stats

MDPI

Article

Neutrosophic F-Test for Two Counts of Data from the Poisson Distribution with Application in Climatology

Muhammad Aslam

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21551, Saudi Arabia; aslam_ravian@hotmail.com

Abstract: This paper addresses the modification of the F-test for count data following the Poisson distribution. The F-test when the count data are expressed in intervals is considered in this paper. The proposed F-test is evaluated using real data from climatology. The comparative study showed the efficiency of the F-test for count data under neutrosophic statistics over the F-test for count data under classical statistics.

Keywords: Poisson distribution; weather data; efficiency; classical statistics; neutrosophic statistics

1. Introduction

The F-test for two counts obtained from the Poisson distribution has been applied for the testing of the hypothesis that two counts differ significantly vs. the alternative hypothesis that there are two counts that do not differ significantly. The F-test for two counts is used to test whether the two counts exposed to different conditions (respectively) are significantly different (see Kanji [1]). The F-test works when the sample is obtained under the same conditions. This type of test is applied for testing the occurrence of counts in a fixed time and when the time interval is different. The details of the F-test when the counts are obtained from the Poisson distribution can be found in Kanji [1]. A powerful test for comparing the means of the Poisson distribution was presented by Krishnamoorthy and Thomson [2]. The application of the count data in the field of education was studied by Hilbe [3]. The application and goodness-of-fit test were studied by Puig and Weiß [4]. Various applications of this kind of test can be seen in [5–10].

Weather and wind forecasting and estimation are conducted using statistical models. The statistical tests are applied for testing the significance of the models used in the estimation and forecasting of wind speed. For the applications of statistical techniques and methods, the reader may refer to [11–20].

The F-test for two counts under classical statistics is usually applied when counts from the Poisson distribution are sure, certain, and determinate. Viertl [21] stated that “statistical data are frequently not precise numbers but more or less non-precise, also called fuzzy. Measurements of continuous variables are always fuzzy to a certain degree”. When the count data are recorded in an interval such as the minimum count and the maximum count, the F-test for two counts cannot be applied. In such a situation, the statistical test designed under fuzzy logic can be applied. The application of statistical tests under fuzzy logic can be seen in [22–29].

The idea of neutrosophic logic was introduced by Smarandache [30]. The efficiency of neutrosophic logic can be found in Smarandache [31]. Some applications of the neutrosophic logic were described in [32–36]. Neutrosophic statistics with applications were discussed by Smarandache [37]. The statistical methods to deal with neutrosophic numbers were discussed by Chen et al. [38,39]. Aslam [40] introduced the F-test to perform an analysis of the variance test. Various neutrosophic tests with applications can be seen in [41,42].

Citation: Aslam, M. Neutrosophic F-Test for Two Counts of Data from the Poisson Distribution with Application in Climatology. *Stats* **2022**, *5*, 773–783. https://doi.org/10.3390/stats5030045

Academic Editor: Wei Zhu

Received: 5 July 2022
Accepted: 10 August 2022
Published: 12 August 2022

Generally, the F-test under classical statistics is applied for testing the equality of variance of two population variances. The existing F-test for two counts (Poisson distribution) is applied when counts are determined or all parameters involved are certain. The F-test under classical statistics cannot be utilized when counts are recorded in an interval. According to the literature, to the best of the author's knowledge, no F-test for two counts (Poisson distribution) can be found when counts are recorded in the interval. The main focus of the paper was to develop the F-test under neutrosophic statistics when two counts are modeled from a Poisson distribution. Data from the weather department are applied for the illustration of the proposed F-test for counts. It is expected that the proposed F-test for counts is more efficient than the existing F-test for counts under classical statistics.

2. Methods

Kanji [1] reported the F-test when two counts follow a Poisson distribution. Kanji [1] presented the statistical test when counts are obtained in a fixed time and obtained over different time intervals. Mathematical proofs of the F-test for two counts when the number of events occurring in a fixed time and different periods of time are shown in the Appendix A (Theorems A1 and A2). Let us assume that $N_{1N} = N_{1L} + N_{1U}I_{1N}; I_{1N}\epsilon[I_{1L}, I_{1U}]$ and $N_{2N} = N_{2L} + N_{2U}I_{2N}; I_{2N}\epsilon[I_{1L}, I_{2U}]$ are two neutrosophic counts consisting of determinate and indeterminate parts. For the measure of $I_{1N}\epsilon[I_{1L}, I_{1U}]$ and $I_{2N}\epsilon[I_{1L}, I_{2U}]$ indeterminacy in counts, see [43–49].

Theorem 1. *The statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ in neutrosophic form when counts are obtained in the same period of time $(t_1 = t_2)$ is given as*

$$F_{N1} = F_{L1} + F_{U1}I_{F_{N1}}; I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right].$$

Proof. The basic operation on neutrosophic numbers can be performed as described below.

Let $F_{N1} = F_{L1} + F_{U1}I_{F_N}$ *and* $F_{N2} = F_{L2} + F_{U2}I_{F_N}$ for $I_{F_N}\epsilon\left[I_{F_L}, I_{F_U}\right]$ be neutrosophic numbers obtained from the neutrosophic F-distribution (see [37]). According to [38], the following operation can be performed on the neutrosophic numbers:

$$F_{N1} + F_{N2} = (F_{L1} + F_{L2}) + (F_{U1} + F_{U2})I_{F_N} = \left[F_{L1} + F_{L2} + F_{U1}I_{F_L} + F_{U2}I_{F_L}, F_{L1} + F_{L2} + F_{U1}I_{F_U} + F_{U2}I_{F_U}\right],$$

or

$$F_{N1} + F_{N2} = \left[(F_{L1} + F_{L2}) + (F_{U1} + F_{U2})I_{F_L}, (F_{L1} + F_{L2}) + (F_{U1} + F_{U2})I_{F_U}\right].$$

Suppose that F_{L1} is the lower value, $F_{U1}I_{F_{N1}}$ is the upper value of the neutrosophic form, and $I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right]$ is the measure of indeterminacy.

