Information **2011**, *2*, 528-545; doi:10.3390/info2030528

OPEN ACCESS

information

ISSN 2078-2489

www.mdpi.com/journal/information

Communication

Pearson-Fisher Chi-Square Statistic Revisited

Sorana D. Bolboacă [1], Lorentz Jäntschi [2,*], Adriana F. Sestraş [2,3], Radu E. Sestraş [2] and Doru C. Pamfil [2]

[1] "Iuliu Haţieganu" University of Medicine and Pharmacy Cluj-Napoca, 6 Louis Pasteur, Cluj-Napoca 400349, Romania; E-Mail: sbolboaca@umfcluj.ro

[2] University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, 3-5 Mănăştur, Cluj-Napoca 400372, Romania; E-Mails: asestras@yahoo.com (A.F.S.); rsestras@yahoo.co.uk (R.E.S.); dpamfil@usamvcluj.ro (D.C.P.)

[3] Fruit Research Station, 3-5 Horticultorilor, Cluj-Napoca 400454, Romania

* Author to whom correspondence should be addressed; E-Mail: lorentz.jantschi@gmail.com; Tel: +4-0264-401-775; Fax: +4-0264-401-768.

Received: 22 July 2011; in revised form: 20 August 2011 / Accepted: 8 September 2011 / Published: 15 September 2011

Abstract: The Chi-Square test (χ^2 test) is a family of tests based on a series of assumptions and is frequently used in the statistical analysis of experimental data. The aim of our paper was to present solutions to common problems when applying the Chi-square tests for testing goodness-of-fit, homogeneity and independence. The main characteristics of these three tests are presented along with various problems related to their application. The main problems identified in the application of the goodness-of-fit test were as follows: defining the frequency classes, calculating the X^2 statistic, and applying the χ^2 test. Several solutions were identified, presented and analyzed. Three different equations were identified as being able to determine the contribution of each factor on three hypothesizes (minimization of variance, minimization of square coefficient of variation and minimization of X^2 statistic) in the application of the Chi-square test of homogeneity. The best solution was directly related to the distribution of the experimental error. The Fisher exact test proved to be the "golden test" in analyzing the independence while the Yates and Mantel-Haenszel corrections could be applied as alternative tests.

Keywords: Chi-square statistics; Fisher exact test; Chi-square distribution; 2 × 2 contingency table

---

1. Introduction

Statistical instruments are used to extract knowledge from the observation of real world phenomena as Fisher suggested "... no progress is to be expected without constant experience in analyzing and interpreting observational data of the most diverse types" (where observational data are seen as information) [1]. Moreover, the amount of information in an estimate (obtained on a sample) is directly related with the amount of information data [2]. Fisher pointed out in [2] that scientific information latent in any set of observations could be brought out by statistical analysis whenever the experimental design is conducted in order to maximize the information obtained. The analysis of information related to associations among data requires specific instruments, the Chi-Square test being one of them. The χ^2 test was introduced by K. Pearson in 1900 [3]. A significant modification to the Pearson's χ^2 test was introduced by R.A. Fisher in 1922 [4] (the degree of freedom was decreased by one unit when applied to contingency tables). Another correction made by Fisher took into account the number of unknown parameters associated to the theoretical distribution, when the parameters are estimated from central moments [5].

The Chi-square test introduced by K. Pearson was subject of debate for much research. A series of papers analyzed Pearson's test [6,7] and its problems were tackled in [8,9].

It is well known that Pearson's Chi-square (χ^2) is a family of tests with the following assumptions [10,11]: (1) The data are randomly drawn from a population; (2) The sample size is sufficiently large. The application of the Chi-square test to a small sample could lead to an unacceptable rate of type II error (accepting the null hypothesis when actually false) [12-14]. There is no accepted cut-off for the sample size; the minimum sample size varies from 20 to 50; and (3) The values on cells are adequate when no more than 1/5 of the expected values are smaller than five and there are no cells with zero count [15,16]. The source of these rules seems to be W. G. Cochran and they appear to have been arbitrarily chosen [17].

Yates' correction is applied when the third assumption is not met [18]. The Fisher's Exact test is the alternative when Yates' correction is not acceptable [19].

Koehler and Larntz suggested the use of at least three categories if the number of observations is at least 10. Moreover, they suggested that the square of the number of observations be at least 10 times higher than the number of categories [20].

The Chi-square test has been applied in all research areas. Its main uses are: goodness-of-fit [21-25], association/independence [26-29], homogeneity [30-33], classification [34-37], *etc.*

The aim of our paper was to present solutions to common problems when applying the Chi-square for testing goodness-of-fit, homogeneity and independence.

2. Material and Methods

The most frequently used Chi-square tests were presented (Table 1) and solutions to frequent problems were provided and discussed.

The main characteristics of these tests are as follows:

- Goodness-of-fit (Pearson's Chi-Square Test [3]):
  - Is used to study similarities between groups of categorical data.

- Tests if a sample of data came from a population with a specific distribution (compares the distribution of a variable with another distribution when the expected frequencies are known) [38].
- Can be applied to any univariate distribution by calculating its cumulative distribution function (*CDF*).
- Has as alternatives the Anderson-Darling [39] and Kolmogorov-Smirnov [40] goodness-of-fit tests.

Table 1. Summary of Chi-square tests.

Type	Aim	Hypotheses	Statistics df H_0 acceptance rule
Goodness-of-fit	- One sample. - Compares the expected and observed values to determine how well the experimenter's predictions fit the data.	H_0: The observed values are equal to theoretical values (expected). (The data followed the assumed distribution). H_a: The observed values are not equal to theoretical values (expected). (The data did not follow the assumed distribution).	$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{i,j}-E_{i,j})^2}{E_{i,j}}$ df = (k – 1) $\chi^2 \le \chi^2_{1-\alpha}$
Homogeneity	- Two different populations (or sub-groups). - Applied to one categorical variable.	H_0: Investigated populations are homogenous. H_a: Investigated populations are not homogenous.	$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{i,j}-E_{i,j})^2}{E_{i,j}}$ df = (r – 1)(c – 1) $\chi^2 \le \chi^2_{1-\alpha}$
Independence	- One population. - Type of variables: nominal, dichotomical, ordinal or grouped interval - Each population is at least 10 times as large as its respective sample [21]	Research hypothesis: The two variables are dependent (or related). H_0: There is no association between two variables. (The two variables are independent). H_a: There is an association between two variables.	$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{i,j}-E_{i,j})^2}{E_{i,j}}$ df = (r – 1) × (c – 1) $\chi^2 \le \chi^2_{1-\alpha}$

The agreement between observation and hypothesis is analyzed by dividing the observations in a defined number of intervals (*f*). The X^2 statistic is calculated based on the formula presented in Equation (1).

$$X^2 = \sum_{i=1}^{f}\frac{(O_i - E_i)^2}{E_i} \approx \chi^2(f - t - 1) \tag{1}$$

where X^2 = value of Chi-square statistics; χ^2 = value of the Chi-square parameter from Chi-square distribution; O_i = experimental (observed) frequency associated to the ith frequency class; E_i = expected frequency calculated from the theoretical distribution law for the ith frequency class; t = number of parameters in theoretical distribution estimated from central moments.

