Journal of Management and Science

13(2) (2023) 59-61
OPEN ACCESS

RESEARCH ARTICLE

DOI:10.26524.jms.13.22

F-test and P-values: A synopsis

Robert Odek [a]*, Gordon Opuodho [a]

Abstract

F-tests are a ratio of variances. F-tests are utilized in comparison of two variances, Analysis of Variance (ANOVA), and multiple regression. On the other hand, P-values depict the extent to which data fit in the design projected by the hypothesis so that P-value test all the assumptions contained on how the data was produced. This article offers a synopsis of what exactly is F-test and P-values, and their usage.

Keywords: F-test, p-value.

Author Affiliation: [a] Department of Economics, Accounting & Finance, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya, East Africa.

Corresponding Author: Robert Odek. Department of Economics, Accounting & Finance, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya, East Africa.

Email: roba.odek@gmail.com

How to cite this article: Robert Odek, Gordon Opuodho, (2023). F-test and P-values: A synopsis, 13(2) 59-61.Retrieved from https://jmseleyon.com/index.php/jms/article/view/668

Received: 21 April 2023 **Revised:** 24 May 2023 **Accepted:** 24 June 2023

1. INTRODUCTION

George W. Snedecor named this statistical tool known as the F-test after its proponent Sir. Ronald Fisher. Studies suggest that F-tests are utilized in comparison of two variances, Analysis of Variance (ANOVA), and multiple regression.[1] Advances that the F-test is a test to ascertain if the linear regression avails the line of best fit to a particular data set. [2] F-tests are therefore a ratio of variances. Consequently, any kind of statistical assessment that employs F-values are referred to as F-test. On the other hand, P-values portray the extent to which data fit in the design anticipated by the hypothesis and all other model assumptions as described in the study under consideration. P-values therefore test all the assumptions encompassed on how the data was generated and not necessarily what was hypothesized.

1.1 Critical conditions for use of F-test

Normality

The data under study need to be normally distributed. [3] Normal distribution implies the data needs to follow the pattern of the Gaussian curve which presents data to be systematic.

Independent and random choice of study variables.

The identification of the study sample should be independent and at the same time representative or random.

1.2 Steps in calculating F-test

a). State the null hypothesis and alternative hypothesis

H0 : share price is a measure of company value (H0 :$\mu1= \mu2$)

H1: share price is not a measure of company value (H1 :$\mu1\neq \mu2$)

b). Calculating the F-value as below

F = (SSE1 – SSE2 / m) / SSE2 / n-k,

where SSE = residual sum of squares,

m = number of restrictions and

k = number of independent variables.

c). Find the F-Statistic (the critical value for this test)

F = explained variance
unexplained variance

d) Decision

Support or reject the hypothesis

F > F critical value then reject the null hypothesis

1.3 F-test, p-value, and Analysis of Variance (ANOVA)-Illustration 1

Consider the below hypothetical data generated from an SPSS data set.

Table 1: ANOVA on governance indicators on Performance Audit Annual Reports in the NG-CDFs, Kenya

	Sum of Squares	df	Mean Square	F-value	Sig.p
Regression	8.66	4	2.165	43.143	.000b
Residual	4.416	88	0.05		
Total	13.076	92			

a. Predictors: (Constant), corruption control, regulation quality, governance effectiveness, accountability

b. Dependent Variable: Performance Audit Annual Reports

Source: Odek & Oyugi (2021) [4]

In view of the foregoing table 1, ANOVA results show that F-value (F=43.143) is significant (p=0.000). In other words, the p-value (0.000) is less than the significance level (F=43.143), hence the above data provide sufficient evidence to conclude that the regression model fits the data better.

Further, the analysis of Variance in Table 1 above is statistically significant suggesting that governance indicators (corruption control, regulation quality, governance effectiveness, accountability) have a statistically significant influence on the performance audit annual reports in the NG-CDFs in Kenya. This is on the basis that F statistic value of 43.143 is greater than the critical value at 0.05 significance level as interpreted below

F statistic = 43.143 > F critical = 2.215 (4, 88).

1.4 F-test, p-value, and Multiple Regression

1.4.1 Objectives of multiple regression

i. To define and comprehend the association among the study variables
ii. To envisage
iii. To adjust and control a process

1.4.1 Multiple regression-Illustration 2 (One way ANOVA)

Table2: Model of fitness representing governance indicators and Performance Audit Annual Reports in the NG-CDFs, Kenya

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.814a	0.662	0.647	0.224

a Predictors: (Constant), corruption control, regulation quality, governance effectiveness, accountability

b Dependent Variable: Performance Audit Annual Reports

Source: Odek & Oyugi (2021) [4]

Where;

R: tells fit of the model and rarely utilised

a(constant): provides the forecasted value for Y other factors 0

b(regression coefficients): specifies the effect of the independent variables on the dependent variable after adjustments. This advises the formation of the equation $Y = \beta 0+ \beta 1X1+\beta 2X2+\beta 3X3+\beta 4X4+\mu$

R square (Coefficient of determination) - measures the strength of the relationship between the study model and the dependent variable. In other words, it portrays the percentage of variation in Y explained by the X variables

1.4.2 Forming Inferences

i. When R2 is smaller than the critical value in the table, then the model is not significant. Accept the null hypothesis that x variables do not aid in predicting y variable.
ii. When R2 is larger than the critical value in the table, then the model is significant. Reject the null hypothesis and accept the alternative hypothesis that x variables do not aid in predicting y variable.

Using table 2 as an example, we intend to utilize the F-test to form inferences in the analyzed data for multiple regression.

This is explained using the hypotheses;

H0 :µ1= µ2= µ3= µ4 (The four independent variables (I.V) will significantly explain the variance in performance audit reports)

H1 :µ1≠ µ2≠ µ3≠ µ4 (The four independent variables (I.V) will not significantly explain the variance in performance audit reports)

Based on the above data, the null hypothesis can be accepted. R2 is 66.2% and this is the explained variance to mean that the four Intendent Variables explain the dependent variable up to 66.2% other factors (Ɛ) being (100-66.2) remains unexplained.

1.5 Benefits of f-tests

Answer Critical Questions

Answer pressing questions, like does a new process or solution lower the variability of another process or treatment?

Measure multiple models

F-tests can measure numerous model terms simultaneously, enabling them to compare different fits for linear models.

Monitor customer patterns

F-tests can evaluate customer patterns, like purchasing and return frequency, to shape customer experience improvements.

Eliminate or add variables

When creating statistical models, f-tests can determine if older variables should be eliminated or new variables added.

Acknowledgement

Nill

Funding

No funding was received to carry out this study.

References

1. S. Onchiri, M.M. Callen, F Test of Overall Significance in Regression Analysis Simplified, Journal of practice of cardiovascular sciences, (2020).
2. N.J. Salkind, Encyclopedia of research design, SAGE Publications, Inc, (2010).
3. L. Sullivan, Essentials of Biostatistics Workbook (2nd ed.), London: Jones and Bartlett Learning, (2003).
4. R. Odek, L. Oyugi, Influence of governance indicators on disclosures in performance audit reports in national government constituencies development funds, International Journal of Economics, Commerce and Management, 9(11) (2021) 172-235.