According to [1,37,38] and Theorems A1 and A2, we have

$$F_{L1} = \frac{N_{1L}}{N_{2L} + 1} \text{ and } F_{U1} = \frac{N_{1U}}{N_{2U} + 1},$$

i.e., $F_{N1}\epsilon\left[\frac{N_{1L}}{N_{2L}+1}, \frac{N_{1U}}{N_{2U}+1}\right]$.

Note that N_{1L} and N_{1U} are the lower and upper counts of the first population, respectively, and N_{2L} and N_{2U} are the lower and upper counts of the second population, respectively.

Using the basic operation, the neutrosophic form of $F_{N1}\epsilon[F_{L1}, F_{U1}]$ can be written as

$$F_{N1} = \frac{N_{1L}}{N_{2L} + 1} + \frac{N_{1U}}{N_{2U} + 1}I_{F_{N1}}; I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right],$$

or

$$F_{N1} = F_{L1} + F_{U1}I_{F_{N1}}; I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right],$$

where $F_{N1}\epsilon[F_{L1}, F_{U1}]$ presents the neutrosophic F-distribution $(2(N_{2N} + 1), 2N_{1N})$ degree of freedom. □

Theorem 2. To test the null hypothesis that $H_{0N}: \mu_{1N} = \mu_{2N}$, the F-test, say $F_{N1}\epsilon[F_{L1}, F_{U1}]$, is given as $F_{N1} = \frac{N_{1N}}{N_{2N}+1}; F_{N1}\epsilon[F_{L1}, F_{U1}]$.

Proof. Let $N_{1N}\epsilon[N_{1L}, N_{1U}]$ and $N_{2N}\epsilon[N_{2L}, N_{2U}]$ be count in intervals for the first and the second populations, respectively. Let $N_{1N} = N_{1L} + N_{1U}I_{1N}; I_{1N}\epsilon[I_{1L}, I_{1U}]$ and $N_{2N} = N_{2L} + N_{2U}I_{2N}; I_{2N}\epsilon[I_{1L}, I_{2U}]$ be the neutrosophic forms of these counts, where $I_{1N}\epsilon[I_{1L}, I_{1U}]$ and $I_{2N}\epsilon[I_{1L}, I_{2U}]$ are the measures of indeterminacy for counts in population 1 and population 2, respectively. By following Theorems A1 and A2, we can proceed as described below.

From [1], the F-test for classical statistics is given by

$$F_1 = \frac{N_1}{N_2 + 1}.$$

The F-test, say $F_{N1}\epsilon[F_{L1}, F_{U1}]$, under neutrosophic statistics is given by

$$F_{N1} = \frac{[N_{1L}, N_{1U}]}{[N_{2L}, N_{2U}] + 1}; I_{1N}\epsilon[I_{1L}, I_{1U}], I_{2N}\epsilon[I_{1L}, I_{2U}].$$

The F-test under indeterminacy is expressed as

$$F_{N1} = \frac{N_{1L} + N_{1U}I_{1N}}{(N_{2L} + N_{2U}I_{2N}) + 1}; I_{1N}\epsilon[I_{1L}, I_{1U}], I_{2N}\epsilon[I_{1L}, I_{2U}].$$

□

Suppose that the first population mean is μ_{1N} and the second population mean is μ_{2N}. To test the null hypothesis that $H_{0N}: \mu_{1N} = \mu_{2N}$, the F-test, say $F_{N1}\epsilon[F_{L1}, F_{U1}]$, based on neutrosophic counts $N_{1N}\epsilon[N_{1L}, N_{1U}]$ and $N_{2N}\epsilon[N_{2L}, N_{2U}]$ is defined as

$$F_{N1} = \frac{N_{1N}}{N_{2N} + 1}; F_{N1}\epsilon[F_{L1}, F_{U1}]. \quad (1)$$

The statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ is modeled by a neutrosophic F-distribution with $(2(N_{2N} + 1), 2N_{1N})$ degrees of freedom (see Aslam [50]). The statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ in neutrosophic form can be expressed as

$$F_{N1} = F_{L1} + F_{U1}I_{F_{N1}}; I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right]. \quad (2)$$

The proposed statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ is a generalization of the existing F-test for two counts of data. The statistic F_{L1} presents the existing F-test for two counts of data. Note that $F_{U1}I_{F_{N1}}; I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right]$ presents the indeterminate part, and $I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right]$ is a measure of uncertainty associated with $F_{N1}\epsilon[F_{L1}, F_{U1}]$. The proposed statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ becomes the existing statistic when $I_{F_{L1}} = 0$.

Theorem 3. *The proposed statistic $F_{N2}\epsilon[F_{L2}, F_{U2}]$ in neutrosophic form over the different periods of time t_1 and t_2 is given as*

$$F_{N2} = F_{L2} + F_{U2}I_{F_{N2}}; I_{F_{N2}}\epsilon\left[I_{F_{L2}}, I_{F_{U2}}\right].$$

Proof. Let t_1 and t_2 be different times with counting rates N_{1N}/t_1 and N_{2N}/t_2, respectively. Suppose that F_{L1} is the lower value, $F_{U1}I_{F_{N1}}$ is the upper value of the neutrosophic form, and $I_{F_{N1}}\epsilon\left[I_{F_{L1}}, I_{F_{U1}}\right]$ is the measure of indeterminacy.

