

Department of Statistics, Universitas Islam Indonesia

Advances on Statistical Methods in Engineering, Science, Economy, Education and Disaster Management

Proceedings of
The First International Conference on
Statistical Methods in Engineering, Science,
Economy and Education (SESEE-2015)
Universitas Islam Indonesia,
Yogyakarta, Indonesia
September 19th-20th, 2015.

ISBN : 978-602-99849-2-7

Advances on Statistical Methods in Engineering, Science, Economy, Education and Disaster Management

Proceedings of
The First International Conference on Statistical Methods in Engineering, Science, Economy and Education (SESEE-2015)
Universitas Islam Indonesia,Yogyakarta, Indonesia
September 19th-20th, 2015.

Editor
Dr. RB. Fajriya Hakim, S.Si., M.Si.
Department of Statistics,
Universitas Islam Indonesia
e-mail: hakimf@uii.ac.id

ISBN : 978-602-99849-2-7

Advances on Statistical Methods in Engineering, Science, Economy, Education and Disaster Management

Proceedings of
The First International Conference on Statistical Methods in Engineering, Science, Economy and Education (SESEE-2015)
Universitas Islam Indonesia,Yogyakarta, Indonesia
September 19th-20th, 2015.

The First International Conference on Statistical Methods in Engineering, Science, Economy and Education (SESEE-2015)

Organized by:
Department of Statistics
Faculty of Mathematics and Natural Sciences.
Universitas Islam Indonesia

Published by
Department of Statistics
Faculty of Mathematics and Natural Sciences.
Universitas Islam Indonesia
Sleman, Yogyakarta
Indonesia.

Conference Organizer

Patron	
Dr. Ir. Harsoyo, M.Sc.	The Rector of Universitas Islam Indonesia
Advisor	
Prof. Drs. Suryo Guritno, M. Stat., Ph.D.	Universitas Islam Indonesia
Prof. Akhmad Fauzy, S.Si., M.Si., Ph.D.	Universitas Islam Indonesia
Drs. Allwar, M.Sc., Ph.D.	Universitas Islam Indonesia
General Chair	
Dr. RB. Fajriya Hakim, S.Si., M.Si.	Universitas Islam Indonesia
Organizing Committee	
Arum Handini Primandari, S.Pd.Si., M.Sc.	Universitas Islam Indonesia
Asmadini Handayani, S.Si., M.M.	Universitas Islam Indonesia
Atina Ahdika, S.Si., M.Si.	Universitas Islam Indonesia
Ayundyah Kesumawati, S.Si., M.Si.	Universitas Islam Indonesia
Muhammad Hasan Sidiq, S.Si., M.Sc.	Universitas Islam Indonesia
Muhammad Muhajir, S.Si., M.Si.	Universitas Islam Indonesia
Rahmadi Yotenka, S.Si., M.Si.	Universitas Islam Indonesia
Tuti Purwaningsih, S.Stat., M.Si.	Universitas Islam Indonesia
Reviewer	
Prof. Dr.rer-nat., Dedi Rosadi, S.Si., M.Sc.	Gadjah Mada University, Indonesia
Dr. Edy Widodo, S.Si., M.Si.	Universitas Islam Indonesia
Dr. Fatekurrohman, S.Si., M.Si.	UNEJ, Indonesia
Iwan Tri Riyadi Yanto, M.Sc.	Ahmad Dahlan University, Indonesia
Dr. Jaka Nugraha, S.Si., M.Si.	Universitas Islam Indonesia
Kariyam, M.Si.	Universitas Islam Indonesia
Dr. Techn. Rohmatul Fajriyah, S.Si., M.Si.	Universitas Islam Indonesia
Sugiyarto, S.Si., M.Si., Ph.D.	Ahmad Dahlan University, Indonesia
Dr. Suhartono, S.Si., M.Sc.	Institut Teknologi Surabaya, Indonesia
Assoc. Prof. Dr. Tutut Herawan	Universiti Malaya, Malaysia
Dr. Yosza Dasril	Universiti Teknikal Malaysia, Malaysia
Student Committee	
M Nasihun Ulwan	Universitas Islam Indonesia
Rudy Prietno	Universitas Islam Indonesia
Septianusa	Universitas Islam Indonesia

ISBN : 978-602-99849-2-7

Preface

We are honored to be part of this special event in The First International Conference on Statistical Methods in Engineering, Science, Economy, and Education (SESEE-2015) together with Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia.

