*Fundamental Journal of Mathematics and Applications*, 7(3) (2024), 169-185
Research Paper / Open Access

**Fundamental Journal of Mathematics and Applications**

*ISSN Online: 2645-8845*
*www.dergipark.org.tr/en/pub/fujma*
*https://doi.org/10.33401/fujma.1432348*

# The Most Powerful Member of the Power-Divergence Family for the Independence Model in Contingency Tables

**Gökçen Altun** [1,†]

[1] *Department of Econometrics, Ankara Hacı Bayram Veli University, Ankara, Türkiye*
[†] *gokcen.altun@hbv.edu.tr*

**Article Information**

**Keywords:** Goodness-of-fit test statistics; Independence model; Power-divergence family; Small sample

**AMS 2020 Classification:** 62Q05

**Abstract**

The family of power-divergence (PD) test statistic contains many well-known test statistics used in the analysis of the contingency tables under the independence model. In this work, we compare the various test statistics for the independence model. The type-I and type-II errors of the test statistics are obtained and compared via simulation study considering the different degree of freedoms and sample sizes. According to the simulation results, we recommend the PD(0.4) test statistic for the small sample size based on its power and type-I error rates. Two applications are given to demonstrate the usefulness of the PD(0.4) test statistic over the chi-square test statistic contingency tables.

## 1. Introduction

The chi-square test was developed by [1] to evaluate the association or difference between categorical variables. The chi-square test is commonly used in social and medical sciences to test the dependence structures of the levels of the categorical variables. The results of the chi-square test are misinterpreted by the researchers because of the lack of statistical knowledge [2]. Besides, the application of the chi-square test is very problematic for the small sample sizes which is ignored in many researches. It is well-known that the test statistic of the chi-square test follows the $\chi^2$ distribution. However, the asymptotic approximation is only valid for the non-sparse contingency tables and large sample sizes. Working with the less observations than needed reduces the power of the test. Therefore, to obtain the higher power value, one should work with required sample size based on the dimension of the contingency table [3, 4]. The determination of the sample size is done based on four inputs: type-I error, type-II error or power, effect size and degree of freedom (df). The type-I and type-II errors are the pre-determined inputs [5].

When the contingency tables have large number of cells, the frequencies of each cell may be very small or has zero frequencies. So, the contingency tables with large numbers of row and column variables yields the less observations in the cells. In this case, these contingency tables are called as sparse contingency tables [6]. The sparse contingency tables occur when the the values of 0 and 1 in many cells of the contingency table and the total number of cells are higher than the sample size [7, 8]. Besides, the sparseness index (SI) is useful to determine the sparse contingency tables. The SI is defined as

$$SI = \frac{n}{RC},$$

where $n$ is the sample size $R$ is the number of rows and $C$ is the number of columns in the contingency tables.

There are various studies in the literature for the comparison of goodness-of-fit test statistics in small samples. [9] performed a study to find a clear answer about what is the minimum value of the expected frequency and sample size to achieve the reasonable approximation to the $\chi^2$ distribution. [10] implemented a simulation study to compare the $\chi^2$, $G^2$ and [11] test statistic for 13 contingency tables. [10] found that the $\chi^2$ and Cressie and Read statistics can be used for smaller sample sizes than suggested by [9]. Several rule of thumb were suggested by researchers for $\chi^2$ approximation of the Pearson and likelihood ratio test statistics. [9] suggested that minimum cell expectation should be higher than $5t_5/t$ where $t_5$ is the number of cells where the expected frequency is smaller than 5 and $t$ is the total number of cells of the corresponding contingency

**Received:** 05-02-2024 **Revised:** 02-09-2024 **Accepted:** 24-09-2024 **Online:** 30-09-2024 **Published:** 30-09-2024

table. [12] suggested that the sample size should be higher than 4 or 5 times $t$. [13] showed that the $\chi^2$ statistic is much more appropriate than $G^2$ statistic for the small sample size. Recently, [14] performed a comprehensive simulation study to asses the small sample accuracy of the seven members of the power-divergence statistics for testing both independence and homogeneity in contingency tables. The results of the study of [14] showed that $G^2$ statistic rejects the null hypothesis too often in both sparse and non-sparse contingency tables. They suggested the non-asymptotic variant of $\chi^2$ statistic removes the deficiency of the Pearson $\chi^2$ statistic for sparse contingency tables. [15] investigated the determination of the power divergence parameter under quasi-independence model. More recently, [16] studied the asymptotic properties of $T^2$ test statistic under the symmetry model and concluded that the approximation of the $T^2$ test statistic to $\chi^2$ distribution is only valid for very large sample sizes. While the chi-square approach gives healthy results in tables with a degree of freedom greater than 1 and a maximum of 20 % of the expected frequencies below 5, this approach is weak in the sparse contingency table [7].

A general class of the test statistics was proposed by [11] and called as power-divergence (PD) family of statistics. The PD statistics contains Pearson's $\chi^2$, likelihood ratio statistic $G^2$, Freeman-Tukey's $T^2$, modified likelihood ratio statistics $GM^2$ and Neyman's modified $\chi^2$ as its sub-models. Note that these test statistics follow $\chi^2$ distribution [12, 17, 18]. This study compares the members of the PD test statistic using the different values of the parameter $\lambda$ based on the independence model. The type-I and power values of the test statistics are compared with simulation studies for different dimensions of the contingency tables. The goal of the simulation study is to find the most powerful test statistic for the independence model considering the sample sizes, type-I and type-II errors.

The other sections of the study is designed as follows. Firstly, the independence model is given in Section 2. The PD family of statistics is given in Section 3. The comparison of the test statistics via simulation studies is presented in Section 4. The recommended test statistic and Freeman-Halton (FH) test statistic is compared in Section 5. The power comparison of the most powerful test statistic and $\chi^2$ test statistic based on the real datasets is given in Section 6. The future work and conclusions of the presented study are given in Section 7.

## 2. Independence model

In the analysis of contingency tables, either ”row and column variables are independent of each other” or ”the constant levels of one of the variables do not differ between the other variable levels” are tested according to the researcher's purpose. The total chi-square of the calculations for the entire R×C table is divided into row, column and relationship components as follows

$$\chi_T^2 = \chi_R^2 + \chi_C^2 + \chi_{RC}^2. \tag{2.1}$$

In two dimensional tables, the independence hypothesis is expressed with (2.2)

$$H_0 : p_{ij} = p_{i.}p_{.j}, \;\; i = 1,2,3,...,R; \;\; j = 1,2,3,...,C. \tag{2.2}$$

The probability density function for the observed frequencies ($n_{ij}$) is as follows

$$P\left(n_{ij} \middle| p_{ij}, n\right) = \frac{n!}{\prod\limits_{i,j} n_{ij}!} \prod_{i,j} {p_{ij}}^{n_{ij}}. \tag{2.3}$$

Substituting $p_{ij} = p_{i.}p_{.j}$ in (2.3), we have

$$\begin{aligned} P\left(n_{ij} \middle| p_{ij}, n\right) &= \frac{n!}{\prod\limits_{i,j} n_{ij}!} \prod_{i,j} {p_{ij}}^{n_{ij}} = \frac{n!}{\prod\limits_{i,j} n_{ij}!} \prod_{i} {p_{i.}}^{n_{i.}} \prod_{j} {p_{.j}}^{n_{.j}} && (2.4) \\ &= \frac{n! \prod\limits_{i,j} {p_{i.}}^{n_{i.}}}{\prod\limits_{i} n_{i.}!} \frac{n! \prod\limits_{i,j} {p_{.j}}^{n_{.j}}}{\prod\limits_{j} n_{.j}!} \frac{\prod\limits_{i} n_{i.}! \prod\limits_{j} n_{.j}!}{n! \prod\limits_{i,j} n_{ij}!} && (2.5) \end{aligned}$$

Equality of $n_{ij} = np_{ij} + e_{ij}$ is written instead of $n$ in (2.5). When the natural logarithm is taken using the Stirling series expansion in (2.5), the three terms on the right side of the equation (2.5) follows approximately the chi-square distribution (see [19].

$$X_T^2 = \sum_{i=1}^{R} \sum_{j=1}^{C} \frac{\left(n_{ij} - np_{i.}p_{.j}\right)^2}{np_{i.}p_{.j}}. \tag{2.6}$$

The quantity in (2.6) follows the chi-square distribution with (RC-1) degrees of freedom. The first part on the right side of the equation is given by

$$\mathrm{X}_R^2 = \sum_i \frac{(n_{i.} - np_{i.})^2}{np_{i.}}, \tag{2.7}$$

which follows the chi-square distribution with R-1 df. The second part is given by

$$\mathrm{X}_C^2 = \sum_j \frac{(n_{.j} - np_{.j})^2}{np_{.j}},$$

which follows the chi-square distribution with C-1 df. The third part is the test statistic calculated for the independence hypothesis which is given by

$$\mathrm{X}_{RC}^2 = \sum_i \sum_j \frac{(n_{ij} - n_{i.}n_{.j}/n)^2}{n_{i.}n_{.j}/n}.$$

The df can be extracted using the relation given in (2.1). So, the df of the independence model is $(R-1)(C-1)$. The likelihood estimates of expected values $e_{ij}$ under independence model is $e_{ij} = n_{i.}n_{.j}/n$.