According to [1,37] and Theorems 1 and 2, we have

$F_{L1} = \frac{\frac{1}{t_1}(N_{2L}+0.5)}{\frac{1}{t_2}(N_{1L}+0.5)}$ and $F_{U1} = \frac{\frac{1}{t_1}(N_{2U}+0.5)}{\frac{1}{t_2}(N_{1U}+0.5)}$, i.e., $F_{N1}\epsilon\left[\frac{\frac{1}{t_1}(N_{2L}+0.5)}{\frac{1}{t_2}(N_{1L}+0.5)},\frac{\frac{1}{t_1}(N_{2U}+0.5)}{\frac{1}{t_2}(N_{1U}+0.5)}\right]$.

Note that N_{1L} and N_{1U} are the lower and upper counts of the first population, respectively, and N_{2L} and N_{2U} are the lower and upper counts of the second population, respectively.

Using the basic operation, the neutrosophic form of $F_{N2}\epsilon[F_{L2},F_{U2}]$ can be written as

$$F_{N2}=\frac{\frac{1}{t_1}(N_{2L}+0.5)}{\frac{1}{t_2}(N_{1L}+0.5)}+\frac{\frac{1}{t_1}(N_{2U}+0.5)}{\frac{1}{t_2}(N_{1U}+0.5)}I_{F_{N2}};I_{F_{N2}}\epsilon\left[I_{F_{L2}},I_{F_{U2}}\right],$$

or

$$F_{N2}=F_{L2}+F_{U2}I_{F_{N2}};I_{F_{N2}}\epsilon\left[I_{F_{L2}},I_{F_{U2}}\right],$$

where $F_{N2}\epsilon[F_{L2},F_{U2}]$ presents the neutrosophic F-distribution with $(2(N_{2N}+1),2N_{1N})$ degrees of freedom. □

Theorem 4. *The statistic $F_{N2}\epsilon[F_{L2},F_{U2}]$ over different periods of time t_1 and t_2 and counting rates N_{1N}/t_1 and N_{2N}/t_2 is given by*

$$F_{N2}=\frac{\frac{1}{t_1}(N_{2N}+0.5)}{\frac{1}{t_2}(N_{1N}+0.5)};F_{N2}\epsilon[F_{L2},F_{U2}].$$

Proof. According to [1] and Theorems A1 and A2, the F-test for classical statistics is given by

$$F_2=\frac{\frac{1}{t_1}(N_2+0.5)}{\frac{1}{t_2}(N_1+0.5)}.$$

The F-test, say $F_{N2}\epsilon[F_{L2},F_{U2}]$, under neutrosophic statistics is given by

$$F_{N1}=\frac{\frac{1}{t_1}\{[N_{2L},N_{2U}]+0.5\}}{\frac{1}{t_2}\{[N_{1L},N_{1U}]+0.5\}+1};I_{1N}\epsilon[I_{1L},I_{1U}],I_{2N}\epsilon[I_{1L},I_{2U}].$$

The F-test under indeterminacy is expressed as

$$F_{N1}=\frac{\frac{1}{t_1}\{N_{2L}+N_{2U}I_{2N}+0.5\}}{\frac{1}{t_2}\{N_{1L}+N_{1U}I_{1N}+0.5\}+1};I_{1N}\epsilon[I_{1L},I_{1U}],I_{2N}\epsilon[I_{1L},I_{2U}],$$

when two counts are observed over diverse periods of time t_1 and t_2, with the counting rates N_{1N}/t_1 and N_{2N}/t_2. The proposed statistic $F_{N2}\epsilon[F_{L2},F_{U2}]$ is defined as

$$F_{N2}=\frac{\frac{1}{t_1}(N_{2N}+0.5)}{\frac{1}{t_2}(N_{1N}+0.5)};F_{N2}\epsilon[F_{L2},F_{U2}]. \tag{3}$$

The proposed statistic $F_{N2}\epsilon[F_{L2},F_{U2}]$ in neutrosophic form can be expressed as

$$F_{N2}=F_{L2}+F_{U2}I_{F_{N2}};I_{F_{N2}}\epsilon\left[I_{F_{L2}},I_{F_{U2}}\right]. \tag{4}$$

□

The existing F-test is a special case of the proposed statistic $F_{N2}\epsilon[F_{L2},F_{U2}]$. The statistic F_{L2} presents the existing F-test for two counts of data. Note that $F_{U2}I_{F_{N2}};I_{F_{N2}}\epsilon\left[I_{F_{L2}},I_{F_{U2}}\right]$ presents the indeterminate part and $I_{F_{N2}}\epsilon\left[I_{F_{L2}},I_{F_{U2}}\right]$ is a measure of the uncertainty associated with $F_{N2}\epsilon[F_{L2},F_{U2}]$. The proposed statistic $F_{N2}\epsilon[F_{L2},F_{U2}]$ becomes the existing statistic when $I_{F_{L2}}=0$.

3. Application

The US daily and monthly weather recorded data which broke records are used to show the application of the proposed F-test. The daily or monthly records are noted in intervals rather than the exact number. The record weather data follow a Poisson distribution as these are count data; the events are independent, the average rate of record broken data during the specified time can be calculated, and two records are assumed to be not noted in the same time period. The low minimum, low maximum, high minimum, and high maximum data for daily and monthly records were accessed on 7 January 2021, through https://www.ncdc.noaa.gov/cdo-web/datatools/records. The daily record and monthly record data are presented in Tables 1 and 2, respectively (see Aslam [50]). The weather data in intervals can be adequately analyzed using the proposed F-test as compared to the F-test under classical statistics. Tables 1 and 2 indicate that the record-breaking data were recorded for the same time periods (daily records/monthly records); therefore, the statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ can be suitably applied when $(t_1 = t_2)$. In Tables 1 and 2, the values of $F_{N1}\epsilon[F_{L1}, F_{U1}]$ are also shown. The proposed test for the daily records (last 30 days) was implemented according to the following steps:

Table 1. US daily records.

