us see the relationship of the F-distribution with other well-known distributions.

5.3 RELATION OF F-DISTRIBUTION TO OTHER DISTRIBUTIONS

The F-distribution is closely related to t and chi-square distributions. Some of the relationships are discussed as follows:

- If a variable t follows the t-distribution with n df, then the square of t follows the F-distribution with (1, n) df i.e. if $t \sim t_{(n)}$ then $t^2 \sim F(1, n)$.
- As the denominator degrees of freedom (n_2) of the F-distribution increases then n_1F follows a chi-square distribution with n_1 degrees of freedom.
- If a random variable X follows the F-distribution with (n_1, n_2) degrees of freedom then 1/X also follows the F-distribution with (n_2, n_1) degrees of freedom.

Now, try to answer the following Self Assessment Question.

SAQ 2

If a statistic t follows Student's t-distribution with 4 df, then what will be the distribution of the square of t? Also, write the pdf of that distribution.

After introducing the F-distribution, one may be interested in knowing the properties of this distribution. We now discuss some of the important properties of the F-distribution in the next section.

5.4 PROPERTIES OF F-DISTRIBUTION

The important properties of the F-distribution are as follows:

1. The value of the F-distribution is always positive, or zero because it is the distribution of the ratio of the variances which are the square of the deviations and hence cannot assume negative values. Its value lies between 0 and ∞.
2. The F-distribution has two parameters n_1 and n_2, that is, the degrees of freedom for numerator and denominator and the shape of the distribution depends on them. Therefore, there is a different F-distribution for each pair of degrees of freedom.

3. The F-distribution is positively skewed. The curve is more positively skewed when n_2 is smaller than n_1.

4. The F-distribution is a uni-modal distribution, that is, it has a single mode.

5. The square of the t-statistic with n df follows the F-distribution with 1 and n degrees of freedom.

6. The mean of the F-distribution with (n_1, n_2) df is $\frac{n_2}{n_2 - 2}$ for $n_2 > 2$.

7. The variance of the F-distribution with (n_1, n_2) df is

$$\frac{2n_2^2\left(n_1 + n_2 - 2\right)}{n_1\left(n_2 - 2\right)^2\left(n_2 - 4\right)} \text{ for } n_2 > 4.$$

8. If we interchange the degrees of freedom n_1 and n_2 then there exists an extremely useful relation as

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

This is called as **Reciprocal Property** of the F-distribution.

9. As the degrees of freedom for the numerator and for the denominator get larger, then the curve approximates the normal curve.

You can try the Self Assessment Question to see how much you learn about the properties of the F-distribution.

SAQ 3

Write any five properties of the F-distribution.

5.5 APPLICATIONS OF F-DISTRIBUTION

After discussing the main properties of the F-distribution in the previous section, we are now discussing some of the important applications of the F-distribution. The F-distribution has the following applications:

1. The F-distribution is used to test the hypothesis about the variances of two normal populations. Also, it is used for the construction of the confidence intervals for the ratio of two population variances.

2. The F-distribution is also used in regression analysis, particularly in the test for the overall significance of a regression model or to compare the variances of the residuals for two or more models.

3. The t-distribution is used to test/assess the statistical significance of the difference between two population means but if we have to test the significance of more than two means then we use the F-distribution. That is the F-distribution is used in ANOVA as well as the design of experiments.

You will study the first application listed above in Units 14 and 18 of this course and the second application in the course MST-017: Applied Regression Analysis. The third application you have already studied in MST-003: Survey Sampling and Design of Experiments-I.

Now, try the following Self Assessment Question.

SAQ 4

Write four applications of the F-distribution.

After understanding the standard sampling distributions such as t, chi-square and F-distributions in detail, next, you are going to learn how to read their tables. As we have seen in Unit 14 of the course MST-012: Probability and Probability Distributions, we can calculate the area or probability of a random variable which follows a normal distribution with the help of a standard normal table. Similarly, we can also use the t, chi-square, and F-distribution tables. The main difference between the standard normal distribution table and the t, chi-square, and F-distribution tables is that the body of the standard normal distribution table represents the probability whereas the body of the t, chi-square, and F-distribution tables represents the critical value or point beyond which the area/probability of the distribution is α. Therefore, these tables are generally used to find out the value of the variable for which area in the tail of the distribution is given. The tabulated values are also known as **critical values** of the standard sampling distributions. They are also used in constructing **confidence intervals** and **testing of hypotheses** which you will study in the next units of this course. We discuss how to read tabulated values of these distributions one at a time.

5.6 TABULATED VALUES OF t-DISTRIBUTION

As you have seen the t-distribution is described with the help of degrees of freedom (n) and for each degree of freedom there is a different t-distribution. To include each t-distribution, the statisticians arrange the t-table as given in **Table V** in the Appendix. The t-table contains the tabulated values (it is also known as critical values, especially in testing of hypothesis) of t- variable for different degrees of freedom (n) such that the area under the probability curve of the t-distribution to its right tail (upper tail) is equal to α (α is also known as level of significance in testing of hypothesis) as shown in Fig. 5.2 (a). In the t-table, the first column of the left-hand side represents the degrees of freedom (n) while the column heading represents the upper (right-hand side) tail area/probability (α) of the probability curve of the t-distribution. The body contains the value of the t-statistic for each particular value of n and α which represents the critical value or point beyond which the area/probability of the t-distribution is α. **The area/probability (α) represents the proportions of the t-distribution contained in the right tail.**

The t-distribution table is also known as Student's t-table, t-table, t-score table, t-value table, or t-test table.

<u>How to Use the t-distribution Table</u>

To read the t-distribution table, you only need to know three values:

- degrees of freedom (n)
- area/probability, that is, α (level of significance) of the t-statistic (common choices are 0.01, 0.05, and 0.10)
- the tail of the t-statistic on which α lies (one tail or two tails).

If the area/probability, that is, alpha (level of significance) of the t-statistic may lie on the right tail, left tail or both tails so the tabulated value of the t-distribution is called right tail value, left tail value or both tail values. We now discuss how to read the tabulated values in each case one at a time.

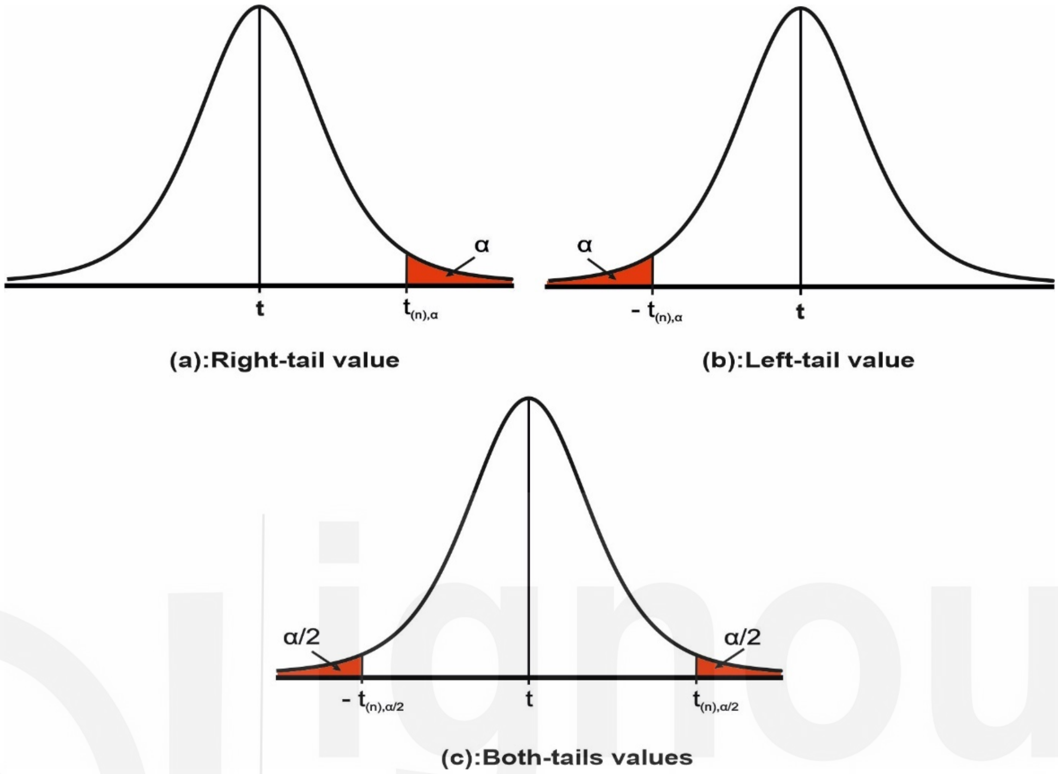


Fig. 5.2: Representation of tabulated value(s) of t-distribution.