## 3. Power-divergence family

The PD family of statistics, $PD(\lambda)$, is given by

$$PD(\lambda) = \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{R} \sum_{j=1}^{C} n_{ij} \left[ \left( \frac{n_{ij}}{e_{ij}} \right)^{\lambda} - 1 \right], \tag{3.1}$$

where $i = 1,2,3,...,R,\ j = 1,2,3,...,C$ and $\lambda \in \Re$. When the $\lambda = 0$ and $\lambda = -1$, the equation (3.1) is not valid. So, the limiting cases of (3.1) for $\lambda = 0$ and $\lambda = -1$ are given as follows

$$\lim_{\lambda \to 0} \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{R} \sum_{j=1}^{C} n_{ij} \left[ \left( \frac{n_{ij}}{e_{ij}} \right)^{\lambda} - 1 \right] = 2 \sum_{i=1}^{R} \sum_{j=1}^{C} n_{ij} \left[ \ln \left( \frac{n_{ij}}{e_{ij}} \right) \right],$$

$$\lim_{\lambda \to -1} \frac{2}{\lambda(\lambda+1)} \sum_{i=1}^{R} \sum_{j=1}^{C} n_{ij} \left[ \left( \frac{n_{ij}}{e_{ij}} \right)^{\lambda} - 1 \right] = 2 \sum_{i=1}^{R} \sum_{j=1}^{C} e_{ij} \left[ \ln \left( \frac{e_{ij}}{n_{ij}} \right) \right],$$

respectively. As given in Section 1, the PD family of statistics contains various known test statistics as its sub-models.

- PD(1) reduces the Pearson's $\chi^2$ test statistics
- PD(0) reduces the likelihood ratio $G^2$ test statistics
- PD(-1/2) reduces the Freeman Tukey's $T^2$ test statistics
- PD(2/3) reduces the Cressie Read test statistic $C^2$

## 4. Simulation studies

Simulation studies are performed to evaluate the performance of the test statistics for the independence model. The multinomial distribution is used to generate contingency tables. The below probability matrices are used to obtain type-I errors of the test statistics. The probability matrices are generated by assuming that null hypothesis, $H_0$ is true.

| 3x3 contingency table | | | 3x4 contingency table | | | | 3x5 contingency table | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.10 | 0.06 | 0.04 | 0.03 | 0.03 | 0.02 | 0.02 | 0.03 | 0.03 | 0.09 | 0.06 | 0.09 |
| 0.15 | 0.09 | 0.06 | 0.06 | 0.06 | 0.04 | 0.04 | 0.02 | 0.02 | 0.06 | 0.04 | 0.06 |
| 0.25 | 0.15 | 0.10 | 0.21 | 0.21 | 0.14 | 0.14 | 0.05 | 0.05 | 0.15 | 0.10 | 0.15 |

**Table 1:** Probability matrices used to detect type-I errors for $R = 3$ and $C = 3,4,5$

| 4x4 contingency table | | | | 4x5 contingency table | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.02 | 0.03 | 0.01 | 0.04 | 0.02 | 0.02 | 0.06 | 0.04 | 0.06 |
| 0.04 | 0.06 | 0.02 | 0.08 | 0.03 | 0.03 | 0.09 | 0.06 | 0.09 |
| 0.06 | 0.09 | 0.03 | 0.12 | 0.03 | 0.03 | 0.09 | 0.06 | 0.09 |
| 0.08 | 0.12 | 0.04 | 0.16 | 0.02 | 0.02 | 0.06 | 0.04 | 0.06 |

**Table 2:** Probability matrices used to detect type-I errors for $R=4$ and $C=4,5$

| 5x5 contingency table | | | | |
|---|---|---|---|---|
| 0.01 | 0.01 | 0.03 | 0.02 | 0.03 |
| 0.01 | 0.01 | 0.03 | 0.02 | 0.03 |
| 0.03 | 0.03 | 0.09 | 0.06 | 0.09 |
| 0.02 | 0.02 | 0.06 | 0.04 | 0.06 |
| 0.03 | 0.03 | 0.09 | 0.06 | 0.09 |

**Table 3:** Probability matrix used to detect type-I errors for $R=5$ and $C=5$

Also, the below matrices are used to obtain power of the test statistics.

| 3x3 contingency table | | | 3x4 contingency table | | | | 3x5 contingency table | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.03 | 0.11 | 0.06 | 0.01 | 0.04 | 0.01 | 0.04 | 0.09 | 0.01 | 0.04 | 0.12 | 0.04 |
| 0.2 | 0.03 | 0.07 | 0.09 | 0.03 | 0.07 | 0.01 | 0.07 | 0.04 | 0.03 | 0.01 | 0.05 |
| 0.15 | 0.22 | 0.13 | 0.15 | 0.3 | 0.05 | 0.2 | 0.1 | 0.12 | 0.15 | 0.05 | 0.08 |

**Table 4:** Probability matrices used to detect powers for $R=3$ and $C=3,4,5$

| 4x4 contingency table | | | | 4x5 contingency table | | | | |
|---|---|---|---|---|---|---|---|---|
| 0.05 | 0.01 | 0.03 | 0.01 | 0.1 | 0.05 | 0.01 | 0.02 | 0.02 |
| 0.01 | 0.02 | 0.1 | 0.07 | 0.1 | 0.1 | 0.05 | 0.03 | 0.02 |
| 0.1 | 0.02 | 0.08 | 0.1 | 0.1 | 0.1 | 0.05 | 0.03 | 0.02 |
| 0.15 | 0.04 | 0.1 | 0.11 | 0.1 | 0.05 | 0.01 | 0.01 | 0.03 |

**Table 5:** Probability matrices used to detect powers for $R=4$ and $C=4,5$

| 5x5 contingency table | | | | |
|---|---|---|---|---|
| 0.015 | 0.015 | 0.04 | 0.01 | 0.02 |
| 0.02 | 0.02 | 0.015 | 0.03 | 0.015 |
| 0.02 | 0.04 | 0.07 | 0.09 | 0.08 |
| 0.03 | 0.04 | 0.03 | 0.07 | 0.03 |
| 0.04 | 0.06 | 0.06 | 0.09 | 0.05 |

**Table 6:** Probability matrix used to detect powers for $R=5$ and $C=5$

These probability matrices are generated by assuming that the alternative hypothesis, $H_1$ is true. The row and column marginal probabilities are degenerated to create the departure from the independence model. The significance level $\alpha$ is determined as 0.05. The interpretation of the simulation results are done based on the 0.06 value. The test statistics having the type-I error above the 0.06 value are considered as inappropriate. The simulation replication is determined as $N=10,000$.

Table 7 shows the effect sizes of the contingency tables used for the power calculation. Note that the effect sizes of the contingency tables used for the type-I error is zero. As reported in Table 7, the small and moderate effect sizes are used to compare the power values of the test statistics.

| Effect size | Dimension | | | | | |
|---|---|---|---|---|---|---|
| | 3x3 | 3x4 | 3x5 | 4x4 | 4x5 | 5x5 |
| w | 0.4341 | 0.3328 | 0.4691 | 0.3328 | 0.2564 | 0.2642 |

**Table 7:** The effect sizes of the contingency tables used for the power calculation.

### 4.1. Type-I error

Figure 1 displays the simulation results for the 3x3 contingency table. We also consider the sparseness index to analyze the behaviours of the test statistics for the small sample sizes. When the indicator SI is below 5 value, we call this contingency table as *sparse table*. So, the contingency table is called as sparse table if the number of observations is below 45 for the $R = 3$ and $C = 3$. This value is plotted in the figures vertically. According to the findings in Figure 1, we evaluate the convergence of the test statistics to $\chi^2$ distribution. From Figure 1, we observe that $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ are above the 0.06 value which is evidence that these test statistics do not converge to $\chi^2$ distribution. When the sample size is above 150, all test statistics work well, except $T^2$.

