Turk J
Field Crops
2016, 21(2), 286-297
DOI: 10.17557/tjfc.33765

THE COCHRAN-ARMITAGE TEST TO ESTIMATE THE SAMPLE SIZE FOR TREND OF PROPORTIONS FOR BIOLOGICAL DATA

Mustafa Agah TEKINDAL[1], Ozlem GULLU[2], Ayse Canan YAZICI[1] and Yasemin YAVUZ[3]*

[1]İzmir University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, İzmir, TURKEY
[2]Ankara University, Faculty of Sciences, Department of Statistics, Ankara, TURKEY
[3]Ankara University, Faculty of Medicine, Department of Biostatistics, Ankara, TURKEY
**Corresponding author: matekindal@gmail.com*

Received: 13.01.2016

ABSTRACT

The biological activity of a substance can be investigated through a series of experiments done with the increased or decreased dosage of it. One of the purposes of such studies is to determine the trend of responses based on dosage. In studies carried out for this purpose, appropriate sample size has an indisputable influence on the reliability of the decisions to be made at the end of the study. There are various statistical methods for determining the trend of proportions. One of them is the Cochran-Armitage test. In a categorical data analysis, the trend between two variables with k categories can be determined through the Cochran-Armitage test. This study aims to explore the sample size calculation method developed by Nam J. (1987) for the Cochran-Armitage test. The power of the test was investigated in different numbers of categories and in different sample sizes for each category when the least biologically significant differences changed as Type I error was taken as 0.05. To this end, the study examined the results obtained by making 10000 repetitions for each case through the Monte Carlo simulation method. When the least biologically significant differences change at the end of simulation studies, the power of test highly varies in different combinations. When the number of categories is 2, determination of trend requires working with very large samples. When the number of categories is 3 or 4, the desired power can be obtained with smaller samples compared to the case where the number of categories is 2. When the number of categories is over 4, a substantial increase is needed in sample size to obtain the desired power. Change in marginal frequencies does not have much influence on sample size.

Keywords: cochran-armitage, exact test, sample size, trend of proportions, type I error rate

INTRODUCTION

In biological research, 2xC size contingency tables are frequently used for the analysis of ordered categorical data. Here, there are C ordered groups in return for the binary response variable (C: amount). The Cochran-Armitage test is frequently used for calculating the trend of binomial proportions (Cochran, 1954; Armitage, 1955; Lachin, 2011). This test is widely used in epidemiological and genetic research, in biomedical studies focusing on dosage-response relationship, in cancer studies, and in toxicological risk assessment. Studies involving this test are also found in agricultural and veterinary studies (Ahn et al., 2007; Zheng and Gastwirth, 2006).

In the period when the U.S. Environmental Protection Agency banned the use of diazinon, Banks et al. (2005) analyzed whether or not the samples collected from the rivers in the rural and urban areas of Denton city in the U.S. state of Texas involved diazinon. They collected a total of 1243 samples between 2001 and 2004. For analysis of variance (ANOVA) and analysis of the categorical data, Mantel–Haenszel Chi-Square test and Cochran–Armitage Trend test were used for "Sx2" contingency tables. According to the obtained results, decrease in diazinon concentration having a value over the determined lower limit was statistically significant between 2001 and 2004 (Mantel–Haenszel Chi-Square test, $p<0.0001$, $n=1243$). The four-year data also indicated that a significant decrease occurred (Cochran–Armitage trend test, $z= 17.94$, $p<0.0001$, $n=1243$). It was concluded that substantial decrease in non-agricultural diazinon use considerably reduced the formation of pesticides in surface waters.

Shen et al. (2014) investigated the influence of sex on the relationship between cardiovascular risk factors and gallstone disease in the Taiwanese population engaging in agriculture and fishing. The research sample consisted of 6511 participants (3971 males and 2540 females) applying to a training and research hospital in 2010 on a voluntary basis. While the risk factors influential on gallstone disease were analyzed through multiple logistic regression, categorical data were analyzed through Chi-Square and the Cochran-Armitage trend test. According to the obtained results, the females, compared to the males, (Chi-Square test p<0.003); those at the age of 85 or over, compared to those at the age of 60 to 64, (Cochran-Armitage trend test, p<0.0001); and those having a metabolic syndrome, compared to those not having a metabolic syndrome, were seen to have a higher risk of having gallstone disease. Thus, sex and age were determined to be significantly influential on this disease.