This first edition of the conference is a continuation of the First Workshop on Big Data and Hadoop Implementation which has been successfully held in the Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia during the period of September 16-18, 2015 which has attracted more than 100 participants.

This first conference event we accept 26 submissions after rigorous review processes by program committee members of SESEE-2015. All registered and presented papers will be published by online proceedings and indexed by Google Scholar. The best selected papers will be invited by Scopus indexed journal.

On behalf of SESEE-2015, we would like to express our highest appreciation to be given the chance to do cooperation with KOMPAS Media for their support. Our special thanks go to the Rector of Universitas Islam Indonesia, the Dean of Faculty of Mathematics and Natural Sciences, Steering Committee, Program Committee Chairs, Organizing Chairs, all Program and Reviewer Committee members and all the additional reviewers for their valuable efforts in the review process that helped us to guarantee the highest quality of the selected papers for the conference.

We also would like to express our thanks to the three keynote speakers, Prof Dr Jemal H. Abawajy from Deakin University, Australia, Prof Dr Mustafa Mat Deris from Universiti Tun Hussein Onn Malaysia and Mr. Younes Sa'adi, MT., from Batna Universite, Algeria. We are also very pleased to have with us our best special coordinator in conducting this Conference, Dr. Tutut Herawan from Universiti Malaya, Malaysia.

We cordially thank all the authors for their valuable contributions and other participants of this conference. The conference would not have been possible without them. We wish to thank the members of the student committee for their very substansial work, especially whose who played essential roles, Septianusa, Rudy and Ulwan.

We hope that the papers become very useful references to enrich the ideas that can support the advancement of statistical method in various application.

SESEE-2015 Chair

RB Fajriya Hakim

Table of Contents

Cover Publishing Statement Preface

1 The Performance of Radial Basis Function Neural Network Model in Forecasting Foreign Tourist Flows To Yogyakarta..... 1
Dhoriva Urwatul Wutsqa and Syalsabila Mei Yasfi

2 Estimating The Break-Point of Segmented Simple Linear Regression Using Empirical Likelihood Method..... 7
Muhammad Hasan Sidiq Kurniawan and Zulaela

3 Risk Factor of Formaldehyde Detection on Sales Location of Jambal Roti Salted Fish (*Arius Thalassinus*) in Yogyakarta..... 19
Roza Azizah Primatika

4 Cluster Analysis and Its Various Problems..... 24
Erfiani

5 Volatility Modelling Using Hybrid Autoregressive Conditional Heteroskedasticity (ARCH) - Support Vector Regression (SVR)..... 30
Hasbi Yasin, Tarno and Abdul Hoyyi

6 Optimization of Fuzzy System Using Point Operation Intensity Adjustment for Diagnosing Breast Cancer..... 36
Kurrotul A'yun and Agus Maman Abadi

7 Thurston Method, Area Development Project Impact Evaluation in Pasaman Barat..... 43
Aam Alamudi, Kusman Sadik and Khairil Anwar Notodiputro

8 Simulation Study of Robust Regression in High Dimensional Data Through the Lad-Lasso..... 46
Septian Rahardiantoro and Anang Kurnia

9 An Implementation of Genetic Algorithm To Generate The Most Compromised Decision When Information of The Alternatives is Incomplete..... 49
Bagus Sartono and Septian Rahardiantoro

10 Estimation of Median Growth Charts for Children Based on Biresponse Semiparametric Regression Model by Using Local Linear Estimator..... 53
Nur Chamidah and Marisa Rifada

11 Feature Reduction of Wayang Golek Dance Data Using Principal Component Analysis (PCA)..... 60
Joko Sutopo, Adhi Susanto, Insap Santosa and Teguh Barata Adji

12 The Ability The Chi Squares Statistics to Rejecting The Null Hypothesis on Contingency Tables 2x2 66
Jaka Nugraha

13 Predictive Simulation of Amount of Claims With Zero Truncated Negative Binomial Distribution..... 72
Heri Kurniawan

14 Deterministic and Probabilistic Seismic Hazard Risk Analysis in Bantul Regency..... 77
Septianusa, Maulina Supriyaningsih and Atina Ahdika