Right tail value

The t-table contains the tabulated values (it is also known as critical values, especially in testing of hypothesis) of the t-statistic for different degrees of freedom (n) such that the area under the probability curve of the t-distribution to its right tail (upper tail) is equal to α. To read the right tail tabulated value, we follow the following steps:

Step 1: We start with the first column of the t-table, that is degrees of freedom and downward headed 'n' until the required degree of freedom is reached.

Step 2: After that, we proceed right to the column headed α up to the required α is reached.

Step 3: We get the required right tail tabulated value in the cell of the table at the intersection of required degrees of freedom n and α.

We represent the right tail tabulated value of the t-statistic for n degrees of freedom and for right tail area/probability (α) as $t_{(n),\alpha}$. The right tail probability value represents the probability that the t-statistic would be greater than $t_{(n),\alpha}$ is α.

Suppose we want to read the tabulated value of the t-statistic for which the area/probability on the right tail is 0.05 and the degrees of freedom is 6. To see the tabulated value, we start with the first column of the t-table, that is, degrees of freedom and downward headed 'n' until entry 6 is reached and then proceed right to the column headed α = 0.05. Find the cell in the table at the intersection of degrees of freedom n = 6 and α =0.05 level. This is the t-distribution value. For your convenience, we give a part of the t-table in Table 5.1.

Thus, we get the required right value of the t-statistic as $t_{(n), \alpha} = t_{(6), 0.05} = 1.943$. The value of the t-statistic equal to 1.943 means, the probability that the t-statistic would exceed (greater than) 1.943 is 0.05 as shown in Fig. 5.3.

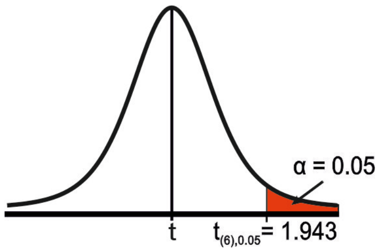


Fig. 5.3

Table 5.1: Part of t-table

α =	0.10	0.05	0.025	0.01	0.005
n (df) =1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499

Left tail value

Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail tabulated values will be equal to the right tail tabulated values in magnitude but opposite in sign (as in the case of normal distribution). Therefore, in the t-table, only right tail values are given. To read the left tail tabulated value, we follow the following steps:

Step 1: First of all, we read the right tail value for the same degrees of freedom (n) and the same area/probability α (level of significance) as discussed.

Step 2: After that, we assign the negative sign to the right tail tabulated value read in Step 1.

We represent the left tail tabulated value of the t-statistic for n degrees of freedom and α level as $t_{(n),(1-\alpha)}$. But the t-distribution is symmetrical about the t = 0 line, therefore, we can represent it as $t_{(n),(1-\alpha)} = -t_{(n),\alpha}$. The left tail tabulated value represents the probability that the t-statistic would be greater than $t_{(n),(1-\alpha)} = -t_{(n),\alpha}$ is 1 – α and less than it is α as shown in Fig. 5.2(b).

Suppose in the above example, if we want to find the value of the t-statistic such that the left area is 0.05 for 6 df then due to the symmetry of the t-distribution the value of the t-statistic will be $-t_{(n), \alpha} = -t_{(6), 0.05} = -1.943$. The tabulated value is shown in Fig. 5.4.

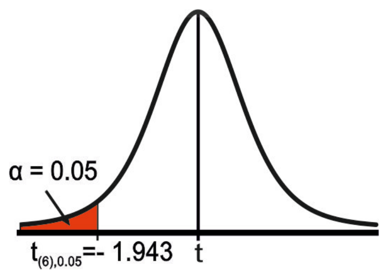


Fig. 5.4

Both tails values

After learning how to read tabulated values for right and left tails, we now discuss how to read it for two tails. For two-tails, there are two tabulated values. To read both tails tabulated value, we follow the following steps:

Step 1: First of all, we half the total area on both tails. If it is α then the half area i.e. α/2 lies in both tails as shown in Fig. 5.2(c).

Step 2: We then read the right-tail and left-tail tabulated values for the same degrees of freedom (n) and α/2 (area/probability lie on right and left tails) instead of α as discussed in the case of right and left tails tabulated values.

We represent both tails tabulated values of the t-statistic for n degrees of freedom and α level as $\pm t_{(n),\, \alpha/2}$.

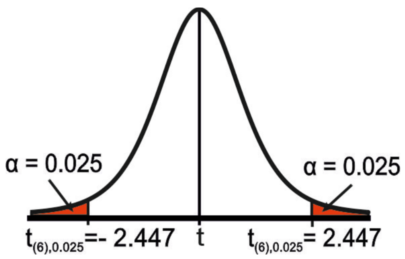


Fig. 5.5

For example, if we want to find out the values of the t-statistic for which the area on both tails is 0.05 and the degrees of freedom is 6. Since the total area on both tails is 0.05, therefore, the area on the right tail as well as on the left tail will be 0.05/2 = 0.025. Thus, we start with the first column of the t-table and downward headed n until entry 6 is reached. Then proceed right to the column headed α = 0.025. We get $t_{(n),\, \alpha/2} = t_{(6),\, 0.025} = 2.447$. Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail value will be the same as the right tail value but in a negative sign. Therefore, $-t_{(n),\, \alpha/2} = -t_{(6),\, 0.025} = -2.447$. So the required values of the t-statistic are $\pm t_{(n),\, \alpha/2} = \pm t_{(6),\, 0.025} = \pm 2.447$. We also show the tabulated values in Fig. 5.5.

Let us take an example.

Example 4: Find the tabulated value of the t-statistic in each case for which the degrees of freedom (n) and the area (level of significance) are given as follows:

(i) n = 10 and α = 0.01 (right tail)

(ii) n = 8 and α = 0.05 (left tail)

(iii) n = 14 and α = 0.10 (both tails)

Solution:

(i) Here, we want to read the tabulated value of the t-statistic for

n = 10 and α = 0.01 (right tail)

Therefore, we start from the first column of the t-table (Table V) given in the Appendix at the end of this volume and downward headed n until entry 10 is reached. Then proceed right to the column headed α = 0.01. So we get the required tabulated value of the t-statistic as $t_{(n),\, \alpha} = t_{(10),\, 0.01} = 2.764$.

(ii) Here, we are given that

n = 8 and α = 0.05 (left tail)

First of all, we read the right tail value by proceeding the same way as part (i) $t_{(n),\, \alpha} = t_{(8),\, 0.05} = 1.860$. Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail value is $-t_{(n),\, \alpha} = -t_{(8),\, 0.05} = -1.860$.