The probability to reject the null hypothesis is calculated based on the theoretical distribution (χ^2). The null hypothesis is accepted if the probability to be rejected ($\chi^2_{CDF}(X^2, f\text{-}t\text{-}1)$) is lower than 5%.

The Chi-square test is the most well known statistics used to test the agreement between observed and theoretical distributions, independency and homogeneity. Defining the applicability domain of the Chi-square test is a complex problem [38].

At least three problems occur when the Chi-square test is applied in order to compare observed and theoretical distributions:

- Defining the frequency classes.
- Calculating the X^2 statistic.
- Applying the χ^2 test.

- The test of homogeneity:
  - Is used to analyze if different populations are similar (homogenous or equal) in terms of some characteristics.
  - Is applied to verify the homogeneity of: data, proportions, variance (more than two variances are tested; for two variances the F test is applied), error variance, sampling variances.

The Chi-square test of homogeneity is used to determine whether frequency counts are identically distributed across different populations or across different sub-groups of the same population. An important assumption is made for the test of homogeneity in populations coming from a contingency of two or more categories (this is the link between the test of homogeneity and the test of independence): the observable under assumption of homogeneity should be observed in a quantity proportional with the product of the probabilities given by the categories (assumption of independence between categories). When the number of categories is two, the expectations are calculated using the $E_{i,j}$ formula (mean of the expected value (for Chi-square of homogeneity) or frequency counts (Chi-square test of independence) for (i,j) pair of factors) [41].

The observed contingency table is constructed; the values for the first factor/population/subgroup are in the rows and the values for the second variable/factor/population/subgroup are in the columns. The observed frequencies are counted at the intersection of rows with columns and the hypothesis of homogeneity is tested.

The value of X^2 statistic is computed using the formula presented in Equation (2).

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{i,j}-E_{i,j})^2}{E_{i,j}} \approx \chi^2((r-1)(c-1)) \tag{2}$$

where r = number of rows in contingency table; c = number of columns in contingency table; $1 \le i \le r$ = indices of observations associated to the first factor; $1 \le j \le c$ = indices of observations associated to the second factor; $O_{i,j}$ = mean of the observed value (for chi-square of homogeneity) or frequency counts (Chi-square test of independence) for (i,j) pair of factors; $E_{i,j}$ = mean of the expected value (for Chi-square of homogeneity) or frequency counts (chi-square test of independence) for (i,j) pair of factors; X^2 = the value of Chi-square statistic; χ^2 = Chi-square critical parameter (from chi-square distribution).

- The test of independence (also known as Chi-square test of association):
  - Is used to determine whether two characteristics are dependent or not.
  - Compares the frequencies of one nominal variable for different values of a second nominal variable.

- Is an alternative to the G-test of independence (also known as the Likelihood Ratio Chi-square test) [43].
- Fisher's exact test of independence [44] is preferred whenever small expected values are presented.

The chi-square test of independence is applied in order to compare frequencies of nominal or ordinal data for a single population/sample (two variables at the same time).

The Chi-square for independence also faced some difficulties when applied on experimental data [39]. Fisher exact test [43] was proposed by Fisher as an alternative to the Chi-square test [44]; the Fisher exact test is based on the calculation of marginal probabilities (which unfortunately has an exact calculation formula only for 2 × 2 contingencies).

Glass and Hopkins [45] consider that the Chi-square test of association is equivalent to the Chi-square test of independence and to the Chi-square test of homogeneity.

3. Results and Discussion

3.1. Chi-Square Test of Goodness-of-Fit

The first problem of the Chi-square test of goodness-of-fit is how to establish the number of frequency classes. At least two approaches could be applied:

- The number of frequency classes (discreet number) is computed from Hartley's entropy [46] of observed versus expected data: $log_2(2n)$, where n = number of observations. The EasyFit software (MathWave Technologies. http://www.mathwave.com) uses this approach.
- The number of frequency classes is obtained based on the histogram of observed values as estimator of density [47]. The optimal criterion is applied in order to obtain the width of the classes when the histogram is used. For example, Dataplot (National Institute for Standards and Technology. http://www.itl.nist.gov/div898/software/dataplot.html) automatically generates frequency classes using this method: the width of the frequency class is $0.3 \cdot s$ (where s = standard deviation of the sample). The lower and upper bounders are given by $m \pm 6 \cdot s$ (where m = arithmetic mean, s = standard deviation) and the marginal classes of frequencies are omitted.

One rule-of-thumb suggests dividing the sample in a number of frequency classes equal to $2 \cdot n^{2/5}$ (where n = sample size) [48].

The second problem refers to the width of the frequency classes. Two approaches could be applied here:

- Data could be grouped in probability frequency classes (theoretical or observed) with equal width. This approach is frequently used when the observed data are grouped.
- Data could be grouped in intervals with equal width.

The third problem is the number of observations in each frequency class. Every class must contain at least five observations; otherwise the frequencies from two proximity classes are cumulated.

3.2. Chi-Square Test of Homogeneity

The investigation of homogeneity of the values associated to a class (row or column in the

contingency table) could be carried out by decomposing the X^2 expression (see Equation (3)). A hierarchy of irregularities on the contingency table could also be obtained by decomposing the X^2 expression.

$$X^2{}_c = \sum_{i=1}^{r} \frac{(O_{i,c} - E_{i,c})^2}{E_{i,c}} \approx \chi^2(r-1)$$
$$X^2{}_r = \sum_{j=1}^{c} \frac{(O_{r,j} - E_{r,j})^2}{E_{r,j}} \approx \chi^2(c-1) \tag{3}$$

One assumption is that the $O_{i,j}$ observations are the result of multiplying two factors; repeated observations approximate better the effect of multiplication. Thus, the formula of expected frequencies ($E_{i,j}$ [43]) is the consequence of factors' multiplication and it is presented in Equation (4).

$$E_{i,j} = \left(\sum_{k=1}^{r} O_{i,k}\right) \cdot \left(\sum_{k=1}^{c} O_{k,j}\right) \Big/ \left(\sum_{i=1}^{r}\sum_{j=1}^{c} O_{i,j}\right) \tag{4}$$

Three mathematical assumptions could be formulated in terms of square error ($(O_{i,j} - E_{i,j})^2$) of observation:

- The measurement is affected by chance errors, absolute values (S^2, Equation (5));
- The measurement is affected by chance errors, relative values (CV^2, Equation (6));
- The measurement is affected by chance errors on a scale with values (X^2, Equation (7)) between absolute and relative errors.

The first hypothesis (chance errors, absolute values) leads mathematically to the minimization of the variance (S^2) obtained between model and observation.

$$S^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} (O_{i,j} - a_i b_j)^2 = \min \tag{5}$$

where a_i, $1 \le i \le r$ = contribution of first factor to the expected value $E_{i,j}$; b_i, $1 \le j \le c$ = contribution of second factor to the expected value $E_{i,j}$; $E_{i,j}=a_i \cdot b_j$.

The second hypothesis (chance errors, relative values) leads to the minimization of the squared coefficient of variation (CV^2) (see Equation (6)).

$$CV^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{i,j} - a_i b_j)^2}{(a_i b_j)^2} = \min \tag{6}$$

One possible solution for the third hypothesis is the minimization of the X^2 statistic (see Equation (7)).