15 Applying Extrapolation Technique to Flexible Binomial Model for Efficiency of American Option Valuation..... 83
Arum Handini Primandari and Indira Ihnu Brilliant

16 Small Area Estimation Considering Skewness Data and Spatially Correlated Random Area Effects 90
Dian Handayani, Anang Kurnia, Asep Saefuddin and Henk Folmer

17 Comparison of Binary, Uniform and Kernel Gaussian Weight Matrix in Spatial Autoregressive (SAR) Panel Data Model and the Application..... 96
Tuti Purwaningsih, Erfiani and Anik Djuraidah

18 Persistence Process of Stock Price Movements Based on Markov Chain Analysis. Case Study: Indonesia Stock Exchange (Idx)..... 103
Atina Ahdika and Bayun Matsaany

19 Application of Fuzzy Logic To Diagnose Severity of Coronary Heart Disease. Case Study in Dr. Zainoel Abidin General Hospital, Banda Aceh Indonesia..... 110
Zurnila Marli Kesuma, Hizir, Izazi

20 Statistics Application on Terestrial Phenomena of Metallic Mining's Activity.... 116
Eddy Winarno, Ira Mughni Pratiwi and Abdul Rauf

21 Empirical Study of Student's Stage Thinking According to Bloom and Van Hiele Learning Theories in Mathematics Instruction..... 122
Noening Andrijati, Budi Harjo and Zaenal Arifin

22 Service Quality in Religious and Common Tourism..... 127
Fety Ilma Rahmillah and Andi Rahadiyan Wijaya

23 Maximum Likelihood Estimation in Intervention Analysis Model Multy Input Step Function: The Impact of Sea Highway Policies on Stock Price Movements in Field Of Shipping Company (Tmas.Jk)..... 133
Wigid Hariadi and Abdurakhman

24 Indonesia's Province Segmentation Based on Flood Disaster Impact Using *Self Organizing Maps* (*SOM*) Algorithm..... 144
Muhammad Muhajir, Berky Rian Efanna and Reza Aditya Pratama

25 The Utilization Density Functional Theory in Structure Determination and Hydrogen Storage Properties of $Ca(Bh_4)_2{\cdot}2nh_3$ Compounds 151
Muhammad Arsyik Kurniawan S.

26 Statistical Analysis of The Difference Absolut Neutrophil Count (ANC) in The Level Sepsis Patients..... 157
Suharyanto and Rizka Asdie

The Ability The Chi Squares StatisticsTo Rejecting The Null Hypothesis on Contingency Tables 2x2

Jaka Nugraha

Statistics Department, Islamic University of Indonesia, Yogyakarta

jnugraha@uii.ac.id

Abstract: The Chi-Square statistic is the primary statistic used for testing the statistical significance of the cross-tabulation table. The Chi-Square test is based on an approcimation that works best when the expected frequencies are fairly large. There is still a lack of consensus on the optimum method most texts recommend the use of the chi squared test and there is disagreement on the boundary between 'large' and 'small' sample sizes. So in this paper discussed the limitations and requirements for proper use, especially related to the minimum number of sample requirements for table 2x2.
The study focussed on four versions of the Chi Squared test the tests were: Pearson's Chi-Squared test, Yates's Chi-Squared test, Likelihood Ratio Chi-Square test and '*N* - 1' Chi-Squared test. From each of the test statistic equations, composed functions that can explain the relationship between the size of the sample (m), the probability of each category (k) and the amount of difference can be detected (d). Concluded that it takes a larger sample when the value of k nearest 0.5 and smaller when approaching a value of 0 or 1. The four test statistics that have the same pattern, but the statistics Yate's need a sample of the largest while the three others the sample sizes are nearly equal.
Keywords: Pearson Chi-Square; Likelihood Ratio; Normal Distribution;'N - 1' Chi-Squared

1. Introduction

TheContingency table, also known as Cross-tabulation, is a joint frequency distribution of cases based on two or more categorical variables. Contingency table analysis is a common method of analyzing the association between two categorical variables. There are two separate sampling strategies lead to the contingency table analysis. First, Test of Independence. A single random sample of observations is selected from the population of interest, and the data are categorized on the basis of the two variables of interest. Second, Test for Homogeneity [1]. Separate random samples are taken from each of two or more populations to determine whether the responses related to a single categorical variable are consistent across populations.