(iii) Here, we want to read the value of the t-statistic for two tails and it is given that

n = 14 and α = 0.10

Since the total area on both tails is 0.10, therefore, the area on the right tail as well as on the left tail will be 0.05/2 = 0.05. Thus, we start with the first column of the t-table and downward headed n until entry 14 is

ignou THE PEOPLE'S UNIVERSITY

reached. Then proceed right to the column headed 0.025. We get $t_{(n),\, \alpha/2} = t_{(14),\, 0.05} = 1.761$. Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail value will be the same as the right tail value but in a negative sign. Therefore, $-t_{(n),\, \alpha/2} = -t_{(14),\, 0.05} = -1.761$. So the required values of the t-statistic are $\pm t_{(n),\, \alpha/2} = \pm t_{(14),\, 0.025} = \pm 1.761$.

After understanding how to read the tabulated (critical) values for different cases, we now discuss some more facts about the t-table. If we closely inspect the t-table then we observe:

- The t-table contains only the right-tailed tabulated values.
- As the degrees of freedom increase, the tabulated value decreases. The reason is that the tails of the t-distribution shift towards the centre as the degrees of freedom increase as shown in Fig. 5.6.
- The t-distribution table does not include entries for every possible degree of freedom. For example, the table lists continuously up to df = 30 and after that t values for df = 40, 60, 120 and does not list entries for degrees of freedom values between these.

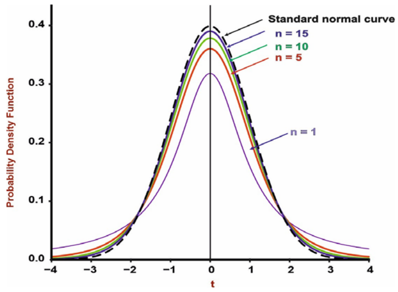


Fig. 5.6

Since the t-table does not include entries for every possible degree of freedom, therefore, a question may arise **"how do we read the tabulated values for df which are not included in the t-table?"** Don't bother about that we can read such values either using software or by the interpolating method which is discussed as follows:

Method of Finding the Values of t-statistic for Degrees of Freedom which are not Listed in the Table

The t-table (Table V) given in the Appendix does not list values for every possible degree of freedom. Therefore, it becomes necessary to know how to find values of the t-statistic for degrees of freedom not listed in the table. Let us discuss the process of finding the values which are not listed in the table with the help of an example.

Suppose we want to find out the tabulated value of the t-statistic for 34 degrees of freedom which is not listed in the table such that the area on the right side is equal to 0.05.

For that, first of all, we read the tabulated values of the t-statistic that are just greater and just smaller than the degrees of freedom for our interest. Thus, from the t-table, we get the values of t-statistic for 40 and 30 degrees of freedom and area α = 0.05 as

$$t_{(40),\, 0.05} = 1.684 \text{ and } t_{(30),\, 0.05} = 1.697$$

Note that the larger the degrees of freedom, the smaller the tabulated value of the t-statistic.

We now calculate how much the t-value changes for each degree of freedom between these two tabulated values. Here, there is a difference of 10(40 – 30) degrees of freedom and a t-value change of 0.013(1.697 – 1.684).

Thus, we can obtain the change in the t-value corresponding to a unit change in degree of freedom as

ignou THE PEOPLE'S UNIVERSITY

$$\frac{0.013}{10} = 0.0013$$

Since we have to obtain the value for 34 degrees of freedom, this is either 4 more than 30 or 6 less than 40. Therefore, we can interpolate from either value. To get from 30 to 34 degrees of freedom there is a difference of 4 (34 – 30). So we multiply this difference by the amount by which the t-value changes per degree of freedom i.e. 0.0013. This result as

$$4 \times 0.0013 = 0.0052$$

Since the larger the degrees of freedom the smaller the tabulated value of the t-statistic, therefore, subtracting this value 0.0052 from $t_{(30),\, 0.05}$ = 1.697 to get the required value. Thus,

$$t_{(34),\, 0.05} = 1.697 - 0.0052 = 1.6918$$

Now, if we interpolate it from 40 degrees of freedom then the difference 6(40 – 34) multiplied by 0.0013 and adding this to 1.684, we get

$$t_{(34),\, 0.05} = 1.684 + 6 \times 0.0013 = 1.6918$$

Thus, we get the same value.

Now, you can try the following Self Assessment Question.

SAQ 5

Find the tabulated values of the t-statistic for which the area and the degrees of freedom are given as follows:

(i) n = 9 and α = 0.01 (right tail)

(ii) n = 15 and α = 0.05 (left tail)

(iii) n = 13 and α = 0.05 (both tails)

5.7 TABULATED VALUES OF CHI-SQUARE DISTRIBUTION

The chi-square distribution table is also known as chi-square table, chi-square score table, chi-square value table, or chi-square test table.

Similar to the t-table, the chi-square table is given in the Appendix as **Table VI**. The chi-square table contains the tabulated values (it is also known as critical values, especially in the testing of hypothesis) of the chi-square statistic for different degrees of freedom (n) such that the area under the probability curve of the chi-square distribution to its right tail (upper tail) is equal to α (α is also known as the level of significance in testing of hypothesis) as shown in Fig. 5.7.

In the chi-square table, the column headings indicate the area on the upper portion (right tail) of the probability curve of the chi-square distribution and the first column on the left-hand side indicates the values of degrees of freedom (n). The body contains the value of the chi-square statistic for each particular value of n and α which represents the critical value or point beyond which the area/probability of the chi-square distribution is α. **The area/probability (α) represents the proportions of the chi-square distribution contained in the right tail.**

How to Use the Chi-square Table

Since the area/probability, that is, α (level of significance) of the chi-square statistic may lie on the right tail, left tail or both tails so the tabulated value of the chi-square distribution is called the right tail value, left tail value or both tail values. We now discuss how to read these tabulated values one at a time.

Right tail value

The chi-square table contains the tabulated values (it is also known as critical values, especially in testing of hypothesis) of the chi-square variable for different degrees of freedom (n) such that the area under the probability curve of the chi-square distribution to its right tail (upper tail) is equal to α.

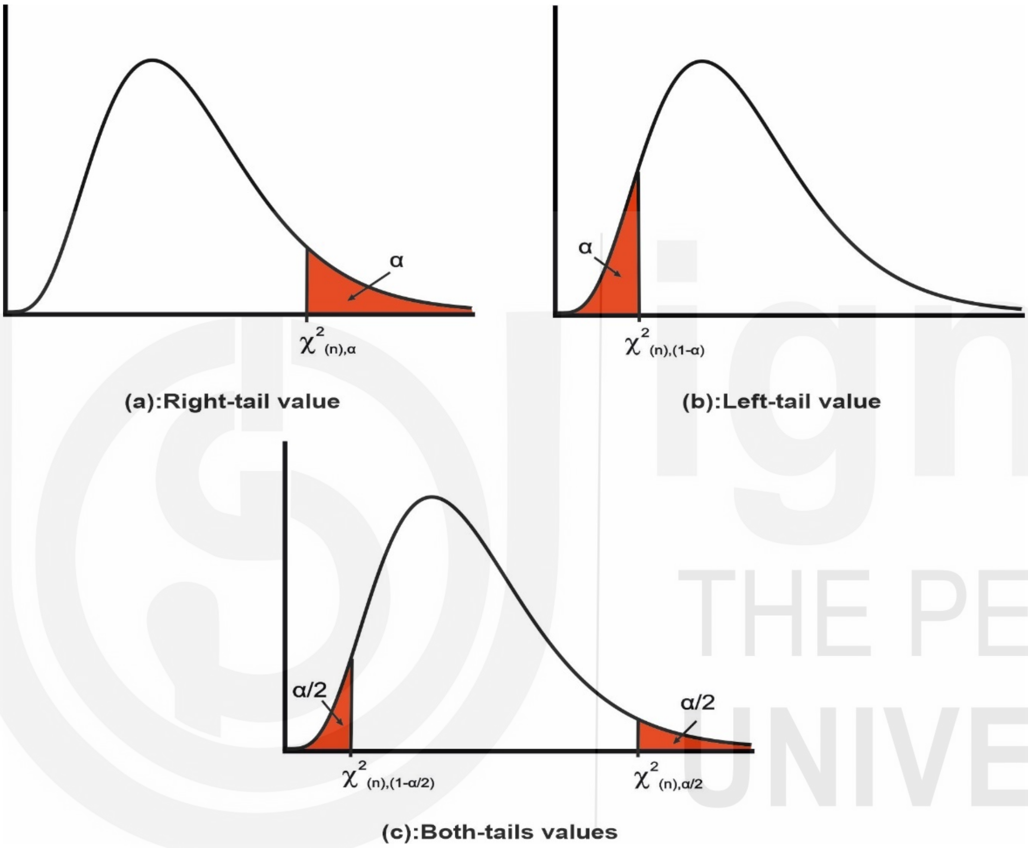


Fig. 5.7: Representation of Tabulated value(s) of chi-square distribution.