Mehta et al. (1998) calculated the exact power and the asymptotic power of the Cochran-Armitage test based on three examples through the method proposed by Nam (1987). Although such samples were hypothetical, they were motivated by realistic study designs. Each sample dose-response relationship was modeled via logistic regression method. The studies were characterized by an unevenly spaced dosing and a small sample size or a large sample size and a low response rate. The first example is a biological study indicating the dose-response relationship of the patients having an advanced chronic disease with a dosage of 1. At low dosage level, the probability of response is 0.001. It is considered that if a daily increase of 0.5 unit in dosage brings about an increase also in response, the medicine can be beneficial. The study was designed in such a way that maximum dosage would be 16 units. Thus, dosage was suggested as 1, 2, 4, 8, and 16 for the sample sizes of 10, 10, 10, 5, and 2 respectively. At 2.5% significance level, asymptotic power estimation was calculated to be 81% for the one-sided Cochran-Armitage trend test. The second example is the long-term follow-up of the subjects exposed to a low dosage of radiation in Japan. The cohort was divided into four dosage groups based on the average radiation undergone: 0, 5, 30, and 75 rad. Each dosage group consisted of 2500, 3600, 1450, and 410 subjects respectively. The one-sided Cochran-Armitage trend test was carried out at 5% significance level. When the response rate of 1 in 10,000 was taken into consideration, it was intended to estimate the power of the test with a logistic scale of 0.049. In the end, the asymptotic power of the test was found to be 75%. The third example is an animal toxicity study conducted by FDA (Food and Drug Administration). A carcinogen dosage of 0, 1, 5, and 50 units was given to the animals respectively. The presence or absence of a particular tumor type was observed in those animals. Each dosing group consisted of 50 subjects. Considering the response rate of 1 in 10,000, an increase of 0.13 occurred in return for each increase of one unit in the dosage. The asymptotic power of the one-sided Cochran-Armitage trend test carried out at 5% significance level was found to be 80%.

The Cochran-Armitage trend test has an asymptotic approach and thus shows a poor performance in very small and unbalanced samples. This test is not recommended to be used for the variables at classification level (Kang, S. H., and Lee, J. W., 2007)

The power of this test is a function of the sample sizes of the groups and the positive rates between the groups and is also defined as the test statistic of non-central distribution. Chapman and Nam (1968) state that the test may be similar, but not equal, to non-central distribution. Nam (1987) defines the power of the Cochran-Armitage test as the linear trend of logit probabilities. Recently, Slager and Schaid (2001) have stated that the power of the Cochran-Armitage test, under the null hypothesis and alternative hypotheses, can correspond to the estimated variance of the trend and Z-test. Among nonparametric measures, Top ,S_i and Rank - sum statistics of nonparametric procedures were found to be useful in detecting the stability of the genotypes Mortazavian S. M. M., Azizi-Nia S. (2014). Since chemical experiments are expensive and time-consuming, it is aimed to determine the most appropriate conditions by acquiring data and doing modeling through predetermined variables and points Tekindal et al. (2014).

The aim of the present study is to explore the sample size calculation method developed by Nam J. (1987) for the Cochran-Armitage test. The power of the test was investigated in different numbers of categories and in different sample sizes for each category when the least biologically significant differences changed as Type I error was taken as 0.05.

MATERIALS AND METHODS

The Cochran – Armitage Test

The Cochran-Armitage test for the trend of proportions is used for testing a linear trend in proportions with its ordinal or quantitative metric or assignable scores over independent groups in C categories. For example, C groups can be ordered as normal, moderately normal, and abnormal, and 1, 2, and 3 can be assigned to them respectively as scores. Likewise, in the study, ordinal numbers can have a real number value or a score value as the daily dosage of a substance.

The sample size and the power of the Cochran – Armitage test are calculated based on the results of Nam (1987). There are asymptotic power and exact power calculations for non-corrected tests and correction for continuity tests. As *X* was a covariate (or doses), it was assumed to follow a linear trend with a logistic measure, and random samples were extracted from k different populations.