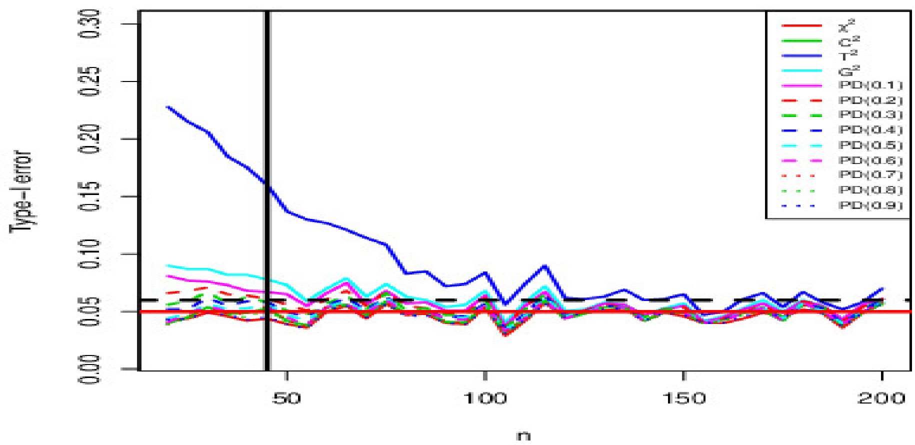


**Figure 1:** Type-I errors of the test statistics for R=3 and C=3.

Figure 2 displays the simulation results for the 3x4 contingency table. From these results, the $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ test statistics do not converge the $\chi^2$ distribution for both sparse and non-sparse contingency tables. The vertical line represents the sample size for the sparseness index which is 60. Additionally, the convergence of the $G^2$ statistic to $\chi^2$ distribution needs high sample sizes for $R = 3$ and $C = 4$ contingency tables. The $C^2$ performs better than the $G^2$ statistic. All test statistics converge to the $\chi^2$ distribution when the sample size is higher than 150, except $T^2$ statistic.

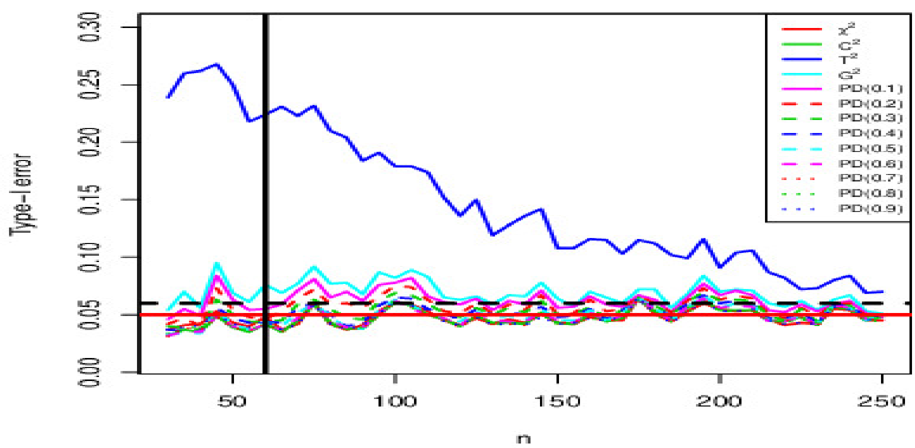


**Figure 2:** Type-I errors for R=3 and C=4.

Figure 3 displays the simulation results for the 3x5 contingency table. Again, the same test statistics fail to converge the $\chi^2$ distribution. Here, the vertical line is 75 for the sample size of sparseness index. The $G^2$ needs higher sample sizes to converge to $\chi^2$ distribution for $R = 3$ and $C = 5$.

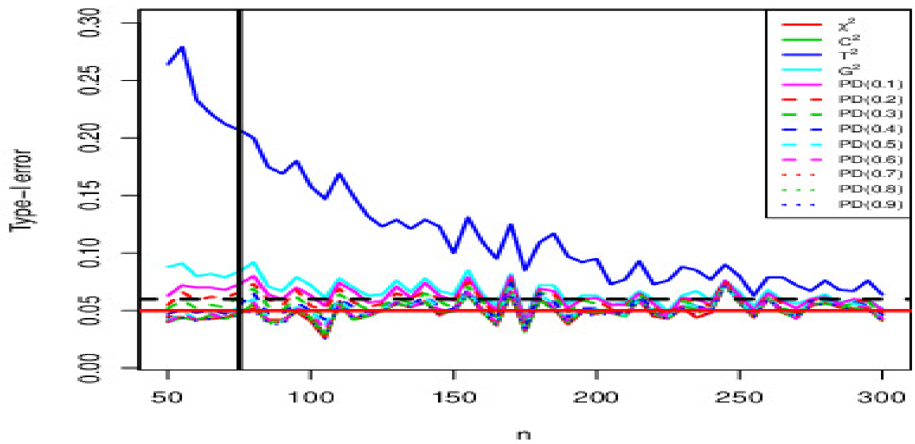


**Figure 3:** Type-I errors for R=3 and C=5.

The interpretation of the results of the 4x4, 4x5 and 5x5 contingency tables are similar to the previous simulation results. The results of these contingency tables are plotted in Figure 4. The vertical lines of the figures are 80, 120 and 125, respectively. From these figures, we conclude the convergence of the $G^2$ to $\chi^2$ distribution is not valid for the small sample sizes.

The below findings are observed based on the simulation results of the test statistics for type-I errors.

- The convergence of the $G^2$ statistic to $\chi^2$ distribution is very problematic for small sample sizes (see [20])
- The $C^2$ statistic performs better than $G^2$ statistic.
- The convergence of the $T^2$ statistic to $\chi^2$ distribution is only valid for the large sample sizes and it cannot be used for any small sample size.
- The dimension of the contingency table effects the convergence of the statistics.
- More sample size is needed for the high dimensional contingency tables.

So, end of the simulation study for the type-I errors of the test statistics, we eliminate the $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ statistics since they do not converge well to $\chi^2$ distribution. In the second step, we compare the power results of the test statistics converges well to $\chi^2$ distribution.

Additionally, we compare the p-values of the test statistics for R=3, C=3 and $n = 50$. Let $F_X$ be the distribution of the test statistic $X$ under the null hypothesis. If $F_T$ is continious, the p-value is distributed as $U(0,1)$ [21]. The distribution of the p-values for test statistics are evaluated via Kolmogorov-Smirnov (KS) test. The histograms of the p-values of the test statistics with the p-values of KS test are displayed in Figure 5. From these figures, it is clear that the distribution of the p-values of the $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ test statistics do not follow the $U(0,1)$ distribution since their p-values are lower than 0.05. It is evidence that these test statistics do not perform well for small sample sizes.

