Published on *Explorable.com* (https://explorable.com)

F-Distribution

The F-distribution, also known as the Snedecor's F-distribution or the Fisher-Snedecor distribution (after R.A. Fisher and George W. Snedecor), is the distribution of ratios of two independent estimators of the population variances.

Suppose we have two samples with n1 and n2 observations, the ratio F = s12 / s22 where s12 and s22 are the sample variances, is distributed according to an F-distribution with v1 = n1-1 numerator degrees of freedom, and v2 = n2-1 denominator degrees of freedom.

For example, if F follows an F-distribution and the degrees of freedom for the numerator are 4 and the degrees of freedom for the denominator are 10, then F ~ F4,10. For each combination of these degrees of freedom there is a different F-distribution. The F-distribution is most spread out when the degrees of freedom are small. As the degrees of freedom increase, the F-distribution is less dispersed.

Properties

The F-distribution has the following properties:

The mean of the distribution is equal to v1 / (v2 - 2). The variance is equal to [v22 * (v1 + 2)] / [v1 * (v2 - 2) * (v2 - 4)]

The F-distribution is skewed to the right, and the F-values can be only positive. The curve reaches a peak not far to the right of 0, and then gradually approaches the horizontal axis. The F-distribution approaches, but never quite touches the horizontal axis.

Uses

The main use of F-distribution is to test whether two independent samples [1] have been drawn for the normal populations [2] with the same variance, or if two independent estimates of the population variance are homogeneous or not, since it is often desirable to compare two variances [3] rather than two averages [4]. For instance, college administrators would prefer two college professors grading exams to have the same variation in their grading. For this, the F-test [5] can be used, and after examining the p-value [6], inference can be drawn on the variation.

Assumptions

In order to perform F-test of two variances, it is important that the following are true:

- The populations [7] from which the two samples are drawn are normally distributed [2].
- The two populations are independent of each other.

If the two populations have equal variances, then s12 and s22 are close in value and F is close to 1. But if the two population variances are very different, s12 and s22 tend to be very different, too.

Choosing s12 as the larger sample variance causes the ratio to be greater than 1. If s12 and s22 are far apart, then F is a large number. Therefore, if F is close to 1, the evidence favours the null hypothesis [8] (the two population variances are equal). But if F is much larger than 1, then the evidence [9] is against the null hypothesis, and we can infer that possibly the population variances differ to a large extent.

Anova and F

In the technique known as Analysis of Variance [10] (ANOVA) which plays a very important role in Design of Experiments [11], the variance ratio test is applied to test the significance [12] of different components of variation against error variation.

For example, a new drug for treating Osteoporosis could need to be field tested. Since severity of this disease is generally a function of age, the new drug could be administered randomly to n patients in each age group. Put differently, this would be an experiment in m age groups and n different dosage levels of the drug allocated randomly to the patients. With figures provided from patients for each age group x dose combination, we can use the variance ratio test (F- test [5]) to test for difference between dose levels and if this variation can be attributed to chance.

The other uses include testing the significance of the correlation [13] ratio between two random variables, and to test the linearity of regression [14].

Source URL: https://explorable.com/f-distribution

Links
[1] https://explorable.com/sample-group
[2] https://explorable.com/normal-probability-distribution
[3] https://explorable.com/statistical-variance
[4] https://explorable.com/arithmetic-mean
[5] https://explorable.com/f-test
[6] https://explorable.com/p-value
[7] https://explorable.com/population-sampling
[8] https://explorable.com/null-hypothesis
[9] https://explorable.com/empirical-evidence
[10] https://explorable.com/anova
[11] https://explorable.com/design-of-experiment
[12] https://explorable.com/significance-test
[13] https://explorable.com/statistical-correlation
[14] https://explorable.com/linear-regression-analysis