It is assumed that there are *k* random binomial variables y_i, refers to factor or dose levels; x_i , n_i refers to sample sizes and p_i refers to the probability of success. For $i = 1, 2, ..., k$ and when $x_1 < x_2 < ... < x_k$, it is defined as follows:

$$N = \sum_{i=1}^{k} n_i \qquad \bar{p} = \frac{1}{N} \sum_{i=1}^{k} y_i \tag{1}$$

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{k} n_i x_i \quad \bar{q} = 1 - \bar{p} \tag{2}$$

Equation (3) is used if it is assumed that success rates follow a linear trend with a logistic scale.

$$p_i = \frac{\exp(\alpha + \beta x_i)}{1 + \exp(\alpha + \beta x_i)} \tag{3}$$

The One-Sided Testing of the Linear Trend Increasing in Proportions

Correction for Continuity Test

Nam (1987) recommended the following asymptotic correction for continuity test statistic for the one-sided testing of the linear trend increasing in proportions.

$$Z_{c.c.} = \frac{\sum_{i=1}^{k} y_i (x_i - \bar{x}) - \frac{\Delta}{2}}{\sqrt{\bar{p}\,\bar{q}\left[\sum_{i=1}^{k} n_i (x_i - \bar{x})^2\right]}} \tag{4}$$

Correction factor for continuity is $\frac{\Delta}{2}$. If the covariate is x_i, it can consist of even spaces or

$$\Delta = x_{i+1} - x_i \quad i < k \tag{5}$$

successive spaces. For unevenly spaced covariates, calculates Δ as follows:

$$\Delta = \frac{1}{k-1}\sum_{i=1}^{k-1} (x_{i+1} - x_i) \tag{6}$$

Continuity for correction was not recommended for the covariates divided into uneven spaces. Thus, non-corrected test must be used in case of the presence of factors divided into unequal spaces.

Non-Corrected Test

The non-corrected test statistic is the $\Delta = 0$ version of the corrected test statistic.

$$Z = \frac{\sum_{i=1}^{k} y_i (x_i - \bar{x})}{\sqrt{\bar{p}\,\bar{q}\left[\sum_{i=1}^{k} n_i (x_i - \bar{x})^2\right]}} \tag{7}$$

The One-Sided Testing of the Linear Trend Decreasing in Proportions

Correction for Continuity Test

Nam (1987) found the asymptotic correction for continuity test statistic for the one-sided testing of the linear trend increasing in proportions. The correction for continuity test is calculated for the one-sided testing of the linear trend decreasing in proportions in the same way. However, here, $\frac{\Delta}{2}$ factor is added, but not subtracted.

$$Z_{c.c.} = \frac{\sum_{i=1}^{k} y_i (x_i - \bar{x}) + \frac{\Delta}{2}}{\sqrt{\bar{p}\,\bar{q}\left[\sum_{i=1}^{k} n_i (x_i - \bar{x})^2\right]}} \tag{8}$$

Non-Corrected Test

The non-corrected test statistic is the $\Delta = 0$ version of the corrected test statistic.

$$Z = \frac{\sum_{i=1}^{k} y_i (x_i - \bar{x})}{\sqrt{\bar{p}\,\bar{q}\left[\sum_{i=1}^{k} n_i (x_i - \bar{x})^2\right]}} \tag{9}$$

The Cochran-Armitage Trend Test Approximate Power Calculation

Power for the One-Sided Testing of the Linear Trend Increasing in Proportions

The critical value $z_{critical}$ is found by use of the standard normal distribution. For the one-sided testing of the alternative hypothesis, p_i is a monotonically increasing function of x_i. For $p = (p_1, p_2, ..., p_k)$, power is calculated according to Equation (10).

$$\begin{aligned} 1 - \beta &= \Pr(z \geq z_{critical} \mid H_1) \\ &= 1 - \Phi(u_U) \end{aligned} \tag{10}$$

In the Equation (10), "Φ" indicates cumulative normal distribution.

$$u_U = \frac{-\left[\sum_{i=1}^{k} n_i p_i (x_i - \bar{x}) - \frac{\Delta}{2}\right] + z_{critical}\sqrt{p(1-p)\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}}{\sqrt{\sum_{i=1}^{k} n_i p_i (1-p_i)(x_i - \bar{x})^2}} \tag{11}$$

$$p = \frac{1}{N}\sum_{i=1}^{k} n_i p_i \tag{12}$$

While power is being calculated for the non-corrected test, it is assumed that $\Delta = 0$.

Power for the One-Sided Testing of the Linear Trend Decreasing in Proportions

The critical value $z_{critical}$ is founded by use of the standard normal distribution. For the one-sided testing of the alternative hypothesis, p_i is a monotonically decreasing function of x_i. For $p = (p_1, p_2, ..., p_k)$, power is calculated according to Equation (13).

$$1-\beta = \Pr(z \le -z_{critical} \mid H_1) = 1 - \Phi(u_L) \tag{13}$$

Here, "Φ" is cumulative normal distribution.

$$u_U = \frac{-\left[\sum_{i=1}^{k} n_i p_i (x_i - \bar{x}) + \frac{\Delta}{2}\right] + z_{critical}\sqrt{p(1-p)\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}}{\sqrt{\sum_{i=1}^{k} n_i p_i (1-p_i)(x_i - \bar{x})^2}} \tag{14}$$

$$p = \frac{1}{N}\sum_{i=1}^{k} n_i p_i \tag{15}$$

While power is being calculated for the non-corrected test, it is assumed that $\Delta = 0$.