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c} \frac{(O_{i,j} - a_i b_j)^2}{a_i b_j} = \min \tag{7}$$

The contribution of each factor ($A = (a_i)_{1 \le i \le r}$, and $B = (b_j)_{1 \le j \le c}$) could be determined through the minimization of values given by Equations (5)–(7). The following formula (Equation (8)) was applied in order to minimize the values form Equations (5)–(7).

$$\left(\frac{\partial\cdot(a_i,b_j)}{\partial a_i}=0\right)_{1\le i\le r};\quad \left(\frac{\partial\cdot(a_i,b_j)}{\partial b_j}=0\right)_{1\le j\le c} \tag{8}$$

where the expression of derivate (a_i, b_j) is the expression of $S^2/CV^2/X^2$ given in Equations (5)–(7).

The calculations revealed the followings:

- The relation in Equation (5) is verified by the values of $(a_i)_{1\le i\le r}$ and $(b_i)_{1\le j\le c}$ from Equation (9).
- The relation in Equation (6) is verified by the values of $(a_i)_{1\le i\le r}$ and $(b_i)_{1\le j\le c}$; from Equation (10).
- The relation in Equation (7) is verified by the values of $(a_i)_{1\le i\le r}$ and $(b_i)_{1\le j\le c}$; from Equation (11).

$$a_i=\left(\sum_{j=1}^{c}b_jO_{i,j}\right)\Bigg/\left(\sum_{j=1}^{c}b_j^{\,2}\right),\ i=1..r;\ b_j=\left(\sum_{i=1}^{r}a_iO_{i,j}\right)\Bigg/\left(\sum_{i=1}^{r}a_i^{\,2}\right),\ j=1..c \tag{9}$$

$$a_i=\left(\sum_{j=1}^{c}\frac{O_{i,j}^{\,2}}{b_j^{\,2}}\right)\Bigg/\left(\sum_{j=1}^{c}\frac{O_{i,j}}{b_j}\right),\ i=1..r;\ b_j=\left(\sum_{i=1}^{r}\frac{O_{i,j}^{\,2}}{a_i^{\,2}}\right)\Bigg/\left(\sum_{i=1}^{r}\frac{O_{i,j}}{a_i}\right),\ j=1..c \tag{10}$$

$$a_i^{\,2}=\left(\sum_{j=1}^{c}\frac{O_{i,j}^{\,2}}{b_j}\right)\Bigg/\left(\sum_{j=1}^{c}b_j\right),\ i=1..r;\ b_j^{\,2}=\left(\sum_{i=1}^{r}\frac{O_{i,j}^{\,2}}{a_i}\right)\Bigg/\left(\sum_{i=1}^{r}a_i\right),\ j=1..c \tag{11}$$

The relations presented in Equations (9)–(11) admit an infinity of solutions and the family of solutions are close to the family of solutions given by Equation (4). Equation (4) was rewritten as presented in Equation (12).

$$a_i\cdot b_j=\left(\sum_{k=1}^{r}O_{i,k}\right)\cdot\left(\sum_{k=1}^{c}O_{k,j}\right)\Bigg/\left(\sum_{i=1}^{r}\sum_{j=1}^{c}O_{i,j}\right) \tag{12}$$

Dealing directly with Equations (9)–(11) without using Equation (12) is ineffective. For example, for r = 2 and c = 3 substituted in Equation (9) leads to Equation (13):

$$\left(\frac{a_2}{a_1}\right)^2+\frac{(O_{1,1}^{\,2}+O_{1,2}^{\,2}+O_{1,3}^{\,2})-(O_{2,1}^{\,2}+O_{2,2}^{\,2}+O_{2,3}^{\,2})}{(O_{1,1}O_{2,1}+O_{1,2}O_{2,2}+O_{1,3}O_{2,3})}\left(\frac{a_2}{a_1}\right)-1=0 \tag{13}$$

This is solvable in (a_2/a_1). Thus, there are an infinity of solutions (for any non-null value of a_1 there is a value a_2 that verifies Equation (13)) and the degree of equation Equation (13) is given by min(r,c). The equations that are obtained by direct substitution are more complicated as the *r* and *c* values increase. For example, if r = 2, and c = 3 the substitutions in Equation (11) lead to the relation presented in Equation (14), which is an equation of fifth degree (r + c).

$$\begin{aligned}&O_{1,1}^{\,2}O_{1,2}^{\,2}(O_{1,1}^{\,2}-O_{1,2}^{\,2})\left(\frac{a_2}{a_1}\right)^5+(O_{1,1}^{\,4}O_{2,2}^{\,2}-O_{1,2}^{\,4}O_{2,1}^{\,2})\left(\frac{a_2}{a_1}\right)^4+\\&+2O_{1,1}^{\,2}O_{1,2}^{\,2}(O_{2,2}^{\,2}-O_{2,1}^{\,2})\left(\frac{a_2}{a_1}\right)^3+2O_{2,1}^{\,2}O_{2,2}^{\,2}(O_{1,2}^{\,2}-O_{1,1}^{\,2})\left(\frac{a_2}{a_1}\right)^2\\&+(O_{1,2}^{\,2}O_{2,1}^{\,4}-O_{1,1}^{\,2}O_{2,2}^{\,4})\left(\frac{a_2}{a_1}\right)+O_{2,2}^{\,2}O_{2,1}^{\,2}(O_{2,1}^{\,2}-O_{2,2}^{\,2})=0\end{aligned} \tag{14}$$

The application of successive approximations using the solution offered by Equation (12) is the indirect way to solve the relations presented in Equations (9)–(11). The relation presented in Equation (12) is used in order to obtain the first approximation (initial approximation) of the solution; in every sequence of approximations the oldest values are replaced on the right side of the relations presented in Equations (9)–(11) in order to obtain the new approximations.

The method of successive approximations rapidly converged towards the optimal solution. Thus, three iterations are necessary in order to obtain a residual value of 282.11735 for the relation presented in Equation (9). Starting with the third iteration, the value of residuals is changed at the level of the 5th decimal. For the relation presented in Equation (11) the same quality of representation of the optimal solution is obtained after the 4th iteration.

The experimental data reported by Fisher [4] were used for exemplification (Table 2). The values suggested by Equation (12) for the $(a_i b_j)_{1 \le i \le 6;\ 1 \le j \le 12}$ are showed in Table 3.

Table 2. Experimental values: response to fertilization with manure on different potato varieties.

TV	UD	KK	KP	TP	ID	GS	AJ	BQ	ND	EP	AC	DY	∑
DS	25.3	28.0	23.3	20.0	22.9	20.8	22.3	21.9	18.3	14.7	13.8	10.0	**241.3**
DC	26.0	27.0	24.4	19.0	20.6	24.4	16.8	20.9	20.3	15.6	11.0	11.8	**237.8**
DB	26.5	23.8	14.2	20.0	20.1	21.8	21.7	20.6	16.0	14.3	11.1	13.3	**223.4**
US	23.0	20.4	18.2	20.2	15.8	15.8	12.7	12.8	11.8	12.5	12.5	8.2	**183.9**
UC	18.5	17.0	20.8	18.1	17.5	14.4	19.6	13.7	13.0	12.0	12.7	8.3	**185.6**
UB	9.5	6.5	4.9	7.7	4.4	2.3	4.2	6.6	1.6	2.2	2.2	1.6	**53.7**
∑	**128.8**	**122.7**	**105.8**	**105**	**101.3**	**99.5**	**97.3**	**96.5**	**81**	**71.3**	**63.3**	**53.2**	**1125.7**

TV: Treatment *vs.* Variety; UD, KK, KP, TP, ID, GS, AJ, BQ, ND, EP, AC, DY: potato varieties (UD = Up to Date; KK= K of K; KP = Kerr's Pink; TP = Tinwald Perfection; ID = Iron Duke; GS = Great Scott; AJ = Ajax; BQ = British Queen; ND = Nithsdale; EP = Epicure; AC = Arran Comrade; DY = Duke of York); DS, DC, DB, US, UC, UB: types of treatment (D* - manure; U* - without manure; S - sulphate; C - chloride; B - basal); ∑ = sum.