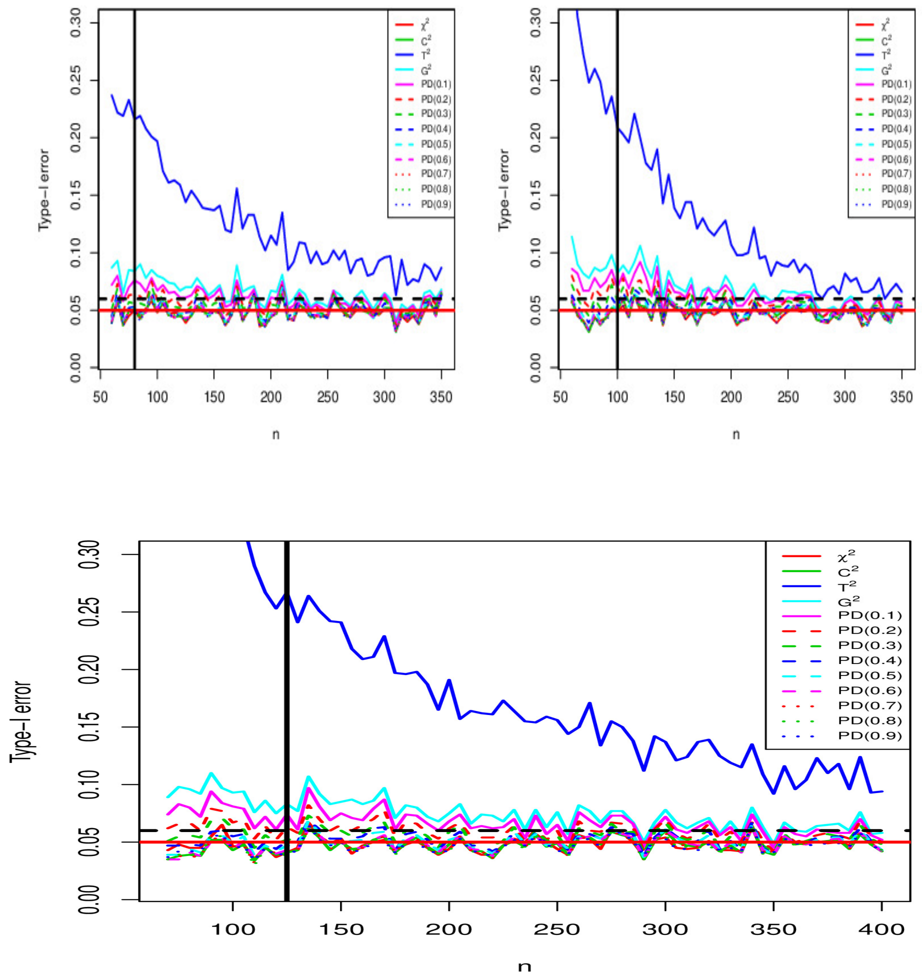


**Figure 4:** Type-I errors for (top-left) R=4 and C=4, (top-right) R=4 and C=5 and (bottom) R=5 and C=5

The below findings are observed based on the simulation results of the test statistics for type-I errors.

- The convergence of the $G^2$ statistic to $\chi^2$ distribution is very problematic for small sample sizes (see [20])
- The $C^2$ statistic performs better than $G^2$ statistic.
- The convergence of the $T^2$ statistic to $\chi^2$ distribution is only valid for the large sample sizes and it cannot be used for any small sample size.
- The dimension of the contingency table effects the convergence of the statistics.
- More sample size is needed for the high dimensional contingency tables.

So, end of the simulation study for the type-I errors of the test statistics, we eliminate the $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ statistics since they do not converge well to $\chi^2$ distribution. In the second step, we compare the power results of the test statistics converges well to $\chi^2$ distribution.

Additionally, we compare the p-values of the test statistics for R=3, C=3 and $n = 50$. Let $F_X$ be the distribution of the test statistic $X$ under the null hypothesis. If $F_T$ is continious, the p-value is distributed as $U(0,1)$ [21]. The distribution of the p-values for test statistics are evaluated via Kolmogorov-Smirnov (KS) test. The histograms of the p-values of the test statistics with the p-values of KS test are displayed in Figure 5. From these figures, it is clear that the distribution of the p-values of the $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ test statistics do not follow the $U(0,1)$ distribution since their p-values are lower than 0.05. It is evidence that these test statistics do not perform well for small sample sizes.

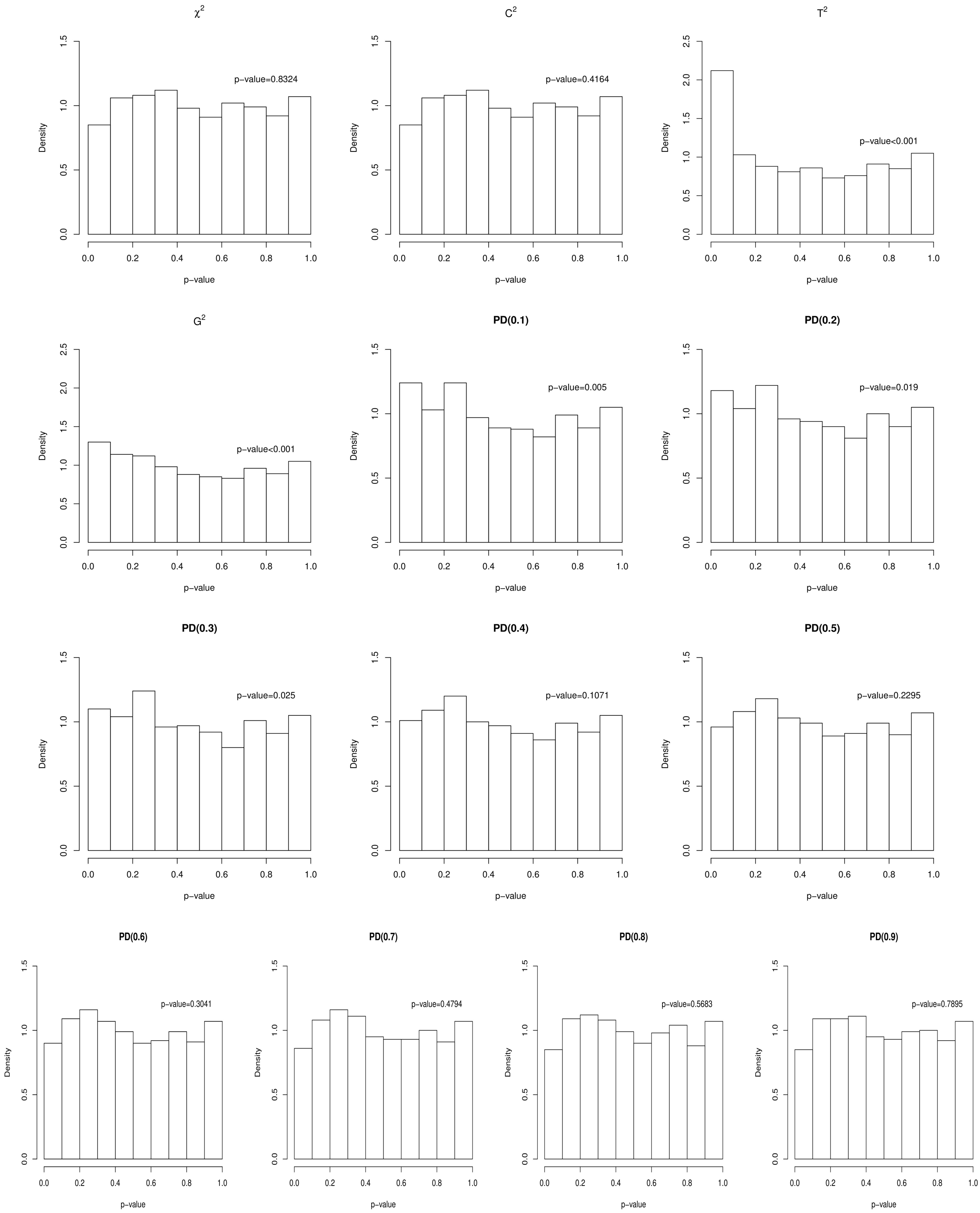


**Figure 5:** The distribution of p-values for the test statistics under R=3 and C=3 and $n = 50$

### 4.2. Power of test

Now, we examine the power results of each test statistics which are the members of the power-divergence family. The contingency tables are generated using the probability matrices given in Section 4.1. According to the results of the type-I errors of the test statistics, $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ do not converge the $\chi^2$ distribution. Although the power of test results are reported for all test statistics, $T^2$, $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ are not considered in evaluation of the power results.

Figure 6 displays the power results of the test statistics for 3x3 contingency table. As seen from these Figure, $T^2$ has the highest power value among others. However, since it does not converge the $\chi^2$ distribution, its power result is not meaningful. Similarly, the power results of the $G^2$, $PD(0.1)$, $PD(0.2)$ and $PD(0.3)$ statistics are also not meaningful. After eliminating these test statistics, the most powerful test statistics is $PD(0.4)$ for 3x3 contingency table. Vertical lines of the Figure 6 shows the minimum required sample size to achieve the 0.80 and 0.90 power values. The minimum sample size is 60 for the power 0.80 and minimum sample size is 75 for the power 0.90.