Scenarios 2x2, 2x3, 2x4, 2x5, 2x6, 2x7, 2x8, 2x9 and 2x10 in table 0.4; 0.5 and 0.6 were determined at increasing proportions. For this purpose, using Monte Carlo simulation method, the results obtained by each case was evaluated again to 10000. Simulations, PASS (Version 11) are made in the program Hintze, J. (2011)

RESULTS AND DISCUSSION

Table 1. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 2 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50	100	0.12339	0.05
0.40;0.50	200	0.24712	0.05
0.40;0.50	300	0.36814	0.05
0.40;0.50	400	0.47982	0.05
0.40;0.50	500	0.57867	0.05
0.40;0.50	600	0.66347	0.05
0.40;0.50	700	0.73445	0.05
0.40;0.50	800	0.79269	0.05
0.40;0.50	900	0.83969	0.05
0.40;0.50	1000	0.87709	0.05

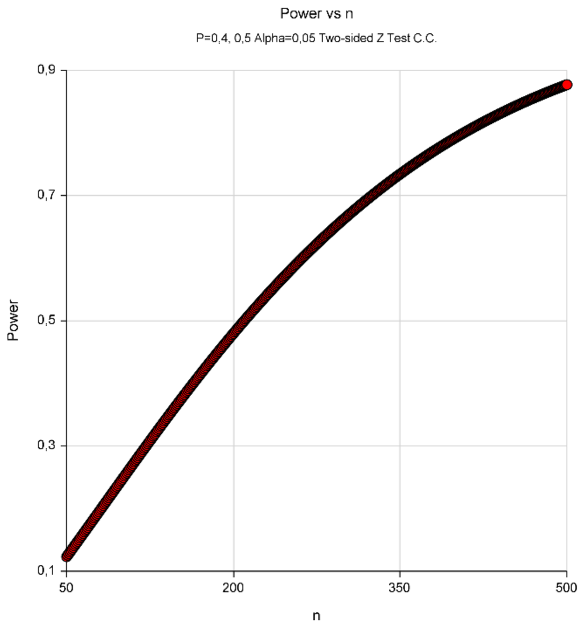

Figure 1. The Cochran-Armitage test power function graph for the 2 x 2 crosstab

Table 2. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 3 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60	150	0.47562	0.05
0.40;0.50;0.60	252	0.71452	0.05
0.40;0.50;0.60	351	0.85438	0.05
0.40;0.50;0.60	450	0.93006	0.05
0.40;0.50;0.60	552	0.96871	0.05
0.40;0.50;0.60	651	0.98621	0.05
0.40;0.50;0.60	750	0.99410	0.05
0.40;0.50;0.60	852	0.99761	0.05
0.40;0.50;0.60	951	0.99903	0.05
0.40;0.50;0.60	1050	0.99961	0.05
0.40;0.50;0.60	1152	0.99985	0.05
0.40;0.50;0.60	1251	0.99994	0.05
0.40;0.50;0.60	1350	0.99998	0.05
0.40;0.50;0.60	1452	0.99999	0.05
0.40;0.50;0.60	1500	0.99999	0.05

Table 3. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 4 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60	200	0.57794	0.05
0.40;0.50;0.60;0.60	300	0.76290	0.05
0.40;0.50;0.60;0.60	400	0.87500	0.05
0.40;0.50;0.60;0.60	500	0.93722	0.05
0.40;0.50;0.60;0.60	600	0.96967	0.05
0.40;0.50;0.60;0.60	700	0.98581	0.05
0.40;0.50;0.60;0.60	800	0.99354	0.05
0.40;0.50;0.60;0.60	900	0.99712	0.05
0.40;0.50;0.60;0.60	1000	0.99874	0.05
0.40;0.50;0.60;0.60	1100	0.99946	0.05
0.40;0.50;0.60;0.60	1200	0.99977	0.05
0.40;0.50;0.60;0.60	1300	0.99991	0.05
0.40;0.50;0.60;0.60	1400	0.99996	0.05
0.40;0.50;0.60;0.60	1500	0.99998	0.05
0.40;0.50;0.60;0.60	1600	0.99999	0.05
0.40;0.50;0.60;0.60	1700	1.00000	0.05
0.40;0.50;0.60;0.60	1800	1.00000	0.05
0.40;0.50;0.60;0.60	1900	1.00000	0.05
0.40;0.50;0.60;0.60	2000	1.00000	0.05