Table 3. Values of $(a_i b_j)_{1 \le i \le 6;\ 1 \le j \le 12}$ calculated with Equation (12) for the response to fertilization on different potato varieties.

TV	UD	KK	KP	TP	ID	GS	AJ	BQ	ND	EP	AC	DY
DS	27.61	26.30	22.68	22.51	21.71	21.33	20.86	20.69	17.36	15.28	13.57	11.40
DC	27.21	25.92	22.35	22.18	21.40	21.02	20.55	20.39	17.11	15.06	13.37	11.24
DB	25.56	24.35	21.00	20.84	20.10	19.75	19.31	19.15	16.07	14.15	12.56	10.56
US	21.04	20.04	17.28	17.15	16.55	16.25	15.90	15.76	13.23	11.65	10.34	8.69
UC	21.24	20.23	17.44	17.31	16.70	16.41	16.04	15.91	13.35	11.76	10.44	8.77
UB	6.14	5.85	5.05	5.01	4.83	4.75	4.64	4.60	3.86	3.40	3.02	2.54

TV: Treatment *vs.* Variety; (D* - manure; U* - without manure; S - sulphate; C - chloride; B - basal). UD, KK, KP, TP, ID, GS, AJ, BQ, ND, EP, AC, DY: potato varieties (UD = Up to Date; KK= K of K; KP = Kerr's Pink; TP = Tinwald Perfection; ID = Iron Duke; GS = Great Scott; AJ = Ajax; BQ = British Queen; ND = Nithsdale; EP = Epicure; AC = Arran Comrade; DY = Duke of York); DS, DC, DB, US, UC, UB: types of treatment;

The values resulted when the iterative approach was applied to obtain the solution for Equations (9)–(11) are presented in Tables 4–6. The summary of the results obtained by all four approaches are presented in Table 7.

Table 4. Optimized value of the $(a_i b_j)_{1 \le i \le 6; 1 \le j \le 12}$ calculated with Equation (9) for the response to fertilization on different potato varieties.

TV	UD	KK	KP	TP	ID	GS	AJ	BQ	ND	EP	AC	DY
DS	27.07	26.42	22.64	21.85	21.85	21.94	20.94	20.63	17.93	15.48	13.54	11.61
DC	26.66	26.02	22.29	21.52	21.52	21.60	20.62	20.32	17.66	15.24	13.33	11.43
DB	24.91	24.32	20.83	20.11	20.11	20.19	19.27	18.99	16.50	14.25	12.46	10.69
US	20.64	20.15	17.26	16.66	16.66	16.73	15.96	15.73	13.67	11.80	10.32	8.85
UC	20.58	20.09	17.21	16.61	16.61	16.68	15.92	15.69	13.63	11.77	10.29	8.83
UB	6.29	6.14	5.26	5.08	5.08	5.10	4.86	4.79	4.17	3.60	3.14	2.70

TV: Treatment *vs.* Variety; UD, KK, KP, TP, ID, GS, AJ, BQ, ND, EP, AC, DY: potato varieties (UD = Up to Date; KK= K of K; KP = Kerr's Pink; TP = Tinwald Perfection; ID = Iron Duke; GS = Great Scott; AJ = Ajax; BQ = British Queen; ND = Nithsdale; EP = Epicure; AC = Arran Comrade; DY = Duke of York); DS, DC, DB, US, UC, UB: types of treatment; (D* - manure; U* - without manure; S - sulphate; C - chloride; B - basal).

Table 5. Optimized value of the $(a_i b_j)_{1 \le i \le 6;\ 1 \le j \le 12}$ calculated with Equation (10) for the response to fertilization on different potato varieties.

TV	UD	KK	KP	TP	ID	GS	AJ	BQ	ND	EP	AC	DY
DS	27.57	26.08	23.04	22.61	21.48	21.61	21.13	20.69	17.66	15.23	13.79	11.56
DC	27.38	25.90	22.88	22.45	21.34	21.46	20.99	20.55	17.54	15.13	13.69	11.48
DB	25.84	24.44	21.59	21.19	20.14	20.26	19.80	19.40	16.56	14.28	12.92	10.83
US	21.23	20.08	17.74	17.40	16.54	16.64	16.27	15.93	13.60	11.73	10.62	8.90
UC	21.47	20.31	17.94	17.61	16.73	16.83	16.46	16.12	13.76	11.86	10.74	9.00
UB	7.02	6.64	5.87	5.76	5.47	5.51	5.38	5.27	4.5	3.88	3.51	2.94

TV: Treatment *vs.* Variety; UD, KK, KP, TP, ID, GS, AJ, BQ, ND, EP, AC, DY: potato varieties (UD = Up to Date; KK= K of K; KP = Kerr's Pink; TP = Tinwald Perfection; ID = Iron Duke; GS = Great Scott; AJ = Ajax; BQ = British Queen; ND = Nithsdale; EP = Epicure; AC = Arran Comrade; DY = Duke of York); DS, DC, DB, US, UC, UB: types of treatment; (D* - manure; U* - without manure; S - sulphate; C - chloride; B - basal).

The analysis of the results presented in Table 7 revealed that each method defined in Equations (9)–(11) increases the value of the objective sum compared with the expression defined in the Equation (12) formula; the methods provided by Equations (9)–(11) also represent corrections of Equations (12). The relation presented in Equation (9) offers a better solution compared to Equation (12) under the hypothesis of experimental errors uniformly distributed among classes (absolute experimental error). The relation presented in Equation (10) obtained a better solution compared to Equation (12) under the hypothesis of experimental errors proportional with the magnitude of the observed phenomena (relative experimental error). The relation presented in Equation (11) obtained a better solution compared to Equation (12) when the aim is to minimize the Pearson-Fisher X^2 statistics (Pearsonian expression of type III [4], p. 337).

Table 6. Optimized value of $(a_i b_j)_{1 \le i \le 6;\ 1 \le j \le 12}$ calculated with Equation (11) for the response to fertilization on different potato varieties.

