SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
VOL. 3, ISSUE 8, FEBRUARY 2026
Available Online : http://shodhprakashan.in/

THE USE OF F-TEST ON CPI DATA

Dr. Kavita Shrivastava[1], Dr. Jyoti Nema[2], Harshita Pandey[3]

[1]HOD, Department of Mathematics

[2]Guest Faculty, Department of Mathematics

[3]B.Sc. IV Year (Honours)

[123]Saroijini Naidu Government Girls P.G. (Autonomous) College, Bhopal

Abstract

This study examines whether there is a statistically significant difference in the variability of corruption levels between European and Asian countries using the Corruption Perception Index (CPI). The CPI scores, published by Transparency International, are used as a proxy for perceived public-sector corruption. An F-test is employed to compare the variances of CPI scores for selected countries from Europe and Asia. Prior to analysis, the data are tested for normality to satisfy the assumptions of the F-test. The results indicate whether the dispersion of corruption perception differs significantly between the two regions, providing insights into regional heterogeneity in governance quality. The findings contribute to comparative institutional analysis and offer empirical evidence on regional differences in corruption perception, which may be useful for policymakers, researchers, and international organizations concerned with governance reforms and anti-corruption strategies.

Introduction

The F-test is a fundamental statistical method used to compare variances and to assess the overall significance of statistical models. Its origin is closely associated with the development of modern statistical theory in the early twentieth century. The test is based on the F-distribution, which describes the ratio of two independent sample variances drawn from normally distributed populations.

Corruption is a pervasive issue that affects economic development, governance, and social welfare across the globe. The Corruption Perceptions Index (CPI) is a widely used metric to measure the perceived levels of corruption in countries worldwide. This study aims to compare the CPI scores of countries in Europe and Asia to determine if there is a significant difference in the levels of corruption between these two regions. Using the F-test, we will examine the variances of CPI scores to assess whether corruption patterns are homogeneous or heterogeneous across European and Asian countries.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
VOL. 3, ISSUE 8, FEBRUARY 2026
Available Online : http://shodhprakashan.in/

Literature Review

The F-test is a fundamental statistical method used to compare variances and to assess the overall significance of statistical models. Its origin is closely associated with the development of modern statistical theory in the early twentieth century. The test is based on the F-distribution, which describes the ratio of two independent sample variances drawn from normally distributed populations.

The F-test was developed by Sir Ronald A. Fisher, a pioneering British statistician, geneticist, and evolutionary biologist. During the 1920s, Fisher introduced the F-distribution while working on experimental data analysis in agriculture at the Rothamsted Experimental Station in England. His work aimed to provide reliable methods for analyzing variability in experimental results and distinguishing systematic effects from random errors. [1]

Theoretical Foundations

The F-test theory in a research paper involves using the F-statistic, a ratio of two variances, to test hypotheses often comparing the distribution of variance. It compares variability *between* groups to variability *within* groups.

Core Concepts

1. **F-Statistic:** A positive value calculated as the ratio of two variances (Variance 1 / Variance 2), often representing between-group variance to within-group variance.
2. **F-Distribution:** A right-skewed probability distribution, defined by two degrees of freedom (numerator and denominator), used for F-tests.
3. **Null Hypothesis :** It is default assumption that no significant effect, relationship, or difference exists between variables being studied. [1]

The interpretation depends on comparing the calculated F-Statistic with a critical value

1. Significant Result ($F > F_{critical}$): Reject the null hypothesis. This indicates that the group of differences is statistically significant and unlikely to occur by chance.
2. Non-Significant result ($F \leq F_{critical}$): Fail to reject the null hypothesis. This suggests it that any observed differences are due to random variation alone.

CLASSICAL F-TEST FOR VARIANCE COMPARISON

Two-Sample F-Test : The most basic application of the F-test is comparing the variances of two independent normally distributed data .