The procedure to read the right tail tabulated value of the chi-square statistic is almost similar to reading the t-table, we follow the following steps:

Step 1: We start with the first column of the chi-square table, that is degrees of freedom and downward headed 'n' until the required degree of freedom is reached.

Step 2: After that, we proceed right to the column headed α up to the required α is reached.

Step 3: We get the required right tail tabulated value in the cell of the table at the intersection of required degrees of freedom n and α.

We represent the tabulated value of the chi-square statistic for n degrees of freedom and for right tail area/probability as $\chi^2_{(n),\alpha}$.

Suppose we want to find out the tabulated value of the chi-square statistic for which the area/probability on the right tail is 0.01 and the degrees of freedom is 4. To read the tabulated value, we start with the first column of the chi-square table, that is degrees of freedom and downward headed 'n' until entry 4 is reached and then proceed right to the column headed α = 0.01. Then we find the cell in the table at the intersection of degrees of freedom n = 4 and α = 0.01 level. For your convenience, we give a part of the chi-square table as shown in Table 5.2.

Table 5.2: The Part of Chi-square Table

α =	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
n(df)=1	---	---	---	---	0.02	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55

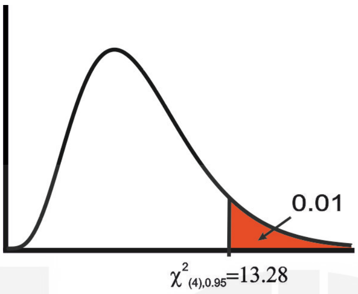


Fig. 5.8

Thus, we get the required right value of the chi-square statistic as $\chi^2_{(n),\alpha} = \chi^2_{(4),0.01} = 13.28$. The value of the chi-square statistic equal to 13.28 means, the probability that the chi-square statistic would exceed (greater than) 13.28 is 0.01 as shown in Fig. 5.8.

Left tail value

Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail values are the same in magnitude and opposite in sign as the right tail value and the t-table contains only tabulated values for α = 0.10, 0.05, 0.025, 0.01 and 0.005. But the chi-square is not symmetrical, therefore, to find out the left tail values using the chi-square table, the tabulated values for α = 0.995, 0.99, 0.975, 0.90, are also given. For α = 0.95, means there exists 0.95 probability/ area on the right side and the rest 1 – 0.95 = 0.05 on the left side of the curve. Hence, to read the left tail value from the chi-square table we have to see the value at 1 – α instead of α. To read the left tail tabulated value of the chi-square statistic, we follow the following steps:

Step 1: We start with the first column of the chi-square table, that is degrees of freedom and downward headed 'n' until the required degree of freedom is reached.

Step 2: After that we proceed right to the column headed α up to 1 – α is reached.

Step 3: We get the required left tail tabulated value in the cell of the table at the intersection of required degrees of freedom n and 1 – α.

We represent the left tail tabulated/critical value of the chi-square as $\chi^2_{(n),(1-\alpha)}$ which represents the probability that the chi-square statistic would be less than $\chi^2_{(n),(1-\alpha)}$ is α and greater than $\chi^2_{(n),(1-\alpha)}$ is 1 – α as shown in Fig. 5.7(b).

Suppose we want to read the tabulated value for which the left area is 0.05 and the degree of freedom is 4. To read this tabulated value, we start with the first column of the chi-square table, that is degrees of freedom and downward headed 'n' until entry 4 is reached and then proceed right to the column headed α to 1 – 0.05 = 0.95 instead of 0.05. Then we find the cell in the table at the intersection of degrees of freedom n = 4 and α = 0.95 level. Thus, we get the required left value of the chi-square statistic as $\chi^2_{(n),(1-\alpha)} = \chi^2_{(4),0.95}$ $= 0.71$. The tabulated value of the chi-square statistic equal to 0.71 means, the probability that the chi-square statistic would be less than 0.71 is 0.05 and greater than 0.71 is 0.95 as shown in Fig. 5.9.

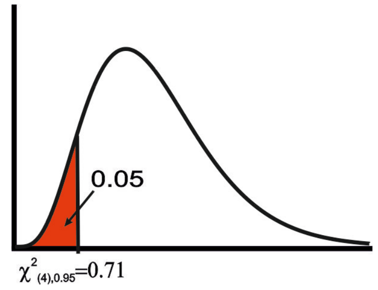


Fig. 5.9

Let us learn how to read both/two tails tabulated values from the chi-square table.

Both tails values

The procedure to read the two tails tabulated value of the chi-square statistic is almost similar to the t-table, we follow the following steps:

Step 1: First of all, we half the total area on both tails. If it is α then the half area i.e. α/2 lies in both tails as shown in Fig. 5.7(c).

Step 2: We then read the right tail and left tail tabulated values for the same degrees of freedom (n) and α/2 (area/probability lie on right and left tails) instead of α as discussed in the case of right and left tails tabulated values.

In this case, we represent the right-tail tabulated value as $\chi^2_{(n),\alpha/2}$ and left-tail as $\chi^2_{(n),(1-\alpha/2)}$.

For example, if we want to read the values of the chi-square statistic for which the area on both tails is 0.05 and the degrees of freedom is 6. For that first of all, we half the total area/probability as 0.05/2 = 0.025. After that, we read the right tail and left tail values in such a way that the area on both the tails remains α/2 = 0.025 as discussed in the case of right and left tails tabulated values cases. We get $\chi^2_{(n),\alpha/2} = \chi^2_{(6),0.025} = 14.45$ and $\chi^2_{(n),(1-\alpha/2)} = \chi^2_{(6),(1-0.025)}$ $= \chi^2_{(6),0.975} = 1.24$. These values are shown in Fig. 5.10.

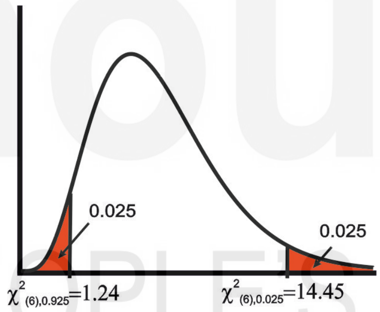


Fig. 5.10

After understanding how to read tabulated (critical) values for different cases, we now discuss some more facts about the chi-square table. If we closely inspect the chi-square table, then we observe:

- The table contains the tabulated values for the right tail as well as the left tail.
- As the degrees of freedom increase, the tabulated value decreases. The reason is that as we increase the degrees of freedom the tails of the chi-square distribution shift towards the right side as shown in Fig. 5.11.
- The chi-square table does not include entries for every possible degree of freedom. For example, the table lists continuously up to df = 30 and after that the chi-square values for df = 40, 60, 120 and does not list entries for degrees of freedom values between these.