TV	UD	KK	KP	TP	ID	GS	AJ	BQ	ND	EP	AC	DY
DS	27.64	26.19	22.85	22.60	21.59	21.44	20.98	20.71	17.49	15.24	13.67	11.47
DC	27.35	25.91	22.61	22.36	21.36	21.22	20.76	20.50	17.30	15.08	13.52	11.35
DB	25.74	24.40	21.28	21.05	20.11	19.97	19.55	19.29	16.29	14.20	12.73	10.68
US	21.17	20.06	17.50	17.31	16.53	16.42	16.07	15.87	13.39	11.68	10.47	8.78
UC	21.40	20.28	17.69	17.50	16.71	16.60	16.25	16.04	13.54	11.80	10.58	8.88
UB	6.57	6.23	5.43	5.37	5.13	5.10	4.99	4.93	4.16	3.63	3.25	2.73

TV: Treatment *vs.* Variety; UD, KK, KP, TP, ID, GS, AJ, BQ, ND, EP, AC, DY: potato varieties (UD = Up to Date; KK= K of K; KP = Kerr's Pink; TP = Tinwald Perfection; ID = Iron Duke; GS = Great Scott; AJ = Ajax; BQ = British Queen; ND = Nithsdale; EP = Epicure; AC = Arran Comrade; DY = Duke of York); DS, DC, DB, US, UC, UB: types of treatment; (D* - manure; U* - without manure; S - sulphate; C - chloride; B - basal).

Table 7. Comparative value for chance experimental errors.

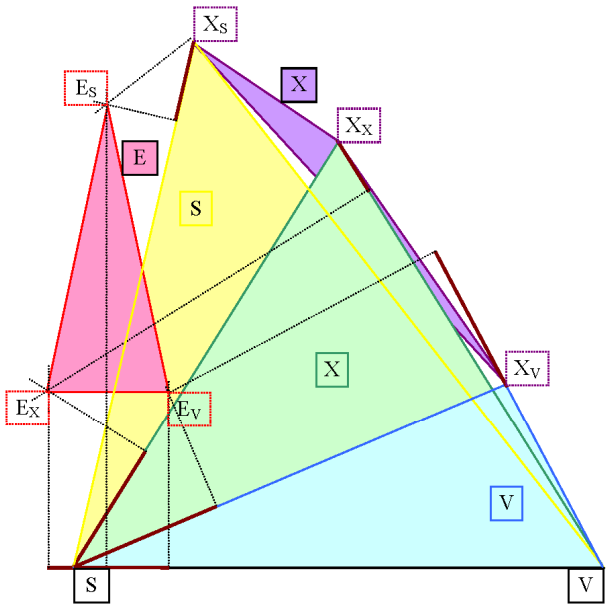
	S^2				X^2				CV^2			
Tt	**Equation (12)**	**Equation (9)**	**Equation (11)**	**Equation (10)**	**Equation (12)**	**Equation (9)**	**Equation (11)**	**Equation (10)**	**Equation (12)**	**Equation (9)**	**Equation (11)**	**Equation (10)**
DS	23.4	18.76	24.12	57.97	1.10	0.937	1.127	2.308	0.056	0.0515	0.0573	0.0971
DC	59.7	48.48	59.86	104.95	3.08	2.497	3.052	4.847	0.164	0.133	0.1611	0.2365
DB	69.8	66.77	71.47	95.21	3.78	3.596	3.796	4.803	0.221	0.2078	0.2167	0.2633
US	41.6	49.03	41.66	35.34	2.72	3.190	2.709	2.358	0.186	0.2158	0.183	0.1635
UC	57.6	59.01	56.53	82.16	3.46	3.660	3.339	4.367	0.218	0.2375	0.2065	0.2444
UB	37.5	40.1	37.13	28.26	7.89	8.295	7.659	5.956	1.751	1.8018	1.6696	1.3512
UD	30.3	26.3	28.20	78.9	2.66	2.35	2.15	3.58	0.335	0.293	0.235	0.232
KK	15.3	13.5	15.80	18.7	0.76	0.64	0.73	0.88	0.045	0.033	0.035	0.044
KP	63.0	62.7	64.00	67.5	3.11	3.15	3.13	3.19	0.155	0.162	0.159	0.155
TP	34.3	31.4	33.30	76.5	2.79	2.69	2.37	3.67	0.357	0.340	0.256	0.242
ID	3.4	3.9	4.00	4.5	0.21	0.27	0.28	0.26	0.017	0.028	0.029	0.021
GS	26.2	25.6	26.90	28.6	2.29	2.45	2.52	2.42	0.319	0.349	0.352	0.327
AJ	45.0	47.0	45.30	43.4	2.56	2.71	2.60	2.44	0.152	0.168	0.164	0.148
BQ	21.5	20.4	21.00	31.8	1.93	1.71	1.67	2.19	0.253	0.205	0.182	0.193
ND	18.3	17.9	19.10	20.5	2.13	2.29	2.35	2.27	0.393	0.424	0.427	0.403
EP	2.9	3.2	3.30	3.8	0.53	0.64	0.66	0.62	0.133	0.158	0.163	0.142
AC	18.2	18.8	18.70	19.3	1.76	1.87	1.84	1.83	0.209	0.232	0.233	0.221
DY	11.1	11.5	11.20	10.6	1.31	1.40	1.39	1.27	0.228	0.255	0.258	0.227
Σ	289.5	**282.2**	290.8	404.1	22.04	22.17	**21.69**	24.62	2.596	2.647	2.493	**2.355**

Tt = type of treatment; S^2 = Equation (5); X^2 = Equation (7); CV^2 = Equation (6)

The values for all three types of experimental errors (square absolute S^2, square relative CV^2 and Pearson's X^2) and for all four analyzed cases are presented in Table 7 (theoretical frequency estimated from the contingency table - Equation (12); theoretical frequency estimated through the minimization of the square absolute error - Equation (9); theoretical frequency estimated through minimization of the square relative error - Equation (10); theoretical frequency estimated through minimization of Pearson-

Fisher statistics - Equation (11)); the values are obtained in a design of experiments with two independent factors (type of treatment and type of potato variety, abbreviated as factors A and B). This experiment allowed the representation of the Euclidian distances between obtained results (see Figure 1).

Figure 1. Euclidian distances among estimations of experimental errors.



$\boxed{S}$ = coordinate of experimental error in hypothesis of S^2=min;
$\boxed{V}$ = coordinate of experimental error in hypothesis of CV^2=min;
X_S = coordinate of absolute square error (S^2) in hypothesis of X^2=min;
X_V = coordinate of relative square error (CV^2) in hypothesis of X^2=min;
X_X = coordinate of X^2 statistics in hypothesis of X^2=min;
E_S = coordinate of experimental absolute square error (S^2) in hypothesis of contingency;
E_V = coordinate of experimental relative square error (CV^2) in hypothesis of contingency;
E_X = coordinate of Pearson-Fisher (X^2) statistics in hypothesis of contingency;
S = triangle of absolute square errors (S^2); V = triangle of relative square errors (CV^2); X = triangle of X^2 statistics; E = triangle of contingency; $\boxed{X}$ = triangle of X^2 statistics variation;

The experimental errors estimated by Equations (9)–(11) are presented in Figure 2 using the Snyder triangle [49] (diagrams frequently used in chromatography in order to represent three or more parameters which depend on two factors).

Figure 2 was obtained by setting the representation at the same scale of error area in ratio with two factors (the distance between the coordinate of experimental error for the hypothesis that S^2 = min and the coordinate of experimental errors for the hypothesis that CV^2 = min was used as reference). The coordinates for the hypothesis that X^2 = min were obtained after maximizing the error area (maximization of the A, V and X triangle area). The coordinates of contingency were obtained so that its projections on the sides of the triangles could split the sides into the ratios observed among the differences in Table 7.