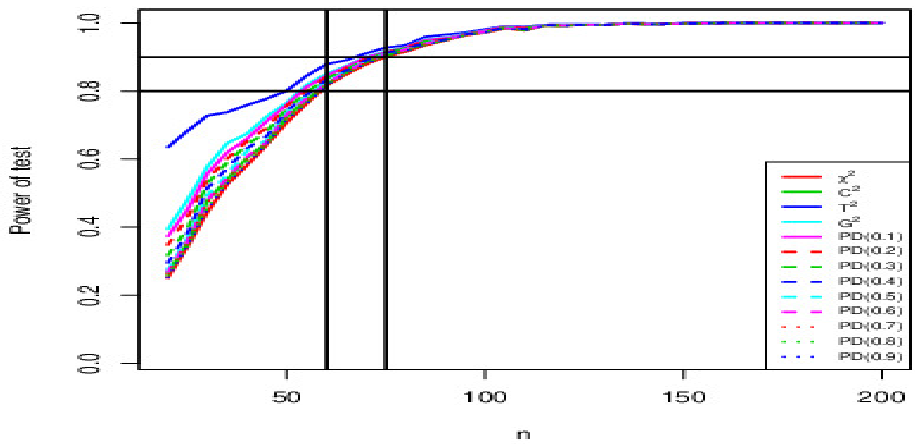


**Figure 6:** Power results under R=3 and C=3.

Figure 7 displays the power results of the test statistics for 3x4 contingency tables. These results are also in favour of the $PD(0.4)$ test statistics. The minimum sample size for the powers 0.80 and 0.90 are 65 and 80, respectively.

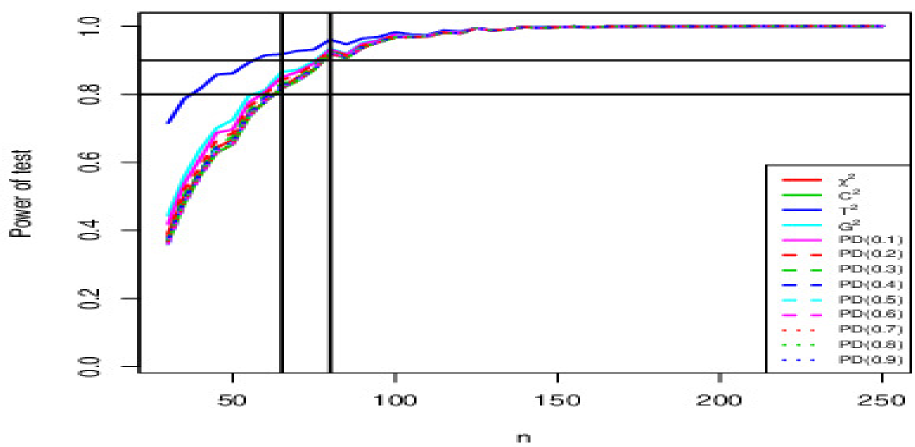


**Figure 7:** Power results under R=3 and C=4.

Similarly, Figure 8 displays the power results of the test statistics for 3x5 contingency table. The most powerful test statistic is $PD(0.4)$. As in previous results, the $PD(0.4)$ test statistics has the highest power among others. The minimum sample size for the powers 0.80 and 0.90 are 70 and 90, respectively.

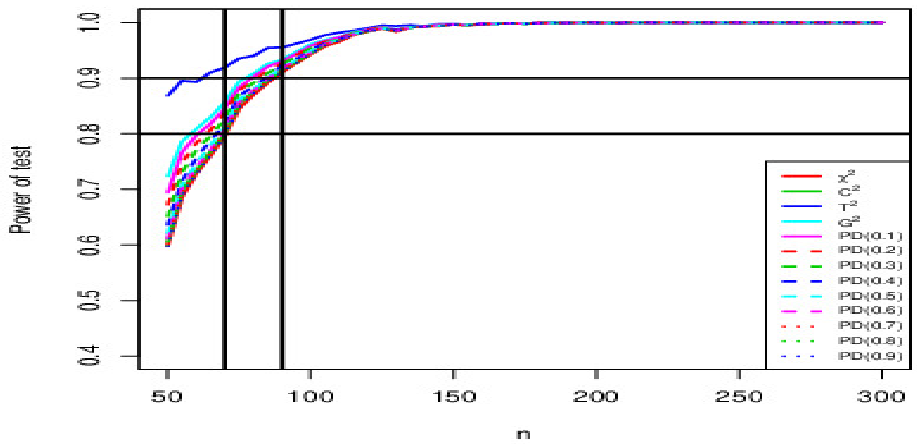


**Figure 8:** Power results under R=3 and C=5.

Figure 9 displays the power results of the test statistics for 4x4 contingency table. $PD(0.4)$ is again the most powerful test statistic among others. From these results, we conclude that the minimum required sample size is 130 to achieve at least 0.80 power and required sample size is 150 to achieve at least 0.90 power.

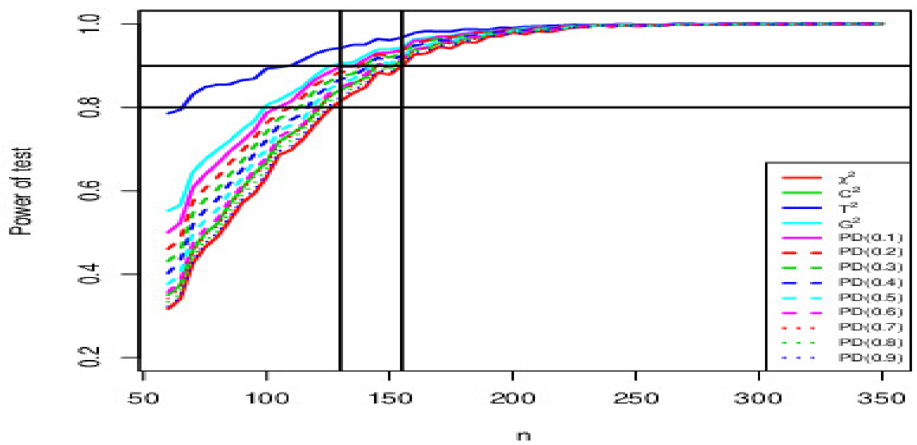


**Figure 9:** Power results under R=4 and C=4.

Figure 10 displays the power results of the test statistics for 4x5 contingency table. Results show that $PD(0.4)$ has the highest power. According to the vertical lines, the required sample size is 160 for the power 0.80 and 195 for the power 0.90.

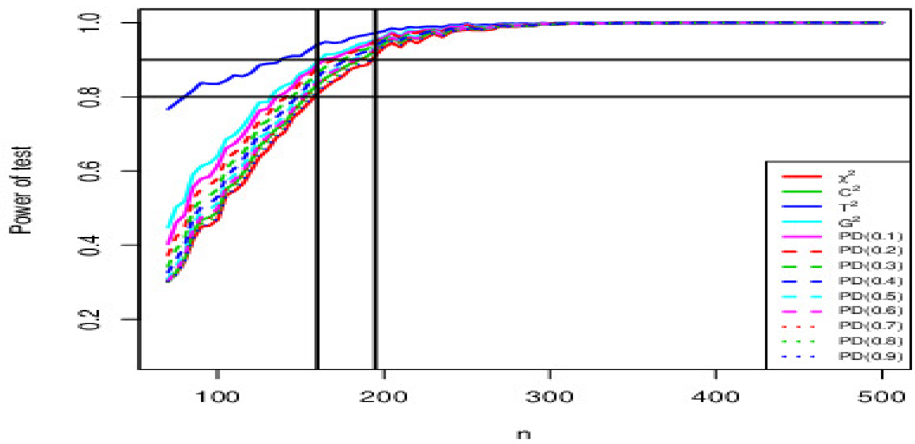


**Figure 10:** Power results under R=4 and C=5.

In a similar vein, Figure 11 displays the power results of the test statistics for 5x5 contingency table. Again, $PD(0.4)$ has the highest value of the power results. The required sample size is 280 for the power 0.80 and 350 for the power 0.90. As seen from these results, once the dimension of the contingency table increases, the required sample size increases to reach higher power values.