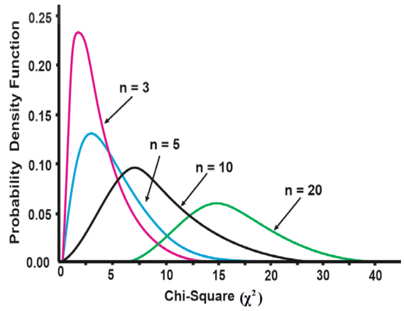


Fig. 5.11

Method of Finding the Values of Chi-square Statistic for Degrees of Freedom which are not Listed in the Table

We obtain the values of the chi-square statistic for degrees of freedom which are not listed in the table in a similar manner as discussed in the case of the t-table. This is explained in part (iv) of Example 5.

Now, let us do one example based on the above discussion.

Example 5: Find the values of the chi-square statistic for which the degrees of freedom (n) and area are given as follows:

(i) n = 2 and α = 0.05 (right tail)

(ii) n = 10 and α = 0.01 (left tail)

(iii) n = 8 and α = 0.05 (both tails)

(iv) n = 64 and α = 0.01 (right tail)

Solution:

(i) Here, we want to read the tabulated value of the chi-square statistic for

n = 2 and α = 0.05 (right tail)

Thus, we start from the first column of the chi-square table given in the Appendix and downward headed n until entry 2 is reached. Then proceed right to the column headed α = 0.05. We get the required tabulated value of the chi-square statistic as $\chi^2_{(n),\alpha} = \chi^2_{(2),0.05} = 5.99$.

(ii) Here, we want to find the value of the chi-square statistic for

n = 10 and α = 0.01 (left tail)

To read the left-tail tabulated value, we start with the first column of the chi-square table, that is, degrees of freedom and downward headed 'n' until entry 10 is reached and then proceed right to the column headed α at 1 – 0.01 = 0.99 instead of 0.01. Then we find the cell in the table at the intersection of degrees of freedom n = 10 and α = 0.99 level. Thus, we get the required left value of the chi-square statistic as $\chi^2_{(n),(1-\alpha)} = \chi^2_{(10),0.99} = 2.56$.

(iii) Here, we want to find the tabulated value of the chi-square statistic for two-tail area α = 0.05 and n = 8.

Since the total area on both tails is 0.05, therefore, half area 0.05/2 = 0.025 lies on both tails. In this case, the chi-square statistic has two values one on the right-tail as $\chi^2_{(n),\alpha/2} = \chi^2_{(8),0.025}$ and one on the left-tail as $\chi^2_{(n),(1-\alpha/2)} = \chi^2_{(8),0.975}$. So by proceeding the same way as above, we get the required values of the chi-square statistic as $\chi^2_{(8),0.025} = 17.53$ and $\chi^2_{(8),0.975} = 2.18$.

(iv) Here, we want to find the tabulated value of the chi-square statistic for the right tail and for n = 64 and α = 0.01.

Since the chi-square table does not have the tabulated value for 64 degrees of freedom so we need to interpolate it. For this, we find the tabulated values of the chi-square statistic that are just greater and just less than the degree of freedom 64 with α = 0.01. Thus, we have

$$\chi^2_{(70),\,0.01} = 100.42 \text{ and } \chi^2_{(60),\,0.01} = 88.38$$

There is a difference of 10 degrees of freedom between these two and a difference of 12.04 $(100.42 - 88.38)$ in the chi-square value. Thus, each degree of freedom has an approximate change in the value of the chi-square statistic as

$$\frac{12.04}{10} = 1.204$$

To get the value of the chi-square statistic for 64 degrees of freedom, we multiply 1.204 by 4 (64−60) and get

$$1.204 \times 4 = 4.816$$

Since the larger the degrees of freedom larger the tabulated value of the chi-square statistic, so adding 4.816 in 88.38, we get the required value as

$$\chi^2_{(64),0.01} = 88.38 + 4.816 = 93.196$$

After understanding how to read the tabulated values of the chi-square distribution, now you can assess yourself by solving the following Self Assessment Question.

SAQ 6

Find the tabulated value of the chi-square statistic for which the degrees of freedom and area are given as

(i) n = 11 and α = 0.01 (right tail)

(ii) n = 19 and α = 0.10 (left tail)

5.8 TABULATED VALUES OF F-DISTRIBUTION

As you have seen the F-distribution is described with the help of two degrees of freedom, one for the numerator (n_1) and the other for the denominator (n_2). Therefore, for each combination of degrees of freedom, the F-statistic has a different tabulated value, therefore, the statisticians organised the F-table somewhat differently than the tables for the other distributions. The statisticians prepare a separate F-table that is associated with different α (the area in the right tail of the distribution) and in each of these tables, the F values are given for various combinations of degrees of freedom. **Table VII** in the Appendix contains the tabulated (critical) values of the F-statistic for various degrees of freedom such that the area under the probability curve of the F-distribution to its right (upper) tail is equal to α = 0.10, 0.05, 0.025, 0.01 and 0.005. A part of the table is shown in Table 5.3.

The F-distribution table is also known as the F-table, F-score table, F-value table, or F-test table.

In the F-table, the first row of the F-table indicates the values of degrees of freedom for the numerator (n_1) and the first column on the left-hand side indicates the values of degrees of freedom for the denominator (n_2). The body of the table contains the value of the F-statistic for each particular pair of (n_1, n_2) and α which represents the critical value or point beyond which the area/probability of the F-distribution is α. **The area/probability (α) represents the proportions of the F-distribution contained in the right tail.**

<u>How to Use the F-table</u>

The area/probability, that is, α (level of significance) of the F-statistic may lie on the right tail, left tail or both tails so the tabulated value of the F-distribution is called right tail value, left tail value or both tail values. We now discuss how to read these tabulated values one at a time.

Right tail value

The F-table contains the tabulated values (it is also known as critical values, especially in testing of hypothesis) of the F-statistic for different degrees of freedom (n_1, n_2) such that the area under the probability curve of the F-distribution to its right tail (upper tail) is equal to α, therefore, for right tail value, we can read the table as such. To read the right tail tabulated value of the F-statistic, we follow the following steps:

Step 1: First of all, we select the table of required α because there is a separate F-table for each α.

Step 2: After selecting the required F-table, we start with the first row of the selected F-table, that is, degrees of freedom for the numerator (n_1) and move right until the required degree of freedom is reached.

Step 3: After that, we proceed downward headed 'denominator degrees of freedom (n_2)' until the required degree of freedom is reached.

Step 4: We get the required right tail tabulated value in the cell of the table at the intersection of the required degrees of freedom n_1 and n_2.

We represent the tabulated value of the F-statistic for (n_1, n_2) degrees of freedom and for right tail area/probability α as $F_{(n_1,n_2),\alpha}$ and is shown in Fig. 5.12 (a).