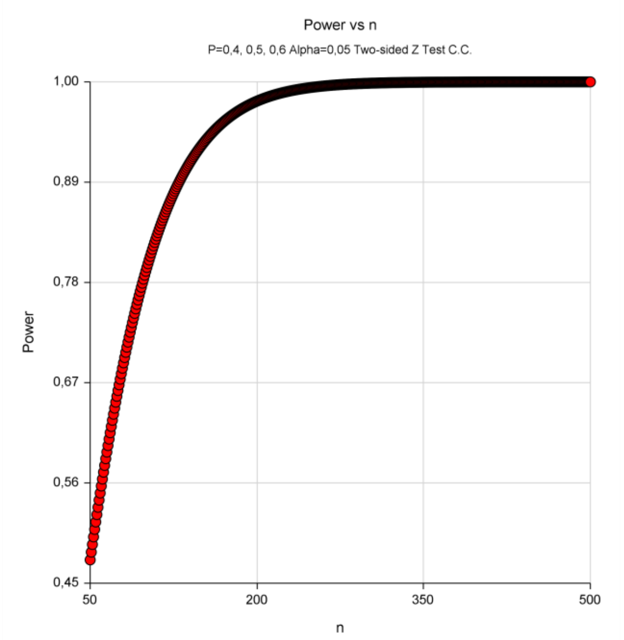

Figure 2. The Cochran-Armitage test power function graph for the 2 x 3 crosstab

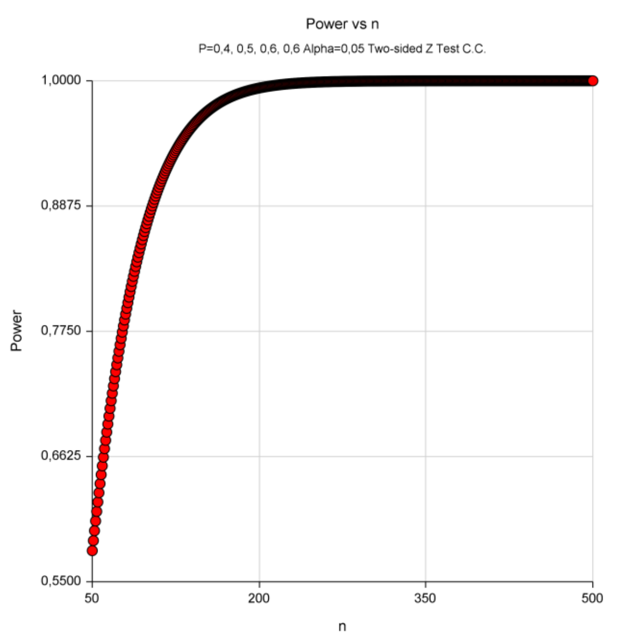

Figure 3. The Cochran-Armitage test power function graph for the 2 x 4 crosstab

Table 4. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 5 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60;0.60	100	0.26738	0.05
0.40;0.50;0.60;0.60;0.60	200	0.49853	0.05
0.40;0.50;0.60;0.60;0.60	300	0.67846	0.05
0.40;0.50;0.60;0.60;0.60	400	0.80370	0.05
0.40;0.50;0.60;0.60;0.60	500	0.88469	0.05
0.40;0.50;0.60;0.60;0.60	600	0.93435	0.05
0.40;0.50;0.60;0.60;0.60	700	0.96358	0.05
0.40;0.50;0.60;0.60;0.60	800	0.98024	0.05
0.40;0.50;0.60;0.60;0.60	900	0.98948	0.05
0.40;0.50;0.60;0.60;0.60	1000	0.99449	0.05
0.40;0.50;0.60;0.60;0.60	1100	0.99716	0.05
0.40;0.50;0.60;0.60;0.60	1200	0.99855	0.05
0.40;0.50;0.60;0.60;0.60	1300	0.99927	0.05
0.40;0.50;0.60;0.60;0.60	1400	0.99964	0.05
0.40;0.50;0.60;0.60;0.60	1500	0.99982	0.05
0.40;0.50;0.60;0.60;0.60	1600	0.99991	0.05
0.40;0.50;0.60;0.60;0.60	1700	0.99996	0.05
0.40;0.50;0.60;0.60;0.60	1800	0.99998	0.05
0.40;0.50;0.60;0.60;0.60	1900	0.99999	0.05
0.40;0.50;0.60;0.60;0.60	2000	1.00000	0.05