The graphical representation in Figure 1 provides qualitative remarks regarding the contingency model defined in Equation (12) in relation with experimental errors:

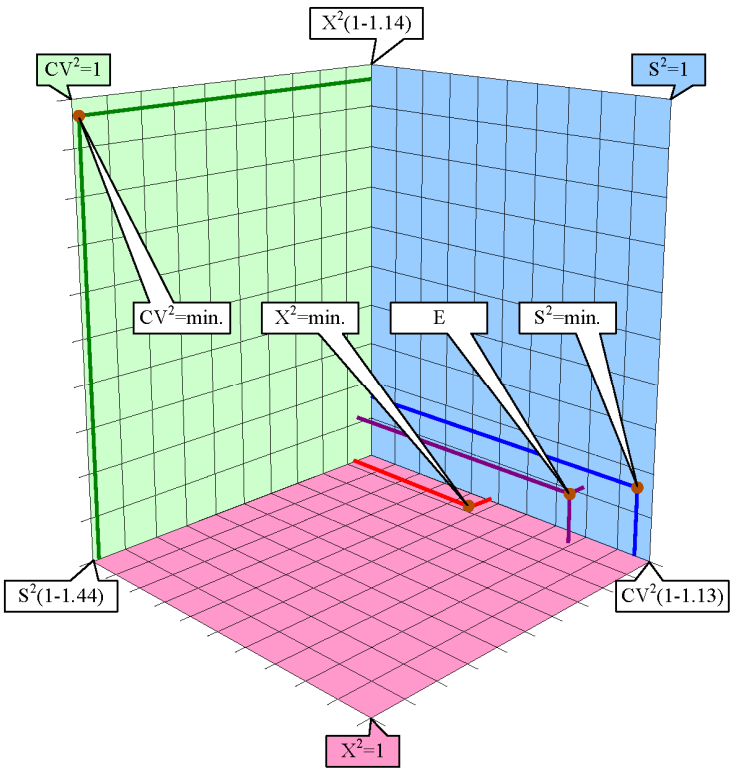
- The intersection between the contingency area and error areas is done through the absolute square error. Therefore, the contingency defined by Equation (12) assured the agreement between

observation and model for the absolute square error only (one out of the three types of errors included in the study).

- The triangle of the X^2 statistics variation intersects only with the X^2 statistics triangle. This fact recommends the use of optimization defined in Equation (5) [5] or the one defined in Equation (7) [39]. Moreover, this demonstrated why the Chi-square test is more exposed to type I errors (the null hypothesis that the row variable is not related to the column variable is rejected even if this hypothesis is true) [50] compared to the Kolmogorov-Smirnov [40,51] and Anderson-Darling [39,43] tests.

The analysis of errors distribution obtained from the above association analysis is presented in Supplementary Material.

Figure 2. Position of empirical estimation (Equation (12)) within minimum relative errors (Equations (9)–(11)).



The relative position of the solution proposed in Equation (12) could be represented in relation to the optimal values obtained using Equations (9)–(11). Therefore, the values presented in Table 7 (last row) were re-arranged and then expressed after being divided to their minimum values. The results are presented in Table 8.

Figure 2 contains the representation of the relative values of errors (error excess) in the coordinates defined by the values of S^2, CV^2 and X^2 for the results obtained through simple estimation (E, Equation (12)), minimization of the absolute square error (S^2 = min, Equation (9)), minimization of the relative square error (CV^2 = min, Equation (10)) and minimization of the X^2 statistics (X^2 = min, Equation (11)).

Table 8. Transformation of the residuals presented in Table 7 in relation to their minimum values.

Absolute value	S^2	X^2	CV^2
E	289.5	22.04	2.596
S^2 = min.	**282.2**	22.17	2.647
X^2 = min.	290.8	**21.69**	2.493
CV^2 = min.	404.1	24.62	**2.355**
Relative value	S^2	X^2	CV^2
E	1.026	1.016	1.102
S^2 = min.	**1**	1.022	1.124
X^2 = min.	1.030	**1**	1.059
CV^2 = min.	1.432	1.135	**1**

E = use of Equation (4) in place of a_ib_j in Equations (5)–(7); S^2 = Equation (5); X^2 = Equation (7); CV^2 = Equation (6).

The results of the representations showed in Figure 2 are consistent with the results of the projections in the areas illustrated in Figure 1. Figure 2 showed that the solution proposed by Equation (12) is very close to the solution proposed by Equation (9) and Equation (11). Moreover, the solution is intermediate between Equation (9) and Equation (11) and far away from the solution proposed by Equation (10).

3.3. Chi-Square Test of Independence

A single degree of freedom is known to exist for a 2 × 2 contingency table.

Table 9 presents such a situation in which the restrictions come from the sums of observations.

Table 9. 2 × 2 contingency table with one degree of freedom.

X^2	**Class A**	**Class Ω_1\A**	**Total Ω_1**
Class B	x	$n_1 - x$	n_1
Class Ω_2\B	$n_2 - x$	$n_3 - n_1 + x$	$n_2 + n_3 - n_1$
Total Ω_2	n_2	n_3	n_2+n_3

X^2 = Chi-Square. Class A = first value of first category. Ω_1 = whole first category. Class B = first value of second category. Ω_2 = whole second category.

The probability to observe the situation presented in Table 9 is given by the multinomial distribution (given by (Equation (15)). The value of the Chi-square statistics (X^2) is given by the relation presented in Equation (16).

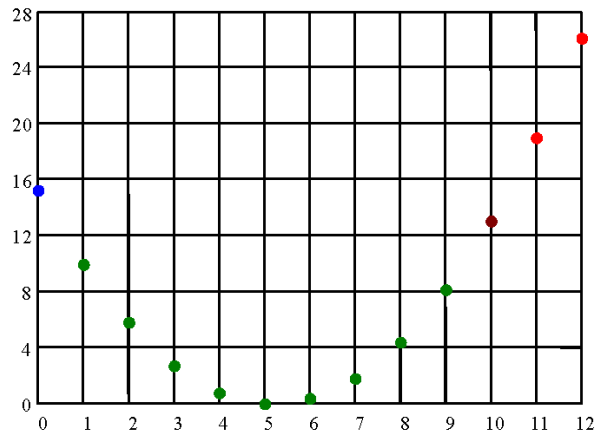
$$p_{MN}(x;n_1,n_2,n_3) = \frac{n_1!\cdot n_2!\cdot n_3!\cdot(n_2+n_3-n_1)!}{x!\cdot(n_1-x)!\cdot(n_2-x)!\cdot(n_3-n_1+x)!\cdot(n_2+n_3)!} \tag{15}$$

$$X^2(x;n_1,n_2,n_3) = \frac{(xn_2+xn_3-n_1n_2)^2(n_2+n_3)}{n_1n_2n_3(n_2+n_3-n_1)} \tag{16}$$

The range in which *x* could take values is [0. Min (n_1,n_2)].

In order to exemplify this problem, the experimental data reported by Fisher in 1935 [19] ($n_1 = 13$, $n_2 = 12$, $n_3 = 18$) for a range of *x* variation from 0 to 12 and with an observed value of *10* was analyzed. The value of X^2 statistic (Equation (16)) was represented in Figure 3.

Figure 3. Value of X^2 statistic as function of independent observation *x*.