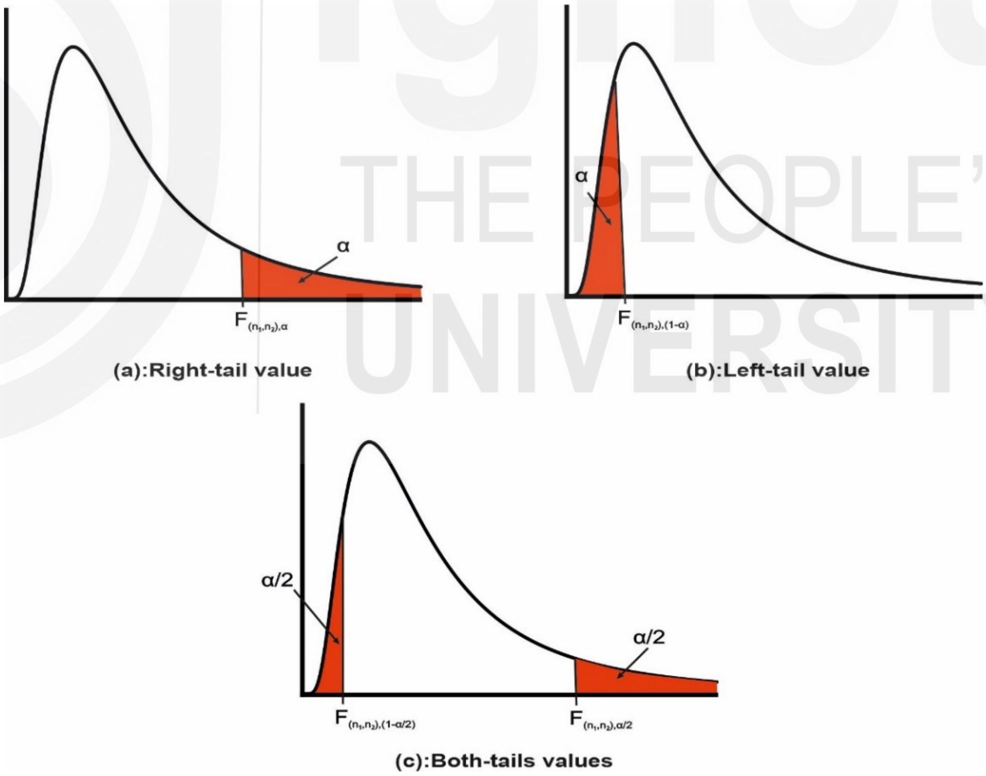


Fig. 5.12: Representation of tabulated value (s) of F-distribution.

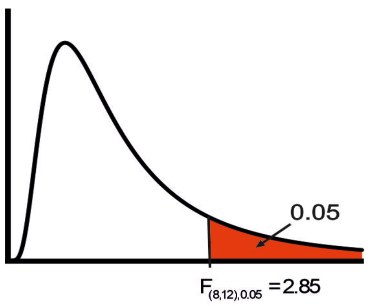


Fig. 5.13

Suppose we want to find out the tabulated value of the F-statistic for which the area/probability on the right tail is 0.05 and the degrees of freedom for the numerator is 8 and for the denominator is 12. To read the tabulated value, we first select the F-table corresponding to α = 0.05. Then we start with the first row of the selected F-table, that is degrees of freedom for the numerator and proceed right up to required $n_1 = 8$ is reached and then downward headed 'denominator degrees of freedom (n_2)' until required $n_2 = 12$ is reached. The

$$F_{(n_1,n_2),\alpha} = F_{(8,12),0.05} = 2.85$$ and is

shown in Fig. 5.13. For your convenience, we give a part of the F-table as follows:

Table 5.3: Part of F-table for α = 0.05

Degrees of freedom for denominator (n_2)	Degrees of freedom for numerator (n_1)											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42

Note 1: The value of the F-statistic (n_1, n_2) is not the same for (n_2, n_1) for the denominator, that is,

$$F_{(n_1,n_2)} \neq F_{(n_2,n_1)}$$

<u>Left tail value</u>

The F-distribution is not symmetrical as the t-distribution, therefore, the left tail value for which the left tail area/probability is α means there exists 1 – α area/probability on the right side of the value. Hence, to read the left tail value from the F-table, we have to see the value at 1 – α instead of α and then we use the relationship of the F-distribution as

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

It means that we can find the left-tail tabulated value of the F-statistic using the right-tail values by interchanging the degrees of freedom. We represent the left tail tabulated/critical value of the F-table as $F_{(n_1,n_2),(1-\alpha)}$. To read the left tail tabulated value of the F-statistic, we follow the following steps:

Step 1: First of all, we select the table of required α because there is a separate F-table for each α.

Step 2: After selecting the required F-table, we first read the tabulated value for the reverse degrees of freedom (n_2, n_1), that is, $F_{(n_2,n_1),\alpha}$ as discussed above.

Step 3: After that, we use the relationship of the F-distribution to find the left-tail tabulated values

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

Suppose we want to find out the tabulated value of the F-statistic for which the area/probability on the left tail is 0.01 and the degrees of freedom for the numerator is (n_1) 6 and for the denominator (n_2) 10. To read the tabulated value, we first select the F-table corresponding α = 0.01. After that we first

ignou THE PEOPLE'S UNIVERSITY

read the tabulated value for reversing the degrees of freedom, that is, $F_{(n_2,n_1),\alpha} = F_{(10,6),0.01}$ as discussed above. We get $F_{(10,6),0.01} = 7.87$. After that, we use the relationship to calculate the required left-tail value as

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

$$F_{(6,10),(1-0.01)} = F_{(6,10),(0.99)} = \frac{1}{F_{(10,6),0.01}} = \frac{1}{7.87} = 0.127$$

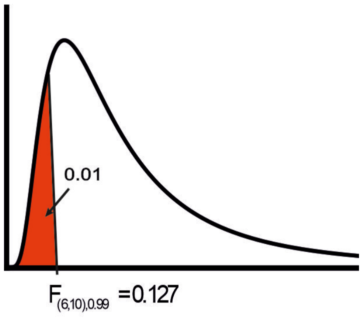


Fig. 5.14

This is also shown in Fig. 5.14.

Let us learn how to read both/ two tails tabulated values from the F-table.

Both tails values

To read the tabulated values of the F-statistic when the total area on both tails is α, we follow the same procedure as in the case of the t-table and chi-square table. We follow the following steps:

Step 1: First of all, we half the total area on both tails. If it is α then the half area i.e. α/2 lies in both tails as shown in Fig. 5.12 (c).

Step 2: We then read the right tail and left tail tabulated values for the same degrees of freedom (n) and α/2 (area/probability lie on right and left tails) instead of α as discussed in the case of right and left tails tabulated values.

In this case, we represent the right-tail tabulated value as $F_{(n_1,n_2),\alpha/2}$ and left-tail as $F_{(n_1,n_2),(1-\alpha/2)}$.

For example, if we want to read the values of the F-statistic for which the area on both tails is 0.05 and the degrees of freedom (4, 12). For that first of all, we half the total area/probability as 0.05/2 = 0.025. After that, we read the right-tail and left-tail values in such a way that the area on both the tails the remaining α/2 = 0.025 as discussed in the case of right and left tails tabulated values then, we get $F_{(n_1,n_2),\alpha/2} = F_{(4,12),0.025} = 4.12$ and $F_{(n_1,n_2),(1-\alpha)} = F_{(4,12),0.975} = \frac{1}{F_{(12,4),0.025}} = \frac{1}{8.75}$ $= 0.11$.

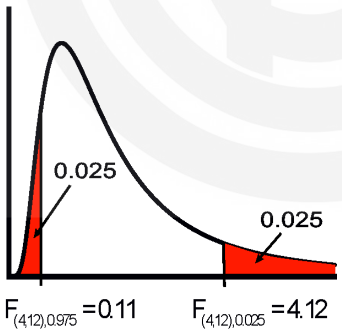


Fig. 5.15

These tabulated values are also shown in Fig. 5.15.

After understanding how to read tabulated (critical) values for different cases, we now discuss some more facts about the F-table. If we closely inspect the F-table then we observe:

- As the degrees of freedom either for the numerator or denominator increase, the tabulated value of the F-statistic decreases. The reason is that as we increase the degrees of freedom the tails of the F-distribution shift towards the left side as shown in Fig. 5.16.