The Chi-Square statistic is the primary statistic used for testing the statistical significance of the cross-tabulation table. The two-way table is set up the same way regardless of the sampling strategy, and the chi-square test is conducted in exactly the same way. Chi-Square tests whether or not the two variables are independent. The chi-square test is based on an approcimation that works best when the expected frequencies are fairly large. No expected frequency should be less than 1 and no more than 20% of the expected frequencies should be less than 5. For tables larger than 2x2, the chi-square distribution with the appropriate degrees of freedom provides a good approximation to the sampling distribution of Pearson's chi-square and the Likelihood Ratio Chi-Square. However, the Chi-Square statistic is only approximated by the chi-square distribution, and that approximation worsens with small expected frequencies. When we have very small expected frequencies, the possible values of the chi-square statistic are quite discrete. The general rule is that the smallest expected frequency should be at least five. However Cochran (1952), who is generally considered the source of this rule, acknowledged that the number "5" seems to be chosen arbitrarily [3].

Statistical test of 2 × 2 tables have been under discussion for a hundred years and dozens of research papers have been devoted to them. However, there is still a lack of consensus on the optimum method most texts recommend the use of the chi squared test for large sample sizes and the Fisher-Irwin test for small sample sizes, but there is disagreement on the boundary between 'large' and 'small' sample sizes [4]. Because the Pearson Chi-Square statistics and Likelihood RatioChi-Square statistics are based on a normal distribution approach, so in this paper discussed the limitations and requirements for proper use, especially related to the minimum number of sample requirements for table 2x2. Two distinct research designs can give rise to 2×2 tables : Model I (Row totals fixed, column totals free to vary), Model II (Both row & column totals free to vary). Model 1

where we wish to test the hypothesis that two proportions are equal. Model II where we wish to test the hypothesis of independence of two variables of classification.

The validity of the chi-square test depends on both the sample size and the number of cells. Several rules of thumb have been suggested to indicate whether the chi-square approximation is satisfactory. One such rule suggested by Cochran (1954) says that the approximation is adequate if no expected cell frequencies are less than one and no more than 20% are less than five. Because of the expected cell frequency criterion in the second sampling strategy, it may be necessary to combine similar categories to lessen the number of categories in your table or to examine the data by subcategories

2. Contingency Table 2x2 Analysis

A number of experiments involve binary outcomes (i.e., 1 and 0, yes and no). Typically, these occur when you are observing the presence or absence of a characteristic such as a disease, flaw, mechanical breakdown, death, failure, and so on. The analysis of the relationship between two bivariate categorical variables results in a 2 × 2 crosstabulation table of counts. There are *2* x *2* possible combinations of responses for these two variables. The *2* × *2* crosstabulation or contingency table has *2* rows and *2* columns consisting of *2* × *2* cells containing the observed counts (frequencies) for each of the *2* × *2* combinations. The observed frequencies are presented in Table 1.

Tabel 1. 2x2 Contingency table

	B	B^c	Total
A	a	b	m_1
A^c	b	d	m_2
Total	n_1	n_2	N

This type of analysis is called a contingency table analysis and is usually accomplished using a chi-square statistic that compares the observed counts with those that would be expected if there were no association between the two variables. Two separate sampling strategies lead to the chi-square contingency table analysis :

1. *Test for Homogeneity (Model I).* Separate random samples are taken from each of two or more populations to determine whether the responses related to a single categorical variable are consistent across populations. In this setting, you have a categorical variable collected separately from two or more populations. The hypotheses are as follows:
   H_0: The distribution of the categorical variable is the same across the populations.
   H_a: The distribution of the categorical variable differs across the populations.
2. *Test of Independence (Model II).* A single random sample of observations is selected from the population of interest, and the data are categorized on the basis of the two variables of interest. In this case, you have two variables and are interested in testing whether there is an association between the two variables. Specifically, the hypotheses to be tested are the following:
   H_0: There is no association between the two variables.
   H_a: The two variables are associated.