Period	Low Min, Low Max	High Min, High Max	$F_{N1}\epsilon[F_{L1},F_{U1}]$
Last 30 days	[115, 152]	[597, 1483]	[0.1923, 0.1024]
Last 365 days	[15679, 19150]	[37143, 33514]	[0.4221, 0.5713]

Table 2. US monthly records.

Period	Low Min, Low Max	High Min, High Max	$F_{N1}\epsilon[F_{L1},F_{U1}]$
Last 30 days	[0, 1]	[14, 25]	[0, 0.0384]
Last 365 days	[586, 985]	[2384, 2363]	[0.2457, 0.4166]

Step-1 State $H_{0N} : \mu_{1N} = \mu_{2N}$ vs. $H_{1N} : \mu_{1N} \neq \mu_{2N}$.

Step-2 Let $\alpha = 5\%$ and the corresponding critical value at $F_{[\{1196,230\},\{2968,304\}]}$ be 1 using the F-table from Kanji [1].

Step-3 For the last 30 days, the values of F_{N1} are calculated as $F_{N1} = \frac{[115,597]}{[152,1484]}$; $F_{N1}\epsilon[0.1923, 0.1024]$.

Step-4 Accept $H_{0N} : \mu_{1N} = \mu_{2N}$ as $F_{N1}\epsilon[F_{L1}, F_{U1}] < 1$.

The computed value of $F_{N1}\epsilon[F_{L1}, F_{U1}]$ falls in the acceptance region, which leads to acceptance of the null hypothesis with the conclusion that there is no significant difference between the two counts for less than 30 days. Similarly, the other results for US daily and monthly records can be interpreted.

4. Comparative Study

A statistical test based on neutrosophic statistics has an edge over the statistical test based on classical statistics. Therefore, in this section, a comparison of the proposed F-test with the existing F-test is presented. The proposed F-test consists of two parts known as the determinate and indeterminate parts, with the latter having the measure of indeterminacy. The proposed F-test is reduced to the existing F-test when record data are measured as exact values. The neutrosophic analysis for the last 30 days and the last 365 days is shown in Table 3 (see Aslam [50]).

Table 3. Neutrosophic analysis of $F_{N1}\epsilon[F_{L1}, F_{U1}]$ two datasets.

Period	Neutrosophic Form	Measure of Indeterminacy
	U.S Daily Records	
Last 30 days	$F_{N1} = 0.1923 - 0.1024 I_{N130D}$	$I_{N130D}[0,\ 0.87]$
Last 365 days	$F_{N1} = 0.4221 + 0.5713 I_{N1365D}$	$I_{N13650D}[0,\ 0.2611]$
	U.S Monthly Records	
Last 30 days	$F_{N1} = 0 + 0.0.084 I_{N130M}$	$I_{N130M}[0,\ 1]$
Last 365 days	$F_{N1} = 0.2457 + 0.4166 I_{N1365M}$	$I_{N1365M}[0,\ 0.4102]$

For example, the neutrosophic form of the statistic $F_{N1}\epsilon[F_{L1}, F_{U1}]$ for the records of the last 365 days (US daily record) is expressed as $F_{N1} = 0.4221 + 0.5713 I_{N1365D}; I_{N1365D}[0,\ 0.2611]$. This neutrosophic form has two parts. The first part 0.4421 denotes the values of statistics under classical statistics. The second value $0.5713 I_{N1365D}$ denotes the indeterminate part, and I_{N1365D} = 0.2611 is the measure of indeterminacy associated with the test. Note here that the proposed F-test is reduced to the F-test under classical statistics when I_{N1365D} = 0. From the analysis, it can be noted that the proposed F-test provides the values of the test statistic in the interval from 0.4221 to 0.5713 rather than the exact value. From this study, it is clear that the existing F-test gives the determinate value of the statistics while the proposed F-test for two counts from the Poisson distribution gives the values in the interval, in addition to information about the measure of indeterminacy.

Another advantage of the proposed test over the existing F-test is that, when α = 5%, the probability of rejecting $H_{0N} : \mu_{1N} = \mu_{2N}$ when it is true is 0.05, the probability of accepting $H_{0N} : \mu_{1N} = \mu_{2N}$ is 0.95, and the measure of uncertainty/indeterminacy associated with the test is 0.2611. This is the case of paraconsistent probability where the total probability (0.05 + 0.95 + 0.2611 = 1.2611) lies in intervals from 1 and 3 (see Smarandache [37]). For the proposed test, the total probability associated with the test can be greater than 1; however, for the existing test, the total probability associated with the test is always equal to 1. Therefore, the proposed test is more informative than the existing F-test in terms of the probability associated with the test.

5. Power of the Test

In this section, we compare the performance of the proposed test with the existing F-test for counts in terms of the power of the test by following Aslam [51]. Suppose that α and β are the type-I error (the chance of rejecting $H_{0N} : \mu_{1N} = \mu_{2N}$ when it is true) and type-II error (the probability of accepting $H_{0N} : \mu_{1N} = \mu_{2N}$ when it is false), respectively. Let $(1 - \beta)$ be the power of the test. A simulation study was performed to calculate the power of the test for various values of $I_{F_{N1}}$. The values of the power of the test of the proposed F-test for counts when α = 0.05 are shown in Table 4. The values of the power of the test of the proposed F-test for counts when α = 0.01 are shown in Table 5. The simulation process to calculate the power of the proposed test is explained as follows:

Table 4. The values of the power of the test when α = 0.05.