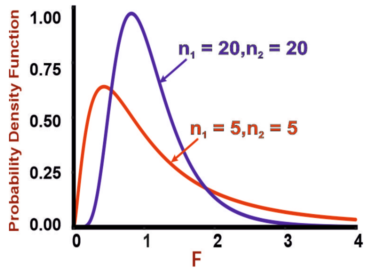


Fig. 5.16

- The value of the F-statistic (n1, n2) is not the same for (n2, n1) for the denominator, that is,

$$F_{(n_1,n_2)} \neq F_{(n_2,n_1)}$$

- The F-table does not include entries for every pair of possible degrees of freedom. For example, the table lists continuously up to $n_1 = n_2 = 30$ and after that for df = 40, 60, 120 and does not list any entries for degrees of freedom values between these.

Method of Finding the Values of F-Statistic for Degrees of Freedom which are not Listed in the Table

We obtain the values of the F-statistic for degrees of freedom which are not listed in the table in a similar manner as discussed in the case of the t-table. This is explained in part (iv) of Example 6.

Now, let us try to read the tabulated values for F-statistic by taking an example.

Example 6: Find the values of the F-statistic in each part for which the degrees of freedom (n_1, n_2) and the area (α) are as follows:

(i) $n_1 = 15$, $n_2 = 8$ and $\alpha = 0.01$ (right tail)

(ii) $n_1 = 7$, $n_2 = 4$ and $\alpha = 0.05$ (left tail)

(iii) $n_1 = 5$, $n_2 = 10$ and $\alpha = 0.10$ (both tails)

(iv) $n_1 = 10$, $n_2 = 32$ and $\alpha = 0.01$ (right tail)

Solution:

(i) Here, we want to find the tabulated value of the F-statistic for which $n_1 = 15$, $n_2 = 8$ and area/probability on the right tail is $\alpha = 0.01$. Thus, we first select the F-table for $\alpha = 0.01$ and then we start with the first row of the selected F-table, that is degrees of freedom for the numerator and proceed right until we reach $n_1 = 15$ and then downward headed 'denominator degrees of freedom (n_2)' until we reach $n_2 = 8$. We get the required tabulated value of the F-statistic as $F_{(n_1,n_2),\alpha} = F_{(15,8),0.01} = 5.52$.

(ii) Here, we want to find the tabulated value of the F-statistic for which $n_1 = 7$, $n_2 = 4$ and area/probability on the left tail is $\alpha = 0.05$.

Since we have to read the tabulated value for the left tail, so we use the relation:

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

We first select the F-table corresponding $\alpha = 0.05$. After that we read the tabulated value by reversing the degrees of freedom, that is, for (4, 7). We get $F_{(n_2,n_1),\alpha} = F_{(4,7),0.05} = 4.12.$ After that, we use the relationship to calculate the required left-tail value as

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

$$F_{(7,4),(1-0.05)} = F_{(7,4),0.95} = \frac{1}{F_{(4,7),0.05}} = \frac{1}{4.12} = 0.243$$

(iii) Here, the total area on both tails is 0.10, therefore, half area 0.10/2 = 0.05 lies on both tails. In this case, the F-statistic has two values one on right-tail as $F_{(n_1,n_2),\alpha/2} = F_{(5,10),0.05}$ $F_{(n_1,n_2),(1-\alpha/2)}$

$= F_{(5,10),(1-0.05)} = F_{(5,10),0.95}$. So, by proceeding the same way as above, we get the required tabulated values of the F-statistic as $F_{(5,10),0.05} = 3.33$ and

$$F_{(5,10),0.95} = \frac{1}{F_{(10,5),0.05}} = \frac{1}{4.74} = 0.21.$$

(iv) Here, we want to find the tabulated value of the F-statistic for

$n_1 = 10$, $n_2 = 32$ and $\alpha = 0.01$(right tail)

Since the F-table for $\alpha = 0.01$ does not have the tabulated value corresponding to degrees of freedom $n_1 = 10$, $n_2 = 32$ so we need to interpolate it.

For this, we find the tabulated values of the F-statistic that are just greater and just less than the degree of freedom 32. Thus, we have

$F_{(10,\ 40),\ 0.01} = 2.80 \quad F_{(10,\ 30),\ 0.01} = 2.98$

There is a difference of 10 degrees of freedom between these two and a difference of 0.18 (2.98 – 2.80) in the F-values. Thus, each degree of freedom has an approximate change in the value of the F-statistic as

$$\frac{0.18}{10} = 0.018$$

To get 32 degrees of freedom, multiplying 0.018 by 2 (32 – 30), we get

$$0.018 \times 2 = 0.036$$

Since the larger the degrees of freedom smaller the tabulated value of the F-statistic, therefore, subtracting this from 2.98, we get the required values as

$$F_{(10,\ 32)} = 2.98 - 0.036 = 2.944$$

Now, try to answer the following Self Assessment Question.

SAQ 7

Find the tabulated value of the F-statistic in the following cases:

(i) $n_1 = 8$, $n_2 = 10$ and $\alpha = 0.05$ (right tail)

(ii) $n_1 = 4$, $n_2 = 12$ and $\alpha = 0.01$ (left tail)

Let us end with a brief look at what we have covered in this unit.

5.9 SUMMARY

In this unit, we have covered the following points:

- The probability density function of the F-distribution with (n_1, n_2) degrees of freedom is given by

$$f(F) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{F^{\frac{n_1}{2}-1}}{\left(1 + \frac{n_1}{n_2}F\right)^{\frac{n_1+n_2}{2}}}; \quad 0 < F < \infty$$

- The F distribution is a positively skewed distribution that has a minimum value of 0, but no maximum value. The shape of the F-distribution for n_1 =1, 2 and n_2 = 1, 2 is inverse J alphabet and the curve starts out high and then subsequently declines. However, for n_1 and n_2 > 2, the shape of the F-distribution first increases and attains the maximum value and after that starts to decrease.
- The mean and variance of the F-distribution with (n_1, n_2) df are given as follows:

  Mean = $\frac{n_2}{n_2 - 2}$ for $n_2 > 2$.

  Variance = $\frac{2n_2^2\left(n_1 + n_2 - 2\right)}{n_1\left(n_2 - 2\right)^2\left(n_2 - 4\right)}$ for $n_2 > 4$.

- If we interchange the degrees of freedom n_1 and n_2 then there exists a very useful relation as

  $$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

- The applications of the F-distribution.
- The method of reading the tabulated value from the t, chi-square and F-distributions tables.

5.10 TERMINAL QUESTIONS

Write down the pdf of the F-distribution in each of the following cases:

(i) (4, 2) degrees of freedom

(ii) (2, 4) degrees of freedom

And show that the pdf of the F-distribution is not the same by reversing the degrees of freedom.

5.11 SOLUTIONS / ANSWERS

Self Assessment Questions (SAQs)

1. If random variable X follows the F-distribution with (n_1, n_2) degrees of freedom, then the probability density function of X is given as

$$f(F) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \frac{F^{\frac{n_1}{2}-1}}{\left(1 + \frac{n_1}{n_2}F\right)^{\frac{n_1+n_2}{2}}}; \quad 0 < F < \infty$$

We now try to convert the given pdf in the form of the standard form of the F-distribution so that we can compare and find the degrees of freedom. Since the coefficient of x in $\left(1 + \frac{x}{4}\right)^{\frac{5}{2}}$ $\frac{1}{4}$

that $n_1 = 1$ and $n_2 = 4$ because the coefficient of x in $\left(1+\frac{n_1}{n_2}x\right)$ is $\frac{n_1}{n_2}$. We try to write the given pdf into the F-distribution as follows:

$$f(x) = \frac{\left(\frac{1}{4}\right)^{\frac{1}{2}}}{B\left(\frac{1}{2},\frac{4}{2}\right)} \frac{x^{\frac{1}{2}-1}}{\left(1+\frac{1}{4}x\right)^{\frac{1+4}{2}}}; \quad 0 < x < \infty$$

$$\left[B\left(\frac{1}{2},2\right) = \frac{\overline{\left|\frac{1}{2}\right.}\overline{|2}}{\overline{\left|\frac{5}{2}\right.}} = \frac{\overline{\left|\frac{1}{2}\right.}\overline{|2}}{\frac{3}{2}\times\frac{1}{2}\times\overline{\left|\frac{1}{2}\right.}} = \frac{4}{3} \right]$$

By comparing the above form with the standard form, we get $n_1 = 1$ and $n_2 = 4$.