The two-way table is set up the same way regardless of the sampling strategy, and the chi-square test is conducted in exactly the same way. The only real difference in the analysis is in the statement of the hypotheses and conclusions. The chi squared test were (1) Pearson's Chi-Squared test (2) Yates's Chi-Squared test (3) Likelihood Ratio Chi-Square test (4) The '*N* - 1' Chi-Squared test.

The original chi-square test, often known as Pearson's chi-square, dates from papers by Karl Pearson in the earlier 1900s. The test serves both as a "goodness-of-fitt" test, where the data are categorized along one dimension, and as a test for the more common "contingency table", in which categorization is across two or more dimensions. The standard Pearson chi-square statistic for 2x2 Contingency table is defined as

$$\chi^2=(ad - bc)^2N / m_1m_2n_1n_2 \quad (1)$$

Pearson's chi-square statistic is not the only chi-square test that we have.

The likelihood ratio chi-square builds on the likelihood of the data under the null hypothesis relative to the maximum likelihood. It is defined as

$$\chi^2 = 2*(a*\log(aN/m_1n_1)+b*\log(bN/m_1n_2)+c*\log(cN/m_2n_1)+d*\mathrm{Log}(dN/m_2n_2)) \quad (2)$$

ISBN : 978-602-99849-2-7
Copyright © 2015by Advances on Science and Technology

If an expected frequency is lower than five, you have alternatives: Yates correction (Yates's Chi-Squared test) and the *N* - 1 chi-square test. [2]

Yates' correction (Yates, 1934) is equivalent to Pearson's chi-square but with a continuity correction.In cases where an expected frequency is below 5, Yates' correction brings the result more in line with the true probability.

$$\chi^2 = (|ad - bc| - \tfrac{1}{2}N)^2 N / m_1 m_2 n_1 n_2 \quad (3)$$

The *N - 1 chi-square test* is another option. Campbell (2007) carried out a very large sampling study on 2x2 tables comparing different chi-square statistics under different sample sizes and different underlying designs. He found that across all sampling designs, a statistic suggested by Karl Pearson's worked best in most situations. The statistic is defined as

$$\chi^2 = (ad - bc)^2 (N - 1) / m_1 m_2 n_1 n_2 \quad (4)$$

3. Calculation ofMinimumSamples Size

Assuming asample size of2mbyeachpopulation ofm, observational datacan be presented inTable2.P(A | B) =khandP(Ac |B) =xwith0<k <1 and0<x <1.

Table2. Tableof contingencyfor thesampleP(A | B) =k

	B	B^c	Total
A	Km	(1-k)m	m
A^c	Xm	(1-x)m	m
Total	(k+x)m	(2-k-x)m	2m

Based onchi squarestatisticformulaas inequation(1) s/d(4), going tolook forthe minimumnumber of samplesthat candetect the differencebetween theproportion ofpopulationI (A) anda population of2(A^c). The hypothesis is

$$H_0: P(A|B) = P(A^c|B) \text{ vs } H_0: P(A|B) = \neq P(A^c|B)$$

The minimumsample sizeto be able to$rejectH_0$withsignificance levelαforeachtest statisticsareasfollows:

1. Pearson's Statistics

$$\frac{(km(1-x)m - xm(1-k)m)^2 2m}{m.m(k+x)m(2-k-x)m.} \geq \chi_\alpha^2$$

$$\Leftrightarrow \quad m \geq \frac{\chi_\alpha^2 (k+x)(2-k-x)}{2(k-x)^2}$$

$$F(k;\alpha;x) = \frac{\chi_\alpha^2 (k+x)(2-k-x)}{2(k-x)^2} \quad (5)$$

$F(k;\alpha;x)$ is a function thatshowsthe minimumsamplerequired tobe able to$rejectH_0$.

2. Yate's Statistics

$$\frac{(|km(1-x)m - xm(1-k)m| - m)^2 2m}{m.m(k+x)m(2-k-x)m.} \geq \chi_\alpha^2$$

$$\Leftrightarrow \frac{(|k-x|m-1)^2 2}{(k+x)(2-k-x)m.} \geq \chi_\alpha^2$$

$$\Leftrightarrow \quad 2(k-x)^2 m^2 - (4|k-x| + (k+x)(2-k-x)\chi_\alpha^2)m + 2 \geq 0$$

$$\Leftrightarrow \quad m \geq \frac{4|k-x|(k+x)(2-k-x)\chi_\alpha^2 + \left(\left(4|k-x| + (k+x)(2-k-x)\chi_\alpha^2\right)^2 - 16(k-x)^2\right)^{1/2}}{2(k-x)^2}$$

$$F(k;\alpha;x) = \frac{4|k-x|(k+x)(2-k-x)\chi_\alpha^2 + \left(\left(4|k-x| + (k+x)(2-k-x)\chi_\alpha^2\right)^2 - 16(k-x)^2\right)^{1/2}}{2(k-x)^2} \quad (6)$$