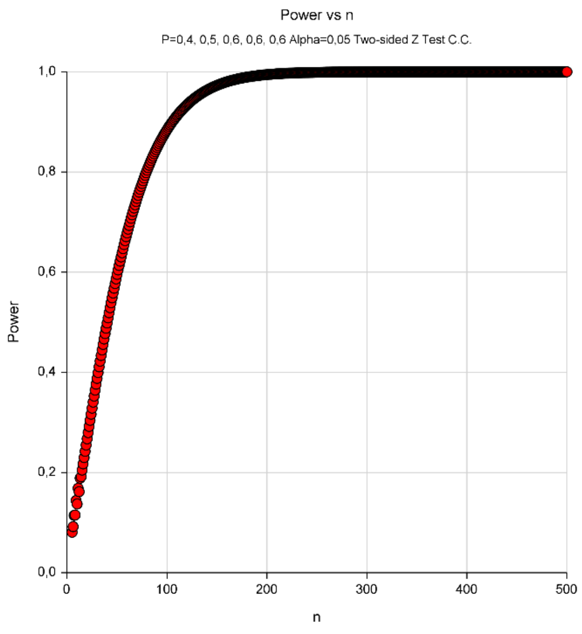

Figure 4. The Cochran-Armitage test power function graph for the 2 x 5 crosstab

Table 5. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 6 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60;0.60;0.60	102	0.23011	0.05
0.40;0.50;0.60;0.60;0.60;0.60	204	0.42772	0.05
0.40;0.50;0.60;0.60;0.60;0.60	300	0.58601	0.05
0.40;0.50;0.60;0.60;0.60;0.60	402	0.71716	0.05
0.40;0.50;0.60;0.60;0.60;0.60	504	0.81261	0.05
0.40;0.50;0.60;0.60;0.60;0.60	600	0.87572	0.05
0.40;0.50;0.60;0.60;0.60;0.60	702	0.92132	0.05
0.40;0.50;0.60;0.60;0.60;0.60	804	0.95110	0.05
0.40;0.50;0.60;0.60;0.60;0.60	900	0.96920	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1002	0.98142	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1104	0.98893	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1200	0.99327	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1302	0.99607	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1404	0.99773	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1500	0.99865	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1602	0.99923	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1704	0.99957	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1800	0.99975	0.05
0.40;0.50;0.60;0.60;0.60;0.60	1902	0.99986	0.05
0.40;0.50;0.60;0.60;0.60;0.60	2004	0.99992	0.05
0.40;0.50;0.60;0.60;0.60;0.60	2100	0.99996	0.05
0.40;0.50;0.60;0.60;0.60;0.60	2202	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60	2304	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60	2400	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60	2502	1.00000	0.05

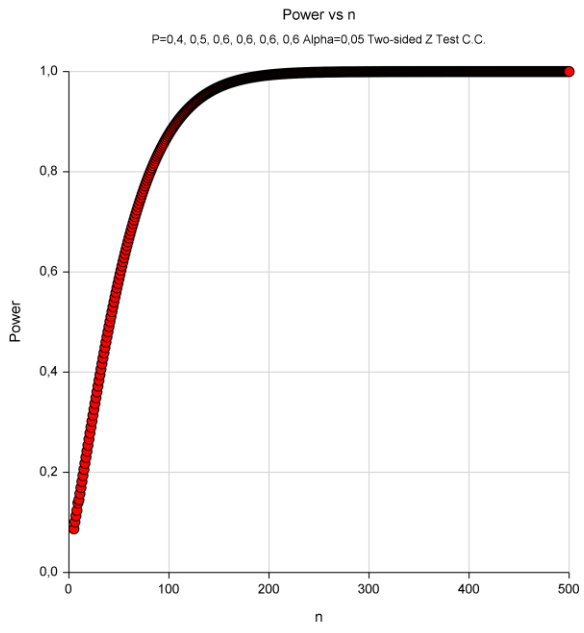

Figure 5. The Cochran-Armitage test power function graph for the 2 x 6 crosstab

Table 6. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 7 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60;0.60;0.60;0.60	105	0.20127	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	203	0.35950	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	301	0.50281	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	406	0.63183	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	504	0.72801	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	602	0.80270	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	700	0.85911	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	805	0.90325	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	903	0.93272	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1001	0.95372	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1106	0.96932	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1204	0.97929	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1302	0.98613	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1400	0.99777	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1505	0.99408	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1603	0.99612	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1701	0.99747	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1806	0.99841	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	1904	0.99897	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2002	0.99934	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2100	0.99958	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2205	0.99974	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2303	0.99983	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2401	0.99990	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2506	0.99994	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2604	0.99996	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2702	0.99997	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2800	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	2905	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	3003	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60	3101	1.00000	0.05