$\boldsymbol{I_{F_{N1}}}$	$(1-\boldsymbol{\beta})$
[0, 0]	[0.90, 0.90]
[0, 0.2]	[0.90, 0.91]
[0, 0.4]	[0.90, 0.92]
[0, 0.6]	[0.90, 0.97]
[0, 0.8]	[0.90, 0.96]
[0, 1]	[0.90, 0.97]

Table 5. The values of power of the test when α = 0.01.

$I_{F_{N1}}$	$(1-\beta)$
0	[0.86, 0.86]
0.2	[0.86, 0.89]
0.4	[0.86, 0.89]
0.6	[0.86, 0.91]
0.8	[0.86, 0.92]
1	[0.86, 0.94]

1. Generate 100 values of counting data at various values of $I_{F_{N1}}$.
2. Specify the values of α = 0.01 and α = 0.05 and select the corresponding table values using the F-table from Kanji [1].
3. Compute the test statistic $F_{N1} \epsilon [F_{L1}, F_{U1}]$ and record the number of values accepting $H_{0N} : \mu_{1N} = \mu_{2N}$ and not accepting $H_{0N} : \mu_{1N} = \mu_{2N}$.
4. The power of the test $(1 - \beta)$ can be computed from the ratio of values accepting $H_{0N} : \mu_{1N} = \mu_{2N}$ to the total number of replications.

In Tables 4 and 5 the first value when $I_{F_{N1}} = 0$ presents the power of the test under classical statistics. From Tables 4 and 5, it can be seen that, as the values of $I_{F_{N1}}$ increased, the values of the power of the test also increased. For example, from Table 4, it can be noted that the value of the power of the test was [0.90, 0.90] when $I_{F_{N1}} = 0$ and the value of the power of the test was [0.90, 0.97] when $I_{F_{N1}} = 0.6$. The curves of power of the proposed F-test for counts are shown in Figures 1 and 2. Figures 1 and 2 also show that as the values of $I_{F_{N1}}$ increased, the values of the power of the test also increased.

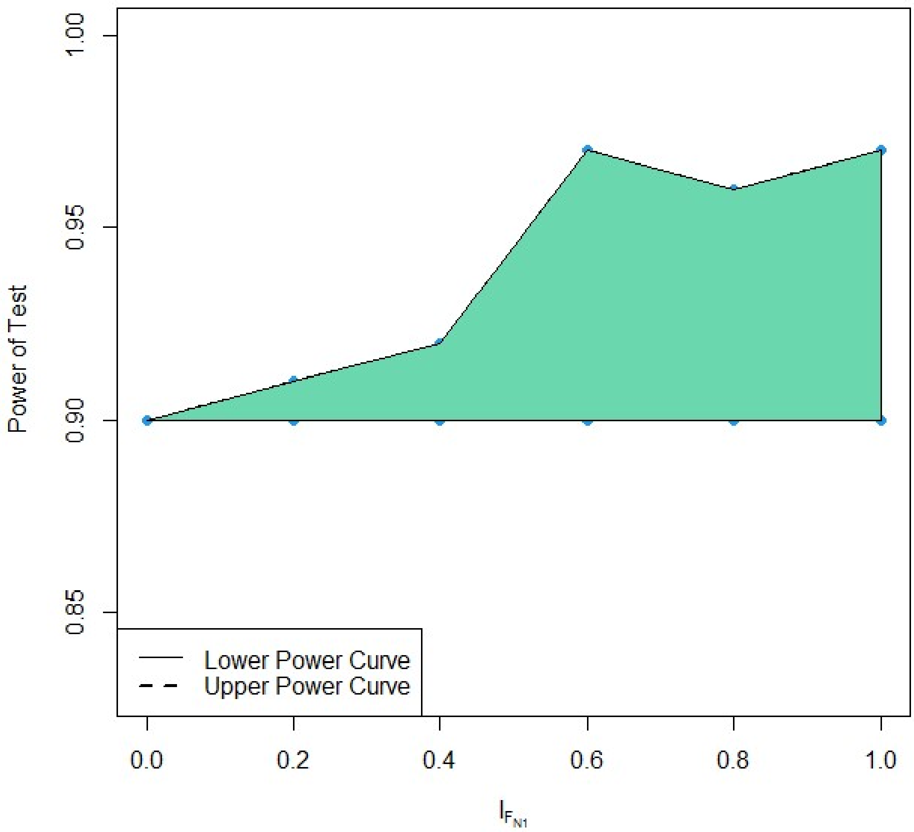

Figure 1. The curve of the power of the test when α = 0.05.