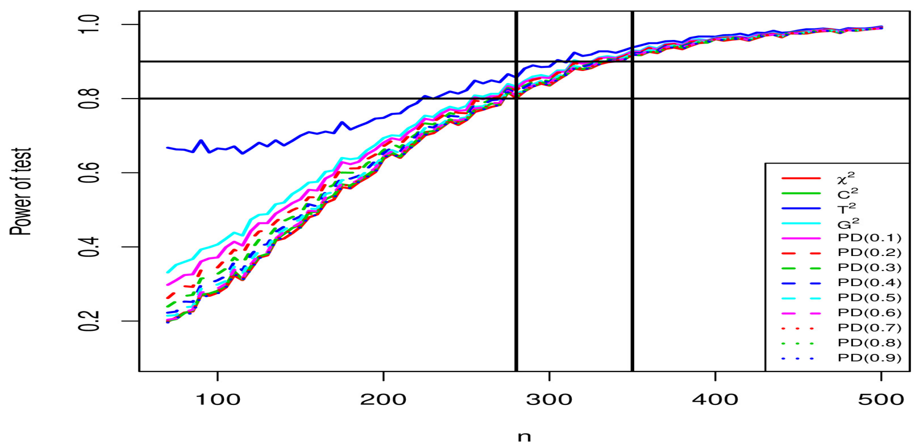


**Figure 11:** Power results under R=5 and C=5.

Table 8 shows the minimum required sample sizes for the contingency tables to reach the minimum 0.80 and 0.90 power values. As seen these results, the required sample size is an increasing function of the dimension of the contingency table. Therefore, higher dimension needs more sample size. The determined effect sizes for each table dimension are reported in Table 7.

| Table dimensions | Power | |
|---|---|---|
| | 0.8 | 0.9 |
| 3x3 | 60 | 75 |
| 3x4 | 65 | 80 |
| 3x5 | 70 | 90 |
| 4x4 | 130 | 150 |
| 4x5 | 160 | 195 |
| 5x5 | 280 | 350 |

**Table 8:** Minimum required sample sizes for the powers 0.8 and 0.9

As given in Section 1, the sample size is function of type-I error, power, df and effect size (see, Section 5). The powers are calculated by considering the different values of the effect size, df and sample sizes for the fixed type-I error 0.05. The results are given in Table 9. From these results, it is seen that when the effect size is low, the required sample size should be large to obtain the high power. Also, when the df is high, the sample size should be large to obtain the high power. Under these results, if Table 8 is revisited, the sample sizes given in this table are determined based on the high effect sizes.

| Sample size (df=4) | w=0.05 | w=0.15 | w=0.30 | w=0.50 | Sample size (df=6) | w=0.05 | w=0.15 | w=0.30 | w=0.50 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.056 | 0.113 | 0.358 | 0.820 | 50 | 0.055 | 0.100 | 0.303 | 0.758 |
| 100 | 0.063 | 0.189 | 0.663 | 0.989 | 100 | 0.060 | 0.161 | 0.589 | 0.980 |
| 150 | 0.069 | 0.272 | 0.852 | 1.000 | 150 | 0.065 | 0.229 | 0.796 | 0.999 |
| 200 | 0.076 | 0.358 | 0.943 | 1.000 | 200 | 0.071 | 0.303 | 0.911 | 1.000 |
| 250 | 0.083 | 0.443 | 0.980 | 1.000 | 250 | 0.076 | 0.378 | 0.965 | 1.000 |
| 500 | 0.121 | 0.773 | 1.000 | 1.000 | 500 | 0.106 | 0.705 | 1.000 | 1.000 |
| Sample size (df=8) | w=0.05 | w=0.15 | w=0.30 | w=0.50 | Sample size (df=9) | w=0.05 | w=0.15 | w=0.30 | w=0.50 |
| 50 | 0.054 | 0.092 | 0.267 | 0.706 | 50 | 0.054 | 0.089 | 0.253 | 0.683 |
| 100 | 0.058 | 0.143 | 0.534 | 0.968 | 100 | 0.058 | 0.137 | 0.510 | 0.962 |
| 150 | 0.063 | 0.202 | 0.747 | 0.998 | 150 | 0.062 | 0.192 | 0.725 | 0.998 |
| 200 | 0.067 | 0.267 | 0.879 | 1.000 | 200 | 0.066 | 0.253 | 0.863 | 1.000 |
| 250 | 0.072 | 0.334 | 0.948 | 1.000 | 250 | 0.070 | 0.317 | 0.939 | 1.000 |
| 500 | 0.097 | 0.650 | 1.000 | 1.000 | 500 | 0.094 | 0.626 | 1.000 | 1.000 |

**Table 9:** The calculated powers for the different values of the effect size, df and sample sizes

## 5. Comparison of PD(0.4) and Fisher-Freeman-Halton exact test statistics

It is well-known that the Fisher exact test is used for R=2 and C=2 contingency tables when more than 20% of cells have expected frequencies less than 5. However, when the table dimension is larger than 2×2, the FH test is used [22]. In this section, we compare the empirical type-I error rates of the PD(0.4) and the FH test statistics based on the simulation study. The same probability matrices given in Section 4 are used. The type-I errors of the PD(0.4) and FH test statistics are reported graphically in Figures 12, 13 and 14. As seen from these figures, it is observed that the PD(0.4) and FH produce similar results in terms of their type-I error rates. Both test statistics can be used for sparse and non-sparse contingency tables. The obtained type-I errors of the PD(0.4) and FH test statistics are below the desired value, 0.05. Also, the empirical power values of the PD(0.4) and FH test statistics are reported in Figures 15, 16 and 17. PD(0.4) and FH test statistics produce similar results for their power values, as in type-I error rates.