Consider two independent samples: -

Sample 1: $X1, X2, \ldots, Xn1 \sim N(\mu 1, \sigma 2\ 1)$

Sample 2: $Y1, Y2, \ldots, Yn2 \sim N(\mu 2, \sigma 2\ 2)$

The null and alternative hypotheses are:

$$H0 : \sigma_1^2 = \sigma_2^2$$
$$H1 : \sigma_1^2 \neq \sigma_2^2$$

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
VOL. 3, ISSUE 8, FEBRUARY 2026
Available Online : http://shodhprakashan.in/

The test statistic is:

$\mathbf{F = S_1^2 / S_2^2}$
where S_1^2 and S_2^2 are the sample variances

$$\mathbf{S_1^2} = \frac{1}{n1-1}\sum_{i=1}^{n}(Xi - X)^2$$

$$\mathbf{S_1^2} = \frac{1}{n2-1}\sum_{i=1}^{n}(Xi - X)^2$$
Under H_0, the test statistic follows an F-distribution with degrees of freedom $(n_1 - 1, n_2 - 1)$.

[2]

Assumptions and Limitations
The validity of the F-test relies heavily on several critical assumptions:

- **Normality**: The samples are drawn must be normally distributed.
- **Independence**: The samples must be drawn randomly and independently.
- **Homogeneity of Variance**: For some F-tests (like in ANOVA), the population variances are assumed to be equal, although the test itself can also be used to test this assumption.

A key limitation of the F-test is its sensitivity to the violation of these assumptions, particularly normality. If data significantly deviate from a normal distribution, the results of the F-test may be unreliable. [3]

Application of F test on sample taken from Corruption Perception Index

The given data of Corruption Perception Index is annually released by Transparency International, ranking countries by perceived public sector corruption (0 = highly corrupt , 100 = very clean)based on expert /business views. Here the data is of countries of western and central Europe and emerging economies of Asia.

CPI Score of Countries of Western and Central Europe [5]

Countries	**Score (X_i)**	**=(X_i - X)**	**= (X_i - X)^2**
Italy	54	3.6	12.96
Poland	53	2.6	6.76
Slovakia	49	-1.4	1.96
Greece	49	-1.4	1.96
Croatia	47	-3.4	11.56
	Mean = 50.4		Mean =35.2

Variance = 35.2/5 =7.04

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
VOL. 3, ISSUE 8, FEBRUARY 2026
Available Online : http://shodhprakashan.in/

CPI Score of Asian Countries [5]

Countries	Score (X_i)	=(X_i - X)	= (X_i - X)^2
Malaysia	50	4.8	23.04
Oman	55	9.8	96.04
China	43	-2.2	4.84
India	38	-7.2	51.84
Kazakhstan	40	-5.2	27.04
	Mean = 45.2		Sum = 203.52

Variance =257.25/5 =51.4

The F – Statistic : The F-value is ratio of the two variances (Always put the larger variance in numerator).

$F = S_1^2 / S_2^2$

Where $S_1^2 = \frac{1}{n1-1}\sum_{i=1}^{n}(Xi - X)^2$ = 35.2/4 = 8.8

$S_1^2 = \frac{1}{n2-1}\sum_{i=1}^{n}(Xi - X)^2$ = 203.52/4 = 50.88

F= 50.88/8.8 = 5.781

The F -critical : The value is taken from the F-Table for the significance level =0.05

Here the significance level means you are willing to accept a 5% risk of concluding that a significant difference exists when it actually does not.

F_{crit} = 6.39

The Statistical decision

From above calculations it is clear that : **F < F_{crit}**

This implies there is no statistically significant difference . It fail to reject the null hypothesis. It demonstrate that there is no statistically significant difference in the variance. This implies the variance (spread) or means (averages) of corruption levels in both the continents is statistically different.

Conclusion

In conclusion , the F-test for compairing the variance is indispensable tool in analysis of different data , providing rigorous method to test hypothesis about the group means and understand the relationships within the group of data. By applying these statistical techniques economists, businessman, policy makers can draw reliable conclusions about the effectiveness of policies across different groups and periods.

SHODH PRAKASHAN
An International Open Access, Peer Reviewed, Refereed Journal
ISSN : 2583-9861
VOL. 3, ISSUE 8, FEBRUARY 2026
Available Online : http://shodhprakashan.in/

References:

[1] Gujarati, D. N., and Porter, D. C. (2009). Basic econometrics (5th ed.) McGraw-Hill.

[2] Mustafa Cavus ,Berna Yazici, comparison of two approaches for the generalized F-test

[3] Robert Odek, Gordon Opuodho, F-test and P- values: A sysnopsis, Journal of Management and science

[4] Sajan Khatri , An analysis of selected public sector banks, ijsrst

[5] The Transparency International, Corruption Perception Index, https://www.transparency.org/en/cpi