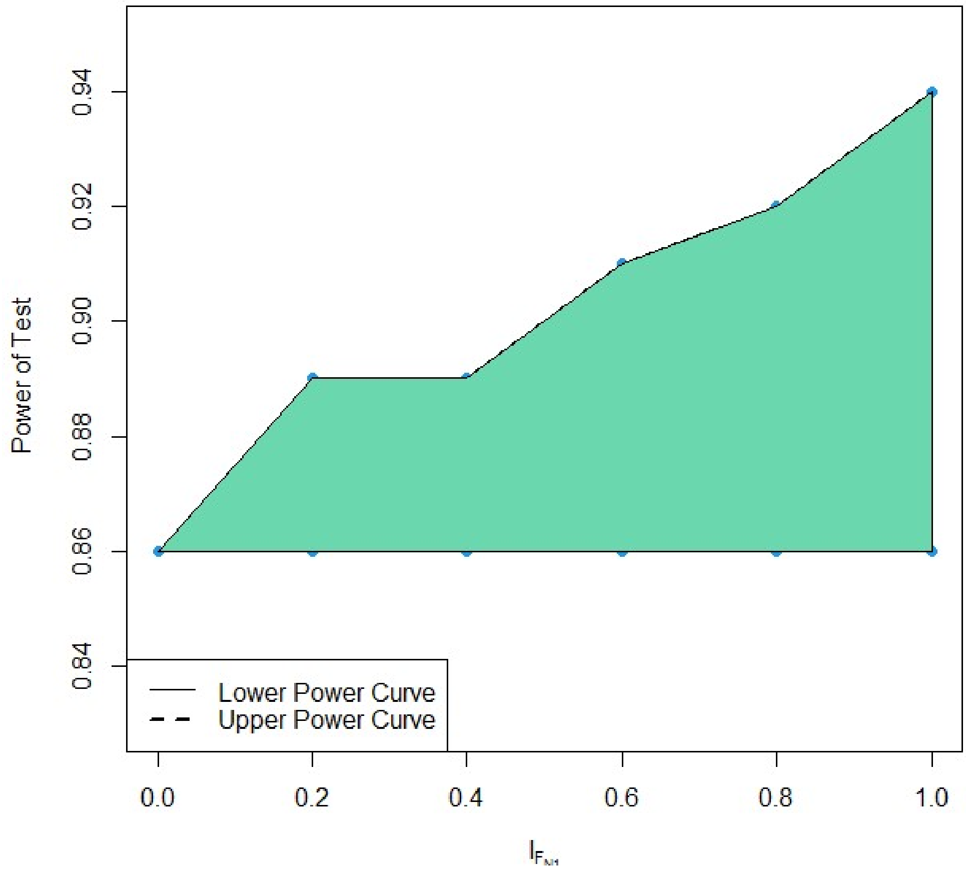

Figure 2. The curve of the power of the test when α = 0.01.

6. Concluding Remarks

The modification of the F-test for two counts of data came from the Poisson distribution presented in the paper. The mathematical proofs to the equations under neutrosophic statistics were added. The theoretical and application studies showed that the proposed F-test for two counts of data is a generalization of the existing F-test. From the analysis, it was noted that the proposed F-test performed better than the existing F-test in terms of information and probabilities associated with the test. The proposed F-test can be extended using a double sampling scheme for future research.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data is taken from https://www.ncdc.noaa.gov/cdo-web/datatools/records (accessed on 4 July 2022).

Acknowledgments: The author is deeply thankful to the editor and reviewers for their valuable suggestions to improve the quality and presentation of the paper.

Conflicts of Interest: The author declare no conflict of interest.

Appendix A

Theorem A1. *Show that* $F = \frac{t_1(n_2+0.5)}{t_2(n_1+0.5)}$.

Proof. From Cox (1953), in direct Poisson sampling, the number of events occurring in a fixed time is recorded, whereby we have

$$prob(no.\ of\ events \geq n) = \sum_{r=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^r}{r!} = prob\left(\frac{1}{2\lambda}\chi^2_{2n} < t\right),$$

and

$$prob(no.\ of\ events \geq n+1) = prob\left(\frac{1}{2\lambda}\chi^2_{2n+2} < t\right).$$

According to Cox (1953), "if we wish to make an approximation to probability (no. of events > *n)* in which the number of events is treated as a continuous variate, it is reasonable to take a quantity intermediate between above equations". A natural choice is

$$prob(no.\ of\ events \geq n+1) = prob\left(\frac{1}{2\lambda}\chi^2_{2n+1} < t\right),$$

i.e., we calculate probabilities as if $2\lambda t$ is distributed as χ^2_{2n+1}.

As mentioned in Cox (1953), "the suggestion is that, if we observe n events in a fixed time t, we may work approximately as if n is fixed and $2\lambda t$ is distributed as χ^2_{2n+1}. The remainder of the paper is concerned with the consequences using the above equation". If we sample two populations with rates of occurrence λ_1, λ_2, and if we observe n_1, n_2 events in times t_1, t_2, then, "$2\lambda t$ is distributed as χ^2_{2n+1}".

We can write that

$$\frac{t_1(n_2+0.5)\lambda_1}{t_2(n_1+0.5)\lambda_2}$$

is distributed approximately as F with $(2n_1+1, 2n_2+1)$ degrees of freedom. Thus, we may test the hypothesis that $\lambda_1 = \lambda_2$ by referring to

$$F = \frac{t_1(n_2+0.5)}{t_2(n_1+0.5)}.$$

Hence, Theorem A1 is proven.

Note: In Kanji [1], the formulas should be read as above. □

Theorem A2. *Show that* $F = \frac{n_1}{n_2+1}$.