The space of possible observations regarding the X^2 statistics as function of the independent variable *x* is discrete as it can be observed from Figure 3. The observed value (x = 10) is situated into the vicinity of one boundary (x = 12) having only two less favorable observations (with an X^2 value higher than the observed value) compared with the observed value in the same vicinity (x = 11 and x = 12) and a less favorable observation in the opposite vicinity (x = 0).

Two possible approaches could be applied in relation to the objective of the comparison in a contingency table:

(1) If higher distances from homogeneity than the observed gives the statistic, then the probability associated to observation is obtained by cumulating the probabilities for x = 0, x = 10, x = 11 and x = 12 (red and blue dots on Figures 3 and 4).

(2) If higher distances from homogeneity than the observed strictly in the sense of the observed gives the statistic, then the probability associated to observation is obtained by cumulating the probabilities for x = 10, x = 11 and x = 12 (red dots on Figures 3 and 4).

Figure 4 present graphically the probability of the observation (calculated from Equation (15)).

Table 10 presents the values of three probabilities: the probability of χ^2 distribution (p_{x2}), the probability to observe a higher distance from homogeneity in the direction of the observed value (p_{O2}) and the probability to observe a higher distance from the homogeneity in any direction (p_{D2}). The probability obtained from the χ^2 distribution (p_{X2}) estimates a higher distance from homogeneity in any direction (p_{D2}).

Figure 4. Value of the statistical probability of the observed according to the observable.

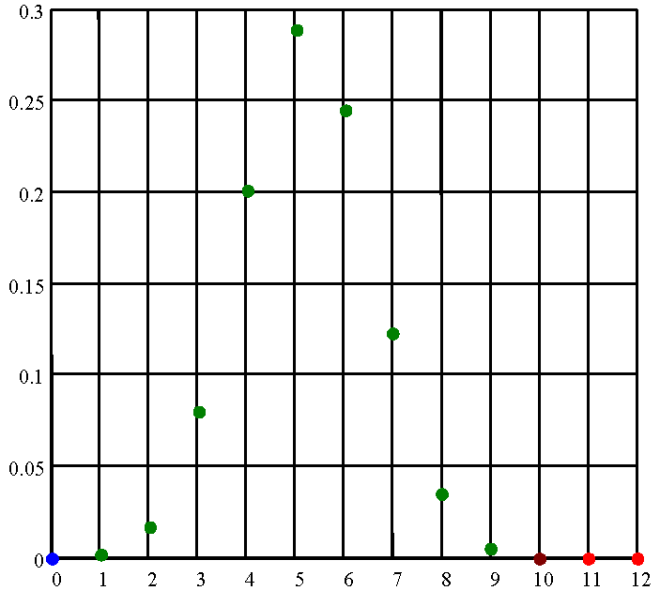


Table 10. Probability of observation.

Probability	Expression of calculus	Value
p_{X2}	$\chi^2_{CDF}(X^2 = 13.03, df = 1)$	3.063×10^{-4}
p_{O2} $(x^2 \geq X^2)$	$p_{MN}(10,13,12,18) + p_{MN}(11,13,12,18) + p_{MN}(12,13,12,18)$	4.625×10^{-4}
p_{O2} $(x^2 > X^2)$	$p_{MN}(11,13,12,18) + p_{MN}(12,13,12,18)$	1.548×10^{-5}
p_{D2} $(x^2 \geq X^2)$	$p_{O2}(x^2 \geq X^2) + p_{MN}(0,13,12,18)$	5.367×10^{-4}
p_{D2} $(x^2 > X^2)$	$p_{O2}(x^2 > X^2) + p_{MN}(0,13,12,18)$	8.702×10^{-5}

p_{X2} = probability of χ^2 distribution; p_{O2} = probability of observing a higher distance from homogeneity in the direction of the observed value; p_{D2} = probability of observing a higher distance from homogeneity in any direction; χ^2_{CDF} = probability of cumulative distribution function; p_{MN} = probability from multinomial distribution.

Table 10 shows how the Chi-square test is in error when the values in the contingency table are far from the imposed conditions for expected counts or frequencies (no more than 20% of the cells in the contingency table should have counts/frequencies lower that 5). Table 10 also shows how in this case the Chi-square test is exposed to type I errors (giving a lower observation probability than the real one; the risk is to accept the alternative hypothesis even if this is not true).

Frank Yates proposed in 1934 [18] a continuity correction in order to correct the statistical significance in a contingency table. If this correction is applied to Equations (1)–(3), a 0.5 value must be subtracted from the absolute difference between observed and expected frequencies in the hypothesis of independence (the middle of the frequency interval). Mantel and Haenszel proposed in 1959 [52] a correction of Chi-square test by dividing its value to df/(df − 1), where df = degree of freedom.

4. Conclusions

The application of the Chi-square test is directly related with some assumptions and with the design of the experiment. Three problems were identified in the application of Chi-square goodness-of-fit and solutions were identified, presented and analyzed.

Three different equations were identified as able to determine the contribution of each factor on three hypothesizes (minimization of variance, minimization of square coefficient of variation and minimization of X^2 statistic) in the application of the Chi-square test of homogeneity. The best solution proved to be directly related to the distribution of the experimental error.

The Fisher exact test proved to be the "golden test" in analyzing the independence while the Yates and Mantel-Haenszel corrections could be applied as alternative tests.

Acknowledgements

The study was supported by UEFISCSU/ID1105/2008 for R. Sestraş and by POSDRU/89/1.5/S/62371 through a fellowship for L. Jäntschi.

References

1. Fisher, R.A. Biometry. *Biometrics* **1948**, *4*, 217-219.
2. Fisher, R.A. Statistics. In *Scientific Thought in the Twentieth Century*; Heath, A.E., Ed.; Watts: London, UK, 1951, pp. 31-55.
3. Pearson, K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philos. Mag.* **1900**, *50*, 157-175.
4. Fisher, R.A. On the interpretation of χ^2 from contingency tables, and the calculation of P. *J. R. Stat. Soc.* **1922**, *85*, 87-94.
5. Fisher, R.A. The conditions under which χ^2 measures the discrepancy between observation and hypothesis. *J. R. Stat. Soc.* **1924**, *87*, 442-450.
6. Mirvaliev, M. The components of chi-squared statistics for goodness-of-fit tests. *J. Sov. Math.* **1987**, *38*, 2357-2363.
7. Plackett, R.L. Karl pearson and the Chi-squared test. *Int. Statist. Rev.* **1983**, *51*, 59-72.
8. Baird, D. The fisher/pearson Chi-squared controversary: A turning point for inductive inference. *Br. J. Philos. Sci.* **1983**, *34*, 105-118.
9. Cochran, W.G. Some methods for strengthening the common chi-square tests. *Biometrics* **1954**, *10*, 417-451.
10. Agresti, A. *Introduction to Categorical Data Analysis*; John Wiley and Sons: New York, NY, USA, 1996; pp. 231-236.
11. Levin, I.P. *Relating Statistics and Experimental Design*; Sage Publications: Thousand Oaks, CA, USA, 1999.
12. Neyman, J.; Pearson, E.S. *On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference, Part I.*; reprinted at pp. 1-66 in *Joint Statistical Papers*; Neyman, J., Pearson, E.S. Eds.; Cambridge University Press: Cambridge, UK, (originally published in 1928), 1967.
13. Neyman, J.; Pearson, E.S. *The Testing of Statistical Hypotheses in Relation to Probabilities a Priori*; reprinted at pp. 186-202 in *Joint Statistical Papers*; Neyman, J., Pearson, E.S., Eds.; Cambridge University Press: Cambridge, UK, (originally published in 1933), 1967.