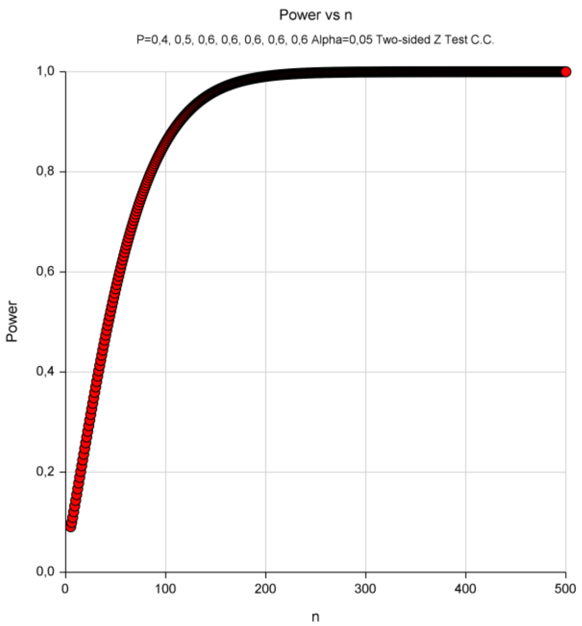

Figure 6. The Cochran-Armitage test power function graph for the 2 x 7 crosstab

Table 7. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 8

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	104	0.17290	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	200	0.30230	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	304	0.43445	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	400	0.54323	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	504	0.64406	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	600	0.72128	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	704	0.78901	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	800	0.83860	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	904	0.88053	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1000	0.91028	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1104	0.93477	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1200	0.95174	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1304	0.96543	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1400	0.97475	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1504	0.98214	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1600	0.98709	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1704	0.99097	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1800	0.99353	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	1904	0.99551	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2000	0.99681	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2104	0.99781	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2200	0.99846	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2304	0.99895	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2400	0.99926	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2504	0.99950	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2608	0.99966	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2704	0.99976	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2800	0.99984	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	2904	0.99989	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3000	0.99992	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3104	0.99995	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3200	0.99992	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3304	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3408	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3504	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3600	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60	3704	1.00000	0.05

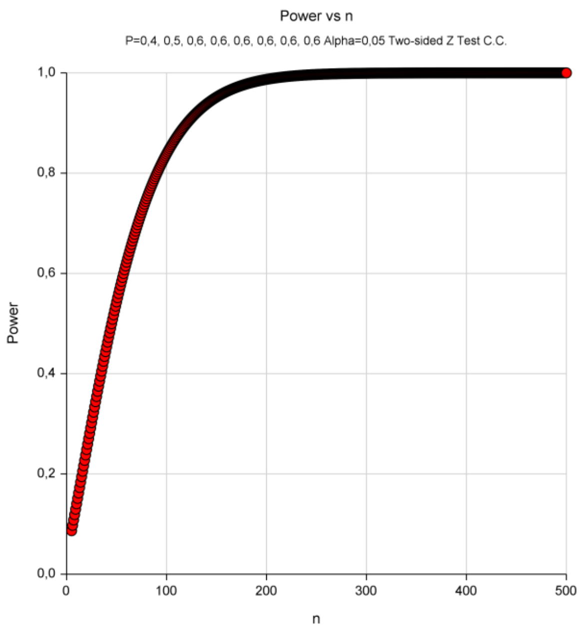

Figure 7. The Cochran-Armitage test power function graph for the 2 x 8 crosstab

Table 8. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 9 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	108	0.15676	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	207	0.26868	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	306	0.37692	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	405	0.37692	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	504	0.56695	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	603	0.64541	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	702	0.71258	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	801	0.76910	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	900	0.81601	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1008	0.85757	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1107	0.88814	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1206	0.91268	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1305	0.93221	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1404	0.94764	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1503	0.95975	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1602	0.96920	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1701	0.97653	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1800	0.98218	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1908	0.98686	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2007	0.99010	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2106	0.99256	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2205	0.99443	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2304	0.99584	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2403	0.99690	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2502	0.99770	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2601	0.99830	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2700	0.99874	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2808	0.99910	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2907	0.99934	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3006	0.99951	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3105	0.99964	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3204	0.99974	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3303	0.99981	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3402	0.99986	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3501	0.99990	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3600	0.99993	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3708	0.99995	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3807	0.99996	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3906	0.99997	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4005	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4104	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4203	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4302	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4401	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4500	1.00000	0.05