Proof. By following Theorem 1 and Kanji [1], we can write that

$$\frac{t_1 2n_1\lambda_1}{2(n_2+1)t_2\lambda_2}$$

is distributed approximately as F with $(2(n_2+1), 2n_1)$ degrees of freedom. Thus, we may test the hypothesis that $\lambda_1 = \lambda_2$ when $(t_1 = t_2)$ by referring to

$$F = \frac{n_1}{(n_2+1)}.$$

Hence, Theorem A2 is proven. □

References

1. Kanji, G.K. *100 Statistical Tests*; Sage: Thousand Oaks, CA, USA, 2006.
2. Krishnamoorthy, K.; Thomson, J. A more powerful test for comparing two Poisson means. *J. Stat. Plan. Inference* **2004**, *119*, 23–35. [CrossRef]
3. Hilbe, J.M. The statistical analysis of count data/El análisis estadístico de los datos de recuento. *Cult. Educ.* **2017**, *29*, 409–460. [CrossRef]
4. Puig, P.; Weiß, C.H. Some goodness-of-fit tests for the Poisson distribution with applications in Biodosimetry. *Comput. Stat. Data Anal.* **2020**, *144*, 106878. [CrossRef]
5. White, G.C.; Bennetts, R.E. Analysis of Frequency Count Data Using the Negative Binomial Distribution. *Ecology* **1996**, *77*, 2549–2557. [CrossRef]
6. Coxe, S.; West, S.G.; Aiken, L.S. The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives. *J. Pers. Assess.* **2009**, *91*, 121–136. [CrossRef]
7. Salinas-Rodriguez, A.; Manrique-Espinoza, B.; Sosa-Rubi, S.G. Statistical analysis for count data: Use of healthcare services applications. *Salud Publica Mex.* **2009**, *51*, 397–406.
8. Pham, T.V.; Jimenez, C.R. An accurate paired sample test for count data. *Bioinformatics* **2012**, *28*, i596–i602. [CrossRef]
9. Hawinkel, S.; Rayner, J.C.W.; Bijnens, L.; Thas, O. Sequence count data are poorly fit by the negative binomial distribution. *PLoS ONE* **2020**, *15*, e0224909. [CrossRef]
10. Böhning, D.; Sangnawakij, P. Count outcome meta-analysis for comparing treatments by fusing mixed data sources: Comparing interventions using across report information. *Adv. Stat. Anal.* **2021**, *105*, 75–85. [CrossRef]
11. Akpinar, E.K.; Akpinar, S. A statistical analysis of wind speed data used in installation of wind energy conversion systems. *Energy Convers. Manag.* **2005**, *46*, 515–532. [CrossRef]
12. Yilmaz, V.; Çelik, H.E. A Statistical Approach to Estimate the Wind Speed Distribution: The Case of Gelibolu Region. *Doğuş Üniversitesi Derg.* **2008**, *9*, 122–132. [CrossRef]
13. Ali, S.; Lee, S.-M.; Jang, C.-M. Statistical analysis of wind characteristics using Weibull and Rayleigh distributions in Deokjeok-do Island—Incheon, South Korea. *Renew. Energy* **2018**, *123*, 652–663. [CrossRef]
14. Arias-Rosales, A.; Osorio-Gómez, G. Wind turbine selection method based on the statistical analysis of nominal specifications for estimating the cost of energy. *Appl. Energy* **2018**, *228*, 980–998. [CrossRef]
15. Akgül, F.G.; Şenoğlu, B. Comparison of wind speed distributions: A case study for Aegean coast of Turkey. *Energy Sources Part A Recover. Util. Environ. Eff.* **2019**, 1–18. [CrossRef]
16. Haq, M.A.U.; Rao, G.S.; Albassam, M.; Aslam, M. Marshall-Olkin Power Lomax distribution for modeling of wind speed data. *Energy Rep.* **2020**, *6*, 1118–1123. [CrossRef]
17. Bludszuweit, H.; Dominguez-Navarro, J.A.; Llombart, A. Statistical Analysis of Wind Power Forecast Error. *IEEE Trans. Power Syst.* **2008**, *23*, 983–991. [CrossRef]
18. Brano, V.L.; Orioli, A.; Ciulla, G.; Culotta, S. Quality of wind speed fitting distributions for the urban area of Palermo, Italy. *Renew. Energy* **2011**, *36*, 1026–1039. [CrossRef]
19. Katinas, V.; Gecevicius, G.; Marciukaitis, M. An investigation of wind power density distribution at location with low and high wind speeds using statistical model. *Appl. Energy* **2018**, *218*, 442–451. [CrossRef]
20. Zaman, B.; Lee, M.H.; Riaz, M.; Abujiya, M.R. An improved process monitoring by mixed multivariate memory control charts: An application in wind turbine field. *Comput. Ind. Eng.* **2020**, *142*, 106343. [CrossRef]
21. Viertl, R. Univariate statistical analysis with fuzzy data. *Comput. Stat. Data Anal.* **2006**, *51*, 133–147. [CrossRef]
22. Filzmoser, P.; Viertl, R. Testing hypotheses with fuzzy data: The fuzzy *p*-value. *Metrika* **2004**, *59*, 21–29. [CrossRef]
23. Tsai, C.-C.; Chen, C.-C. Tests of quality characteristics of two populations using paired fuzzy sample differences. *Int. J. Adv. Manuf. Technol.* **2006**, *27*, 574–579. [CrossRef]
24. Taheri, S.M.; Arefi, M. Testing fuzzy hypotheses based on fuzzy test statistic. *Soft Comput.* **2008**, *13*, 617–625. [CrossRef]
25. Jamkhaneh, E.B.; Ghara, A.N. *Proceedings of the 2010 International Conference on Intelligent Computing and Cognitive Informatics, Kuala Lumpur, Malaysia, 22–23 June 2010*; IEEE: Piscataway Township, NY, USA; pp. 86–89.
26. Chachi, J.; Taheri, S.M.; Viertl, R. Testing Statistical Hypotheses Based on Fuzzy Confidence Intervals. *Austrian J. Stat.* **2012**, *41*, 267–286. [CrossRef]
27. Kalpanapriya, D.; Pandian, P. Statistical hypotheses testing with imprecise data. *Appl. Math. Sci.* **2012**, *6*, 5285–5292.
28. Montenegro, M.; Casals, M.A.R.; Lubiano, M.A.A.; Gil, M.A.A. Two-sample hypothesis tests of means of a fuzzy random variable. *Inf. Sci.* **2001**, *133*, 89–100. [CrossRef]
29. Park, S.; Lee, S.-J.; Jun, S. Patent Big Data Analysis using Fuzzy Learning. *Int. J. Fuzzy Syst.* **2017**, *19*, 1158–1167. [CrossRef]
30. Smarandache, F. *Neutrosophy. Neutrosophic Probability, Set, and Logic, ProQuest Information & Learning*; Infinite Study: Ann Arbor, MI, USA, 1998; Volume 105, pp. 118–123.
31. Smarandache, F. *Introduction to Neutrosophic Measure, Neutrosophic Integral, And Neutrosophic Probability*; Infinite Study: Ann Arbor, MI, USA, 2011. [CrossRef]
32. Broumi, S.; Smarandache, F. Correlation Coefficient of Interval Neutrosophic Set. *Appl. Mech. Mater.* **2013**, *436*, 511–517. [CrossRef]
33. Guo, Y.; Sengur, A. NCM: Neutrosophic c-means clustering algorithm. *Pattern Recognit.* **2015**, *48*, 2710–2724. [CrossRef]