ISBN : 978-602-99849-2-7
Copyright © 2015by Advances on Science and Technology

3. Campbell's Statistics

$$\frac{(km(1-x)m - xm(1-k)m)^2(2m-1)}{m.m(k+x)m(2-k-x)m.} \geq \chi_\alpha^2$$

$$\Leftrightarrow \quad m \geq \frac{\chi_\alpha^2(k+x)(2-k-x))}{2(k-x))^2} + \frac{1}{2}$$

$$F(k;\alpha;x) = \frac{\chi_\alpha^2(k+x)(2-k-x))}{2(k-x))^2} + \frac{1}{2} \qquad (7)$$

4. Likelihood's Statistics

$$ln\left(\left(\frac{2km}{(k+x)m}\right)^{2km}\left(\frac{2(1-k)m}{(2-k-x)m}\right)^{2(1-k)m}\left(\frac{2xm}{(k+x)m}\right)^{2xm}\left(\frac{2(1-x)m}{(2-k-x)m}\right)^{2(1-x)m}\right) \geq \chi_\alpha^2$$

$$\Leftrightarrow \quad 2m.ln\left(\left(\frac{2k}{(k+x)}\right)^{k}\left(\frac{2(1-k)}{(2-k-x)}\right)^{(1-k)}\left(\frac{2x}{(k+x)}\right)^{x}\left(\frac{2(1-x)}{(2-k-x)}\right)^{(1-x)}\right) \geq \chi_\alpha^2$$

$$\Leftrightarrow \quad m \geq \frac{\chi_\alpha^2}{2.ln\left(\left(\frac{2k}{(k+x)}\right)^{k}\left(\frac{2(1-k)}{(2-k-x)}\right)^{(1-k)}\left(\frac{2x}{(k+x)}\right)^{x}\left(\frac{2(1-x)}{(2-k-x)}\right)^{(1-x)}\right)}$$

$$F(k;\alpha;x) = \chi_\alpha^2\left(2.ln\left(\left(\frac{2k}{(k+x)}\right)^{k}\left(\frac{2(1-k)}{(2-k-x)}\right)^{(1-k)}\left(\frac{2x}{(k+x)}\right)^{x}\left(\frac{2(1-x)}{(2-k-x)}\right)^{(1-x)}\right)\right)^{-1} \qquad (8)$$

4. Properties of Functions F(*K;α;X*)

The properties ofthe function F(k; α; x) illustrated using Figure 1 and Figure 2 by taking the value of α=0.05. Figure 1 and Figure 2 describes the effects of changes in the value ofk, namely P(A∩B) and the difference between the value of P(A|B) and P(A^c|B).