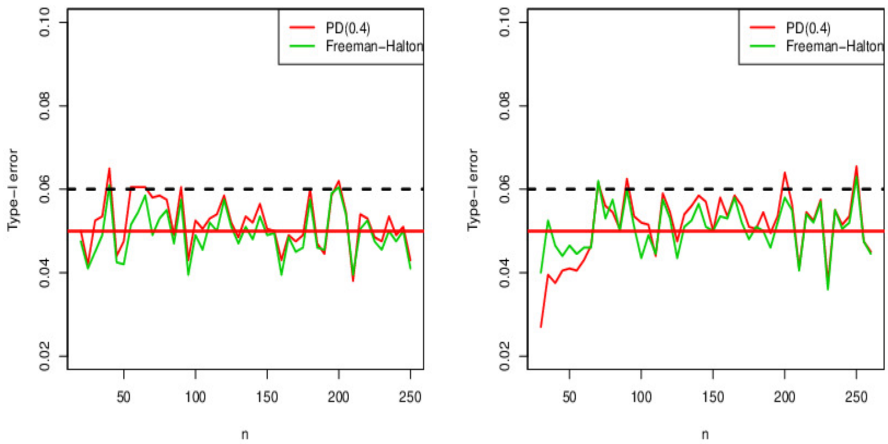


**Figure 12:** Type-I errors of the PD(0.4) and FH test for R=3 and C=3(left) and R=3 and C=4 (right)

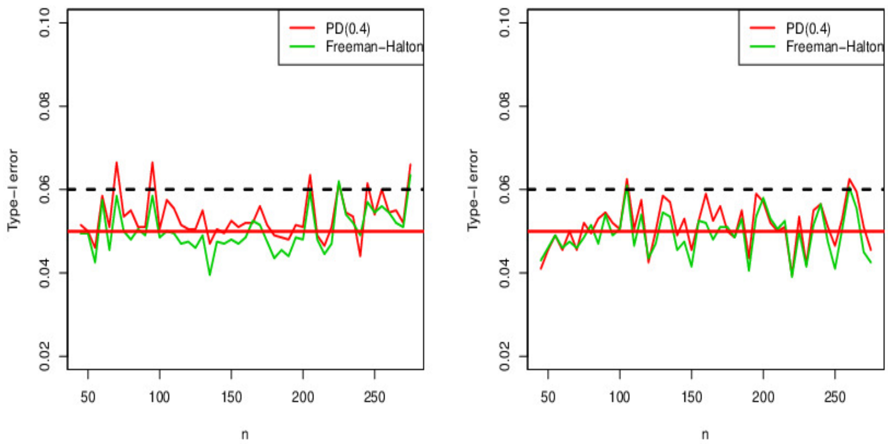


**Figure 13:** Type-I errors of the PD(0.4) and FH test for R=3 and C=5(left) and R=4 and C=4 (right)

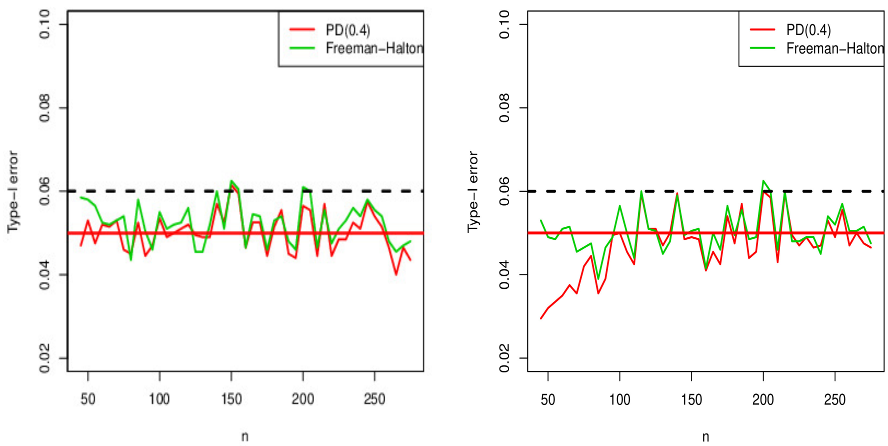


**Figure 14:** Type-I errors of the PD(0.4) and FH test for R=4 and C=5(left) and R=5 and C=5 (right)

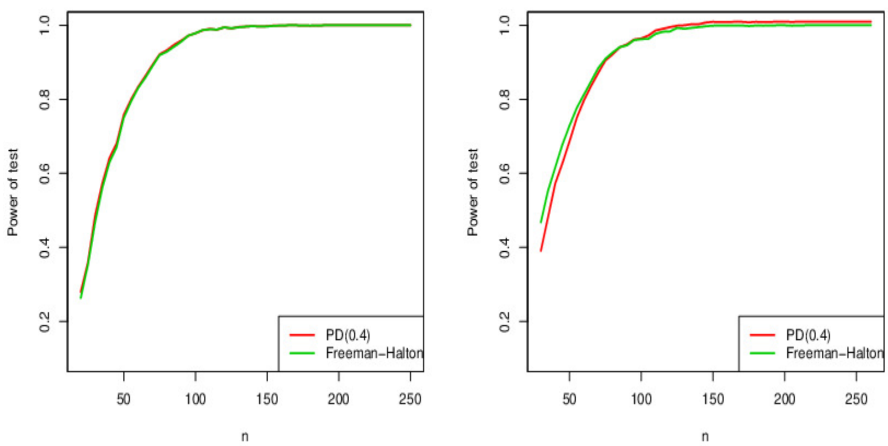


**Figure 15:** Power values of the PD(0.4) and FH test for R=3 and C=3(left) and R=3 and C=4 (right)

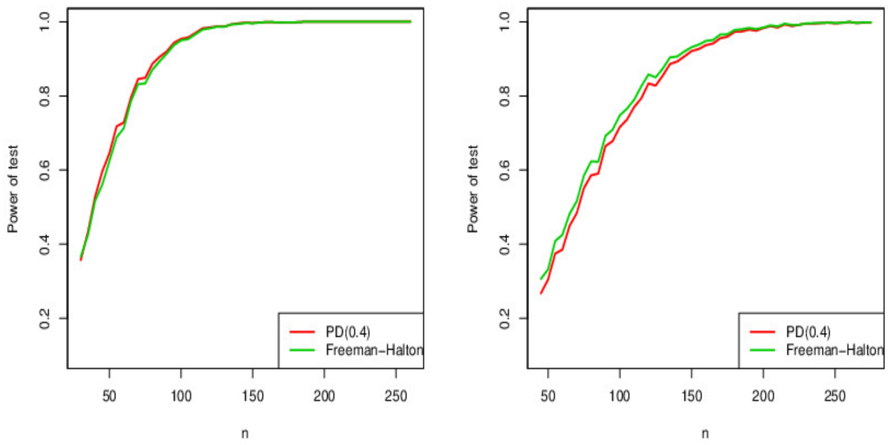


**Figure 16:** Power values of the PD(0.4) and FH test for R=3 and C=5(left) and R=4 and C=4 (right)

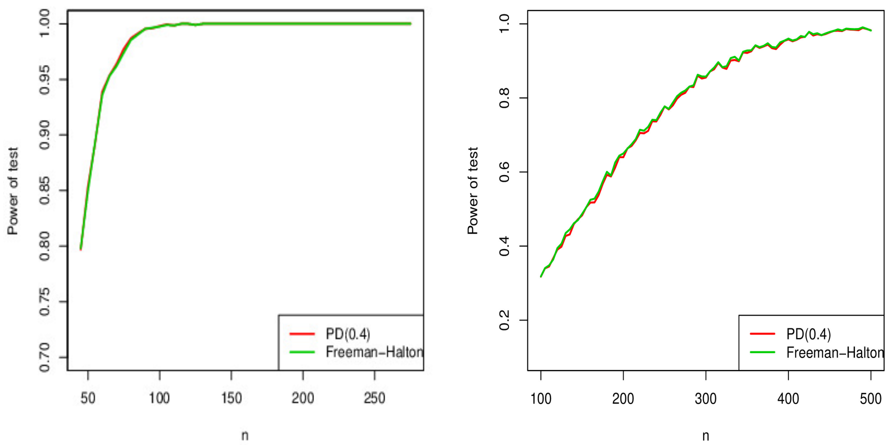


**Figure 17:** Power values of the PD(0.4) and FH test for R=4 and C=5(left) and R=5 and C=5 (right)

## 6. Power comparison of the PD(0.4) and $\chi^2$ test statistics via real data application

The sample size determination is an important step of any field work. Before collecting the data, the researcher should know how many observations is needed to reach the desired power value. The sample size is a function of three parameters. These are effect size, type-I error and power to detect $H_1$ hypothesis.

Let $\mathrm{PD}\,(0.4)_c$ be the calculated value of the $\mathrm{PD}\,(0.4)$ test statistic which is calculated by

$$\mathrm{PD}\,(0.4)_c = \frac{2}{0.4\,(0.4+1)} \sum_{i=1}^{R} \sum_{j=1}^{C} \left[ \left( \frac{n_{ij}}{e_{ij}} \right)^{0.4} - 1 \right],$$

where $n_{ij}$ and $e_{ij}$ are the observed and expected frequencies, respectively . When the null hypothesis ($H_0$) is true, the test statistic is distributed as $\chi^2$ distribution with $(R-1)(C-1)$ df. The null hypothesis is rejected when $\mathrm{PD}\,(0.4)_c > \chi^2_{(R-1)(C-1),\alpha}$ where $\alpha$ is the significance level which is called as type-I error. When the null hypothesis is not true, the distribution of $\mathrm{PD}\,(0.4)_c$ follows the non-central $\chi^2$ distribution with non-centrality parameter $\lambda$ and df $(R-1)(C-1)$. The non-centrality parameter $\lambda$ is a function of $n$ and effect size $w$. We have the following equation to calculate the parameter $\lambda$ (see [3])

$$\lambda = nw^2. \tag{6.1}$$

The effect size is calculated by $w = \sqrt{\mathrm{PD}\,(0.4)_c / n}$. So, replacing $w$ in (6.1), we have $\lambda = \mathrm{PD}\,(0.4)_c$. So, the power of the $\mathrm{PD}\,(0.4)_c$ test statistic can be obtained by

$$\text{Power} = 1 - \Pr\left(\chi^2_{(R-1)(C-1),\lambda}\left(\text{PD}\left(0.4\right)_c\right) < \chi^2_{(R-1)(C-1),\alpha}\right). \tag{6.2}$$

The power of the $\chi^2$ test statistic can be easily computed by changing the $\text{PD}\left(0.4\right)_c$ in (6.2) with the test statistic value of the $\chi^2$. In the remaining part of these section, we analyze two data sets to compare the PD(0.4) with $\chi^2$ test statistics. Note that the calculated power values in the remaining part of this section are empirical powers.