34. Broumi, S.; Bakali, A.; Talea, M.; Smarandache, F. *Bipolar Neutrosophic Minimum Spanning Tree*; Infinite Study: Ann Arbor, MI, USA, 2018.
35. Abdel-Baset, M.; Chang, V.; Gamal, A. RETRACTED: Evaluation of the green supply chain management practices: A novel neutrosophic approach. *Comput. Ind.* **2019**, *108*, 210–220. [CrossRef]
36. Abdel-Basset, M.; Mohamed, M.; Elhoseny, M.; Son, L.H.; Chiclana, F.; Zaied, A.E.-N.H. Cosine similarity measures of bipolar neutrosophic set for diagnosis of bipolar disorder diseases. *Artif. Intell. Med.* **2019**, *101*, 101735. [CrossRef] [PubMed]
37. Smarandache, F. *Introduction to Neutrosophic Statistics*; Infinite Study: Ann Arbor, MI, USA, 2014. [CrossRef]
38. Chen, J.; Ye, J.; Du, S. Scale Effect and Anisotropy Analyzed for Neutrosophic Numbers of Rock Joint Roughness Coefficient Based on Neutrosophic Statistics. *Symmetry* **2017**, *9*, 208. [CrossRef]
39. Chen, J.; Ye, J.; Du, S.; Yong, R. Expressions of Rock Joint Roughness Coefficient Using Neutrosophic Interval Statistical Numbers. *Symmetry* **2017**, *9*, 123. [CrossRef]
40. Aslam, M. Neutrosophic analysis of variance: Application to university students. *Complex Intell. Syst.* **2019**, *5*, 403–407. [CrossRef]
41. Aslam, M.; Albassam, M. Application of Neutrosophic Logic to Evaluate Correlation between Prostate Cancer Mortality and Dietary Fat Assumption. *Symmetry* **2019**, *11*, 330. [CrossRef]
42. Aslam, M. A new method to analyze rock joint roughness coefficient based on neutrosophic statistics. *Measurement* **2019**, *146*, 65–71. [CrossRef]
43. Ye, J. Multiple-attribute group decision-making method under a neutrosophic number environment. *J. Intell. Syst.* **2016**, *25*, 377–386. [CrossRef]
44. Liu, P.; Liu, X. The neutrosophic number generalized weighted power averaging operator and its application in multiple attribute group decision making. *Int. J. Mach. Learn. Cybern.* **2018**, *9*, 347–358. [CrossRef]
45. Zheng, E.; Teng, F.; Liu, P. Multiple attribute group decision-making method based on neutrosophic number generalized hybrid weighted averaging operator. *Neural Comput. Appl.* **2017**, *28*, 2063–2074. [CrossRef]
46. Pramanik, S.; Roy, R.; Roy, T.K. Teacher Selection Strategy Based on Bidirectional Projection Measure in Neutrosophic Number Environment. In *Neutrosophic Operational Research*; Smarandache, F., Abdel-Basset, M., El-Henawy, I., Eds.; Pons Publishing House: Bruxelles, Belgium, 2017; Volume 2, ISBN 978-1-59973-537-5.
47. Maiti, I.; Mandal, T.; Pramanik, S. Neutrosophic goal programming strategy for multi-level multi-objective linear programming problem. *J. Ambient Intell. Humaniz. Comput.* **2019**, *11*, 3175–3186. [CrossRef]
48. Mondal, K.; Pramanik, S.; Giri, B.C.; Smarandache, F. NN-Harmonic Mean Aggregation Operators-Based MCGDM Strategy in a Neutrosophic Number Environment. *Axioms* **2018**, *7*, 12. [CrossRef]
49. Pramanik, S.; Banerjee, D. Neutrosophic number goal programming for multi-objective linear programming problem in neutrosophic number environment. *MOJ Curr. Res. Rev.* **2018**, *1*, 135–141. [CrossRef]
50. Aslam, M. RETRACTED ARTICLE: Neutrosophic statistical test for counts in climatology. *Sci. Rep.* **2021**, *11*, 17806. [CrossRef]
51. Aslam, M. Enhanced statistical tests under indeterminacy with application to earth speed data. *Earth Sci. Inform.* **2021**, *14*, 1261–1267. [CrossRef]