χ2= 3,8415 and d=|k-x|

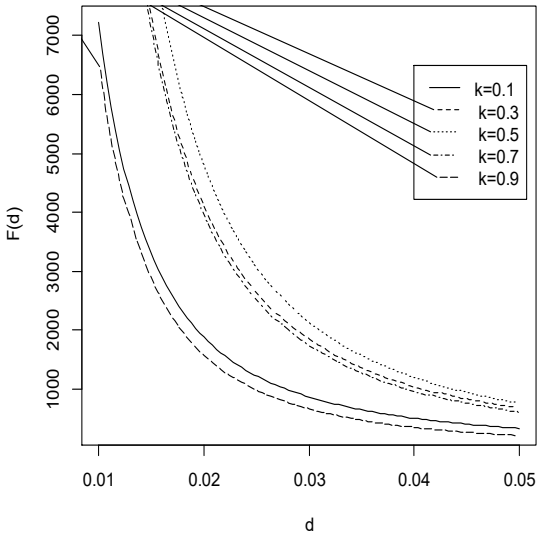


Figure 1. GraphF(k; α; x) on somevaluek

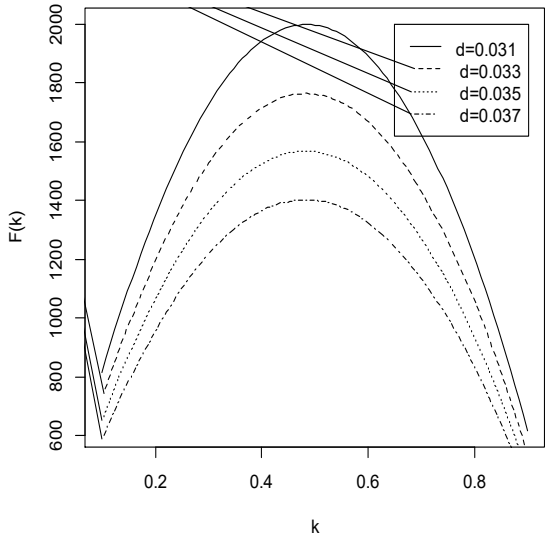


Figure 2. GraphF(k; α; x) on somevalued

As in Figure 1, the greater the ability to detect a difference (d smaller) the number of samples required is also greater. Figure 2 illustrates that the sample also needs greater when k approaches 0.5.
Furthermore, a comparison of several test statistics can be explained in Table 3 and Figure 3 to Figure 6. Table 3 is an example in the case of k = 0.5 and k = 0.4 which shows that (a) The number of

ISBN : 978-602-99849-2-7
Copyright © 2015by Advances on Science and Technology

samples required at Pearson's statistics, Likelihood's statistics and Champbell's statistics relatively similar (b) Yate's statistics require larger samples than the three other statistics. (c) the greater the ability to detect a difference (d smaller) the number of samples required also getting bigger.

Table 3. Tablecontingencyatk=0.4 andk=0.5

Nilai d	k=0,4				k=0,5			
	Pearson	Yate	Likelihood	Champbell	Pearson	Yate	Likelihood	Champbell
0,01	18514	19454	18515	18514	19206	20173	19206	19206
0,02	4647	5117	4647	4646	4800	5283	4801	4800
0,03	2073	2387	2013	2073	2133	2454	2133	2132
0,04	1170	1406	1171	1170	1199	1440	1200	1199
0,05	752	941	752	751	767	959	767	767

In general, to be able to detect the differences of the particular, Yate's statistics require larger samples than the likelihood's statistics, Campbell's statistics and Pearson's statistics. As in Figure 3 and Figure 4, likelihood's statistics, Campbell's statistics and Pearson's statisticshaverelatively similarproperties.. The four statistic has the same pattern (as a function of d or k).