We know that the mean and variance of the F-distribution with (n_1, n_2) degrees of freedom are:

Mean $= \frac{n_2}{n_2 - 2}$ for $n_2 > 2$ and Variance $= \frac{2n_2^2(n_1 + n_2 - 2)}{n_1(n_2-2)^2(n_2-4)}$ for $n_2 > 4$.

Therefore, for $n_1 = 1$ and $n_2 = 4$ degrees of freedom

$$\text{Mean} = \frac{n_2}{n_2 - 2} = \frac{4}{4-2} = 2$$

Since the variance of the F-distribution exists for $n_2 > 4$, therefore, for $n_1 = 4$ it does not exist.

2. As we know if a variable t follows the t-distribution with n df, then the square of t follows the F-distribution with (1, n) df, therefore, for n = 4, the distribution of t square follows the F-distribution with (1, 4) df. The pdf of the F-distribution is given by

$$f(F) = \frac{\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}}}{B\left(\frac{n_1}{2},\frac{n_2}{2}\right)} \frac{F^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}F\right)^{\frac{n_1+n_2}{2}}}; \quad 0 < F < \infty$$

$$f(F) = \frac{\left(\frac{1}{4}\right)^{\frac{1}{2}}}{B\left(\frac{1}{2},\frac{4}{2}\right)} \frac{F^{\frac{1}{2}-1}}{\left(1+\frac{1}{4}F\right)^{\frac{1+4}{2}}} = \frac{3F^{-\frac{1}{2}}}{8\left(1+\frac{1}{4}F\right)^{\frac{5}{2}}} \quad \left[\text{Since } B\left(\frac{1}{2},2\right) = \frac{4}{3}\right]$$

3. Refer to Section 5.4.

4. Refer to Section 5.5.

5. **(i)** Here, we want to find the value of the t-statistic for the right tail corresponding to

n = 9 and α = 0.01

Therefore, we start with the first column of the t-table (Table V) given in the Appendix and downward headed n until entry 9 is reached. Then proceed right to the column headed α = 0.01. We get the required value of the t-statistic as $t_{(n),\alpha} = t_{(9),0.01} = 2.821$.

(ii) In this part, we have to read the tabulated value for the left tail corresponding to

$n = 15$ and $\alpha = 0.05$

First of all, we read the right tail value by proceeding the same way as in part (i) as $t_{(n),\alpha} = t_{(15),0.05} = 1.753$. Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail value will be $-t_{(n),\alpha} = -t_{(15),0.05} = -1.753$.

(iii) Here, we want to read the tabulated values of the t-statistic for the two-tail corresponding to

$n = 13$ and $\alpha = 0.05$

Since the total area on both tails is 0.05, therefore, the area on the right tail as well as on the left tail will be 0.05/2 = 0.025. Thus, we start with the first column of the t-table and downward headed n until entry 13 is reached. Then proceed right to the column headed 0.025. We get $t_{(n),\alpha/2} = t_{(13),0.025} = 2.160$. Since the t-distribution is symmetrical at the t = 0 line, therefore, the left tail value will be the same as the right tail value but in a negative sign. Therefore, $-t_{(n),\alpha/2} = -t_{(13),0.025} = -2.160$. So the required values of the t-statistic are $\pm t_{(n),\alpha/2} = \pm t_{(13),0.025} = \pm 2.160$.

6. **(i)** Here, we want to find the value of the chi-square statistic for

$n = 11$ and $\alpha = 0.01$ (right tail)

We start with the first column of the chi-square table given in the Appendix and downward headed n until entry 11 is reached. Then proceed right to the column headed α = 0.01. We get the required value of the chi-square statistic as $\chi^2_{(n),\alpha} = \chi^2_{(11),0.01} = 24.72$.

(ii) Here, we want to find the value of the chi-square statistic for

$n = 19$ and $\alpha = 0.10$ (left tail)

To read the left-tail tabulated value, we start with the first column of the chi-square table, that is degrees of freedom and downward headed 'n' until entry 19 is reached and then proceed right to the column headed α at 1 – 0.1 0= 0.90 instead of 0.10. Then we find the cell in the table at the intersection of degrees of freedom n = 19 and α = 0.90 level. Thus, we get the required left value of the chi-square statistic as $\chi^2_{(n),(1-\alpha)} = \chi^2_{(19),0.90} = 11.65$.

7. **(i)** Here, we want to find the value of the F-statistic for

$n_1 = 8$, $n_2 = 10$ and $\alpha = 0.05$ (right tail)

Thus, we first select the F-table for α = 0.05 and then we start with the first row of the selected F-table, that is, degrees of freedom for the numerator and proceed right until we reach $n_1 = 8$ and then downward headed denominator degrees of freedom (n_2) until we reach $n_2 = 10$. We get the required tabulated value of the F-statistic as $F_{(n_1,n_2),\alpha} = F_{(8,10),0.05} = 3.07$.

(ii) Here, we want to find the tabulated value of the F-statistic for which $n_1 = 4$, $n_2 = 12$ and area/probability on the left tail is α = 0.01.

Since we have to read the tabulated value for the left tail, so we use the relation

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

We first select the F-table corresponding to α = 0.01. After that we read the tabulated value by reversing the degrees of freedom, that is, for (12, 4). We get $F_{(n_2,n_1),\alpha} = F_{(12,4),0.01} = 14.37$. After that, we use the relationship to calculate the required left-tail value as

$$F_{(n_1,n_2),(1-\alpha)} = \frac{1}{F_{(n_2,n_1),\alpha}}$$

$$F_{(4,12),(1-0.01)} = F_{(4,12),(0.99)} = \frac{1}{F_{(12,4),0.01}} = \frac{1}{14.37} = 0.07$$

Terminal Questions (TQs)

For $n_1 = 4$ and $n_2 = 2$, the pdf of the F-distribution is given by

$$f(F) = \frac{\left(\frac{4}{2}\right)^{\frac{4}{2}}}{B\left(\frac{4}{2},\frac{2}{2}\right)} \frac{F^{\frac{4}{2}-1}}{\left(1+\frac{4}{2}F\right)^{\frac{4+2}{2}}}; \quad 0 < F < \infty$$

$$= \frac{(2)^2}{B(2,1)} \frac{F^{2-1}}{(1+2F)^{\frac{6}{2}}} = \frac{8F}{(1+2F)^3} \left(\text{since } B(2,1) = \frac{1}{2}\right)$$

Similarly, $n_1 = 2$ and $n_2 = 4$, the pdf of the F-distribution is given by

$$f(F) = \frac{\left(\frac{2}{4}\right)^{\frac{2}{2}}}{B\left(\frac{2}{2},\frac{4}{2}\right)} \frac{F^{\frac{2}{2}-1}}{\left(1+\frac{2}{4}F\right)^{\frac{2+4}{2}}}; \quad 0 < F < \infty$$

$$= \frac{\left(\frac{1}{2}\right)^1}{B(1,2)} \frac{F^{1-1}}{\left(1+\frac{1}{2}F\right)^{\frac{6}{2}}} = \frac{1}{\left(1+\frac{1}{2}F\right)^3} \quad \left(\text{since } B(1,2) = \frac{1}{2}\right)$$

Hence, we can say that

$$F_{(n_1,n_2)} \neq F_{(n_2,n_1)}$$