14. Pearson, E.S.; Neyman, J. *On the Problem of Two Samples*; reprinted at pp. 99-115 in *Joint Statistical Papers*; Neyman, J., Pearson, E.S., Eds.; Cambridge University Press: Cambridge, UK, (originally published in 1930), 1967.
15. Rosner, B. *Fundamentals of Biostatistics. Chapter 10. Chi-Square Goodness-of-fit*, 6th ed.; Thomson Learning Academic Resource Center: Duxbury, MA, USA, 2006, pp. 438-441.
16. Roscoe, J.T.; Byars, J.A. An investigation of the restraints with respect to sample size commonly imposed on the use of the chi-square statistic. *J. Am. Stat. Assoc.* **1971**, *66*, 755-759.
17. Cochran, WG. The χ^2 test of goodness of fit. *Ann. Math. Stat.* **1952**, *25*, 315-345.
18. Yates, F. Contingency table involving small numbers and the χ^2 test. *Suppl. J. R. Stat. Soc.* **1934**, *1*, 217-235.
19. Fisher, R.A. The logic of inductive inference. *J. R. Stat. Soc.* **1935**, *98*, 39-54.
20. Koehler, K.J.; Larntz, K. An empirical investigation of goodness-of-fit statistics for sparse multinomials. *J. Am. Stat. Assoc.* **1980**, *75*, 336-344.
21. Li, G.; Doss, H. Generalized pearson-fisher Chi-square goodness-of-fit tests, with applications to models with life history data. *Ann. Stat.* **1993**, *21*, 772-797.
22. Moore, D.S.; Spruill, M.C. Unified large-sample theory of general Chi-squared statistics for tests of fit. *Ann. Stat.* **1975**, *3*, 599-616.
23. Moorea, D.S.; Stubblebinea, J.B. Chi-square tests for multivariate normality with application to common stock prices. *Commun. Stat. Theory Methods* **1981**, *10*, 713-738.
24. Mihalko, D.P.; Moore, D.S. Chi-Square Tests of Fit for Type II Censored Data. *Ann. Stat.* **1980**, *8*, 625-644.
25. Joe, H.; Maydeu-Olivares, A. A general family of limited information goodness-of-fit statistics for multinomial data. *Psychometrika* **2010**, *75*, 393-419.
26. Mantel, N. Chi-square tests with one degree of freedom; extension of the mantel-haenszel procedure. *J. Am. Stat. Assoc.* **1963**, *58*, 690-700.
27. Nathan, G. On the asymptotic power of tests for independence in contingency tables from stratified samples. *J. Am. Stat. Assoc.* **1972**, *67*, 917-920.
28. O'Brien, P.C.; Fleming, T.H. A multiple testing procedure for clinical trials. *Biometrics* **1979**, *35*, 549-556.
29. Tobin, J. Estimation of relationship for limited dependent variables. *Econometrica* **1958**, *26*, 24-36.
30. Overall, J.E.; Starbuck, R.R. F-test alternatives to fisher's exact test and to the Chi-square test of homogeneity in 2×2 tables. *J. Educ. Behav. Stat.* **1983**, *8*, 59-73.
31. Cox, M.K.; Key, C.H. Post hoc pair-wise comparisons for the Chi-square test of homogeneity of proportions. *Key Educ. Psychol. Meas.* **1993**, *53*, 951-962.
32. Pardo, L.; Martín, N. Omogeneity/heterogeneity hypotheses for standardized mortality ratios based on minimum power-divergence estimators. *Biom. J.* **2009**, *51*, 819-836.
33. Andrés, A.M.; Tejedor, I.H. Comments on 'Tests for the homogeneity of two binomial proportions in extremely unbalanced 2×2 contingency tables'. *Stat. Med.* **2009**, *28*, 528-531.
34. Baker, S.; Cousins, R.D. Clarification of the use of Chi-square and likelihood functions in fits to histograms. *Nucl. Instrum. Methods Phys. Res.* **1984**, *221*, 437-442.

35. Elmore, K.L. Alternatives to the Chi-Square Test for Evaluating Rank Histograms from Ensemble Forecasts. *Weather Forecast.* **2005**, *20*, 789-795.
36. Jin-Ting Zhang. Approximate and asymptotic distributions of Chi-squared-type mixtures with applications. *J. Am. Stat. Assoc.* **2005**, *100*, 273-285.
37. Gagunashvili, N.D. Chi-square tests for comparing weighted histograms. *Nucl. Instrum. Methods Phys. Res. Sect. A* **2010**, *614*, 287-296.
38. Snedecor, G.W.; Cochran, W.G. *Statistical Methods*, 8th ed.; Iowa State University Press: Iowa City, IA, USA, 1989.
39. Anderson, T.W.; Darling, D.A. Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Ann. Math. Stat.* **1952**, *23*, 193-212.
40. Kolmogorov, A. Confidence limits for an unknown distribution function. *Ann. Math. Stat.* **1941**, *12*, 461-463.
41. Fisher, R.A.; Mackenzie, W.A. Studies in crop variation. II. The manurial response of different potato varieties. *J. Agric. Sci.* **1923**, *13*, 311-320.
42. Sokal, R.R.; Rohlf, F.J. *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd ed.; Freeman: New York, NY, USA, 1994; pp. 729-739.
43. Fisher, R.A. *Statistical Methods for Research Workers*; Oliver and Boyd, Edinburgh, UK, 1934.
44. Scholz, F.W.; Stephens, M.A. K-sample anderson-darling tests. *J. Am. Stat. Assoc.* **1987**, *82*, 918-924.
45. Glass, G.V.; Hopkins, K.D. *Statistical Methods in Education and Psychology*, 3rd ed.; Allyn and Bacon: Needham Heights, MA, USA, 1996.
46. Hartley, R.V.L. Transmission of Information. *Bell Sys. Tech. J.* **1928**, *1928*, 535-563
47. Scott, D. *Multivariate Density Estimation*; John Wiley: Hoboken, NJ, USA, 1992.
48. Chi-square goodness-of-fit test. In: *NIST/SEMATECH e-Handbook of Statistical Methods*. Available online: http://www.itl.nist.gov/div898/handbook/prc/section2/prc211.htm (accessed 1 November 2010).
49. Snyder, L.R. Classification of the solvent properties of common liquids. *J. Chromatogr. A* **1974**, *92*, 223-230.
50. Steele, M.; Chaseling, J.; Hurst, C. Simulated Power of the Discrete Cramer-von Mises Goodness-of-fit Tests. In *Proceedings of the MODSIM 05 International Congress on Modelling and Simulation. Advances and Applications for Management and Decision Making*, Melbourne, VIC, Australia, 2005; pp. 1300-1304.
51. Smirnov, N.V. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Stat.* **1948**, *19*, 279-281.
52. Mantel, N.; Haenszel, W. Statistical aspects of the analysis of data from retrospective studies of disease. *J. Natl. Cancer Inst.* **1959**, *22*, 719-748.