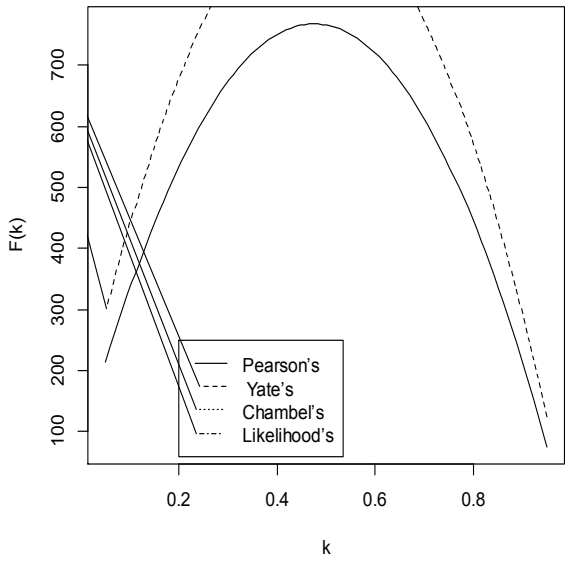


Figure 3.The sample sizeatd=0.05

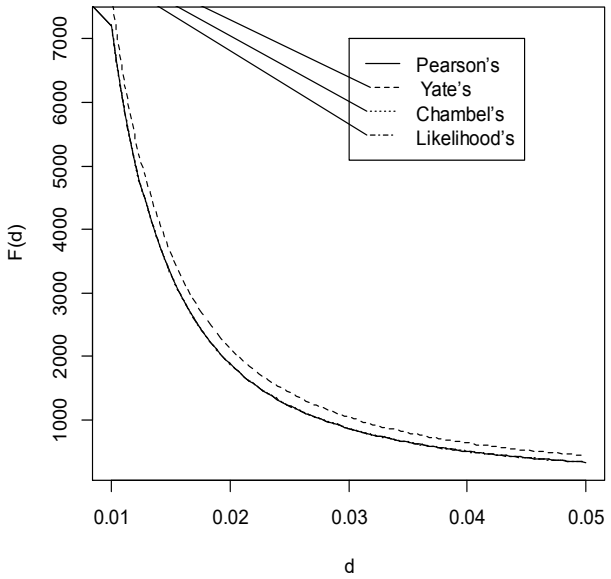


Figure 4.The sample sizeatk=0.1

In moredetail,to determine thestatistical differenceslikelihood's statistics, Campbell's statistics and Pearson's statistics taken short intervals. From Figure 5 and Figure 6, it appears that the order from the smallest sample sizer espectively : likelihood's statistics, and Pearson's statistics and Campbell's statistics.

ISBN : 978-602-99849-2-7
Copyright © 2015by Advances on Science and Technology

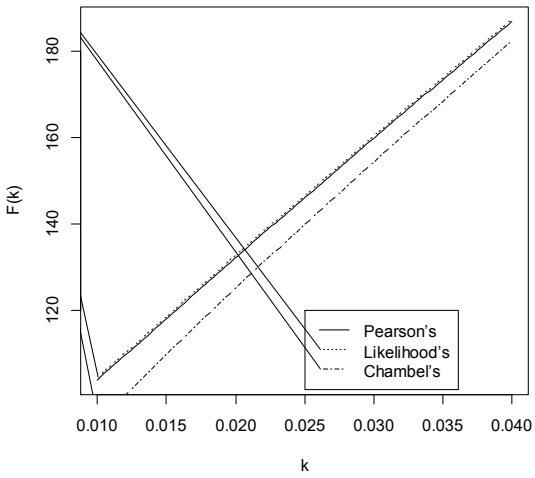


Figure 5.The sample size at k∈ (0.01,0.04)

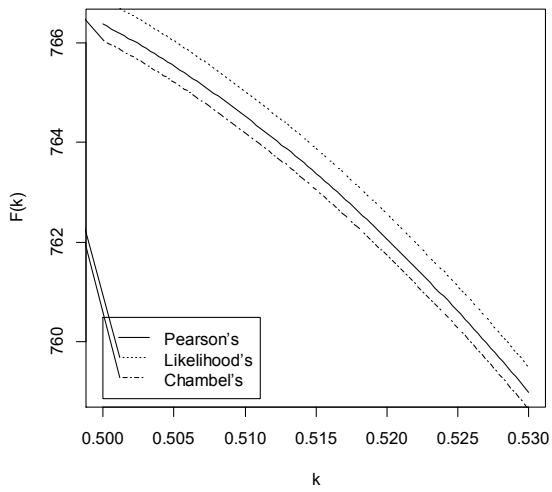


Figure 6.The sample size at k∈ (0.5,0.53)

5. Conclusion

Fromthe above discussionit can be concluded that (a) The higher thesample size, the ability todetect differences (rejecting H_0), the better. (b) The ability ofYate's statisticslowerbe compared toPearson's statistics, Likelihood's statisticsandChampbell's statistics. (c) Ability Likelihood's statistics relatively similar to the Pearson's statistics and better than the Champbell's statistics. should give a summary of:

Acknowledgement. The author would like to thank to Islamic University of Indonesia for supporting this research..

References

[1] Agresti A (2002), *Categorical Data Analysis*, John Wiley and Son
[2] Campbell, I. (2007), Chi-squared and Fisher-Irwin tests of two-by-two tables with small sample recommendations. *Statistics in Medicine*, 26, 3661-3675,.
[3] Cochran, W. G. (1952), The χ^2 test of goodness of fit. Annals of Mathematical Statistics, 25, 315-345.
[4] Yates F. (1984), Tests of significance for 2 × 2 contingency tables (with discussion) *Journal of the Royal Statistical Society Series A*; 147: 426-463.

ISBN : 978-602-99849-2-7
Copyright © 2015by Advances on Science and Technology