### 6.1. Pneumonia data

To compare the power value of the PD(0.4) and $\chi^2$ test statistics, we use the data set on the vaccination program for the pneumonia patients. The data can be found in the work of [23]. Also, the data set is given in Table 10. Here, the research question is that *Does the vaccine protect the individuals from the pneumococcal pneumonia disease?*.

| Health outcome | Unvaccinated | Vaccinated |
|---|---|---|
| Sick with pneumococcal pneumonia | 23 | 5 |
| Sick with non-pneumococcal pneumonia | 8 | 10 |
| No pneumonia | 61 | 77 |

**Table 10:** The data set for vaccination program

The data is analyzed using the PD(0.4) and $\chi^2$ test statistics. Obtained results are given in Table 11. The significance level $\alpha$ is selected 0.05 for both test statistics. According to the Table 11, both of the test statistics reject the null hypothesis. However, the power value of the PD(0.4) test statistic is higher than the $\chi^2$ test statistic. So, we recommend the usage of the PD(0.4) test statistic to obtain higher power value than those of the $\chi^2$ test statistic.

| Test statistics | Value | df | p-value | Power |
|---|---|---|---|---|
| $\chi^2$ | 13.649 | 2 | 0.001 | 0.921 |
| PD(0.4) | 14.095 | 2 | $< 0.001$ | 0.930 |

**Table 11:** Results of the test statistics for the pneumonia data

### 6.2. Epidemiological data

The second data is on the obesity risk of children based on their race. The data set can be found in [24]. Here, the research question is that *Does the obesity risk differ by the race?*. To answer this question, we analyze the data set given in Table 12 with PD(0.4) and $\chi^2$ test statistics.

| Risk | Black | White | Others |
|---|---|---|---|
| At risk | 185 | 140 | 90 |
| Not at risk | 80 | 17 | 23 |

**Table 12:** Epidemiological data for the children

The obtained results are given in Table 13. Based on the results in Table 13, since the power of PD(0.4) is higher than the $\chi^2$, we recommend the PD(0.4) test statistic to analyze the current data set.

| Test statistics | Value | df | p-value | Power |
|---|---|---|---|---|
| $\chi^2$ | 21.595 | 2 | $< 0.001$ | 0.991 |
| PD(0.4) | 22.386 | 2 | $< 0.001$ | 0.992 |

**Table 13:** Results of the test statistics for the epidemiological data

## 7. Conclusion

We compare the various members of the PD family as well as different values of $\lambda$ using the extensive simulation study based on the different settings such as dimensions of the contingency tables, type-I error, sample sizes and powers. When the parameter $\lambda = 0.4$, the test statistic reaches the maximum value of the power. Also, we compare the PD(0.4) test statistic with $\chi^2$ test statistics based on the power values. Two applications to the real datasets show that PD(0.4) provides higher powers than the $\chi^2$ test statistic. As a future work, we plan to develop the web-tool to calculate the required sample size and displays the results of the PD(0.4) test statistic.

## Declarations

**Acknowledgements:** The authors are grateful to the anonymous referee for helpful suggestions to improve the paper.

**Conflict of Interest Disclosure:** The authors declare no conflict of interest.

**Copyright Statement:** Authors own the copyright of their work published in the journal and their work is published under the CC BY-NC 4.0 license.

**Supporting/Supporting Organizations:** This research received no external funding.

**Ethical Approval and Participant Consent:** This article does not contain any studies with human or animal subjects. It is declared that during the preparation process of this study, scientific and ethical principles were followed and all the studies benefited from are stated in the bibliography.

**Plagiarism Statement:** This article was scanned by the plagiarism program. No plagiarism detected.

**Availability of Data and Materials:** Data sharing not applicable.

**Use of AI tools:** The author declares that he has not used Artificial Intelligence (AI) tools in the creation of this article.

### ORCID

Gökçen Altun https://orcid.org/0000-0003-4311-6508

## References

[1] K. Pearson, *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*, Lond. Edinb. Dubl., **50**(302) (1900), 157-175. [CrossRef]

[2] T.M. Franke, T. Ho and C.A. Christie, *The chi-square test: Often used and more often misinterpreted.* Am. J. Evaluation, **33**(3)(2012), 448-458. [CrossRef] [Scopus] [Web of Science]

[3] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, England, Routledge, (1988). [CrossRef] [Web of Science]

[4] K.R. Murphy and B. Myors, *Testing the hypothesis that treatments have negligible effects: Minimum-effect tests in the general linear model*, J. Appl. Psychol., **84**(2)(1999), 234. [CrossRef] [Scopus] [Web of Science]

[5] G.M. Oyeyemi, A.A. Adewara, F.B. Adebola and S.I. Salau, *On the estimation of power and sample size in test of independence*, Asian J. Math. Stat., **3**(3) (2010), 139-146. [CrossRef]

[6] A. Agresti, *Categorical Data Analysis* John Wiley & Sons, **482**(2002). [CrossRef]

[7] A. Agresti, *An Introduction to Categorical Data Analysis*, John Wiley & Sons (2007). [CrossRef]

[8] P. Burman, *On some testing problems for sparse contingency tables*. J. Multivariate Anal., **88**(1)(2004), 1-18. [Scopus] [Web of Science]

[9] J.K. Yarnold, *The minimum expectation in $\chi^2$ goodness of fit tests and the accuracy of approximations for the null distribution*, J. Amer. Statist. Assoc., **65**(330) (1970), 864-886. [CrossRef] [Scopus]

[10] T. Rudas, *A Monte Carlo comparison of the small sample behaviour of the Pearson, the likelihood ratio and the Cressie-Read statistics*, J. Stat. Comput. Simul., **24**(2)(1986), 107-120. [CrossRef] [Scopus] [Web of Science]

[11] N. Cressie and T.R. Read, *Multinomial goodness-of-fit tests*, J. R. Stat. Soc. Ser. B. Stat. Methodol., **46**(3)(1984), 440-464. [CrossRef] [Web of Science]

[12] S.E. Fienberg, *The use of chi-squared statistics for categorical data problems*, J. R. Stat. Soc. Ser. B. Stat. Methodol., **41**(1)(1979), 54-64. [CrossRef]

[13] K. Larntz, *Small-sample comparisons of exact levels for chi-squared goodness-of-fit statistics*, J. Amer. Statist. Assoc., **73**(362)(1978), 253-263. [CrossRef] [Scopus]

[14] M.A. Garcia-Perez and V. Nunez-Anton, *Accuracy of power-divergence statistics for testing independence and homogeneity in two-way contingency tables*, Commun Stat-Simul C., **38**(3)(2009), 503-512. [CrossRef] [Scopus] [Web of Science]

[15] S. Aktaş, *Power divergence statistics under quasi independence model for square contingency tables*, Sains Malays., **45**(10)(2016), 1573-1578. [Web]

[16] G. Altun, *A study on Freeman-Tukey test statistic under the symmetry model for square contingency tables*, Cumhuriyet Sci J., **42**(2)(2021), 413-421. [CrossRef]

[17] S. D. Horn, *Goodness-of-fit tests for discrete data: a review and an application to a health impairment scale*, Biometrics, **33**(1)(1977), 237-247. [CrossRef] [Scopus]

[18] G.S. Watson, *Some recent results in chi-square goodness-of-fit tests*, Biometrics, **15**(3) (1959), 440-468. [CrossRef]

[19] M. Kendall and A. Stuart, *Handbook of Statistics*, (1979).

[20] K. Larntz, *Small Sample Comparison of Likelihood-Ratio and Pearson Chi-Square Statistics for the Null Distribution*, University of Minnesota, (1973). [CrossRef]

[21] R. Durrett, *Probability: Theory and Examples*, Cambridge university press, **49**(2019). [Scopus]

[22] G.H. Freeman and J.H. Halton, *Note on an exact treatment of contingency, goodness of fit and other problems of significance*, Biometrika, **38**(1/2)(1951), 141-149. [CrossRef]

[23] M. L. McHugh, *The chi-square test of independence*, Biochem. Med., **23**(2)(2013), 143-149. [CrossRef] [Scopus] [Web of Science]

[24] P.R. Nelson, P.S. Wludyka and K.A. Copeland, *The Analysis of Means: A Graphical Method for Comparing Means, Rates, and Proportions*, SIAM, (2005). [CrossRef] [Web of Science]

All open access articles published are distributed under the terms of the CC BY-NC 4.0 license (Creative Commons Attribution-Non-Commercial 4.0 International Public License as currently displayed at http://creativecommons.org/licenses/by-nc/4.0/legalcode) which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original work is properly cited.

**How to cite this article:** G. Altun, *The most powerful member of the power-divergence family for the independence model in contingency tables*, Fundam. J. Math. Appl., **7**(3) (2024), 169-185. DOI 10.33401/fujma.1432348