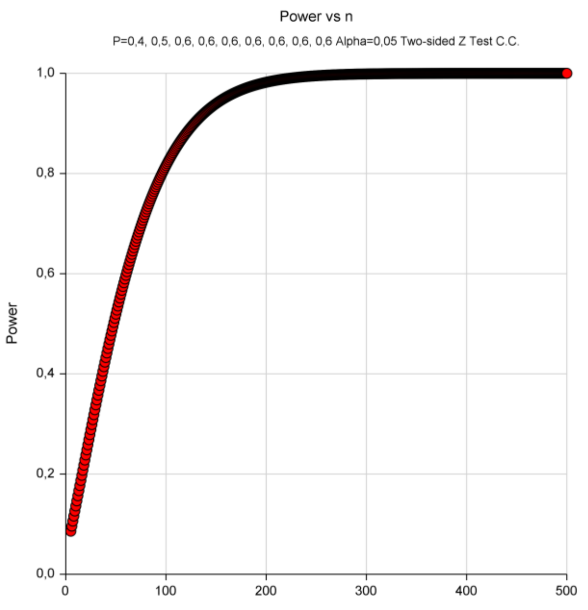

Figure 8. The Cochran-Armitage test power function graph for the 2 x 9 crosstab

Table 9. The power of the Cochran-Armitage trend test for the trend of the proportions in different sample sizes in the 2 x 10 crosstab

p	n	Power (1-β)	Type I Error (α)
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	100	0.13231	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	200	0.22804	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	300	0.32256	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	400	0.41259	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	500	0.49598	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	600	0.57153	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	700	0.63875	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	800	0.69769	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	900	0.74870	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1000	0.79239	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1100	0.82944	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1200	0.86060	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1300	0.88661	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1400	0.90817	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1500	0.92594	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1600	0.94050	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1700	0.95237	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1800	0.96200	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	1900	0.96978	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2000	0.97604	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2100	0.98105	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2200	0.98506	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2300	0.98825	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2400	0.99078	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2500	0.99278	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2600	0.99436	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2700	0.99561	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2800	0.99658	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	2900	0.99735	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3000	0.99794	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3100	0.99841	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3200	0.99877	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3300	0.99905	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3400	0.99927	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3500	0.99944	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3600	0.99957	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3700	0.99967	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3800	0.99975	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	3900	0.99981	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4000	0.99985	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4100	0.99989	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4200	0.99991	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4300	0.99994	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4400	0.99995	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4500	0.99996	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4600	0.99997	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4700	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4800	0.99998	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	4900	0.99999	0.05
0.40;0.50;0.60;0.60;0.60;0.60;0.60;0.60;0.60;0.60	5000	0.99999	0.05

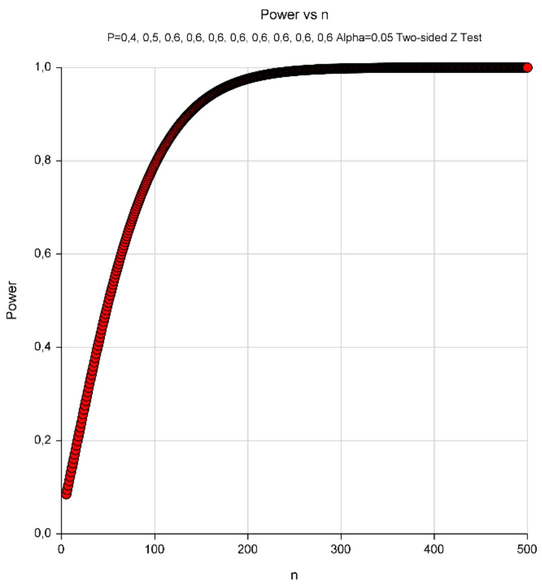

Figure 9. The Cochran-Armitage test power function graph for the 2 x 10 crosstab

When all tables and figures are analyzed;

Randomized studies in biological sciences do not have adequate quality in terms of sample size and power analysis. Thus, attention should be focused on the concepts of sample size and power analysis. The stages of a research are respectively as follows: formulating hypotheses, research design, data collection, and statistical analysis. It is finalized with the evaluation of the validity of the proposed hypotheses through statistical tests. In other words, statistical analysis is the stage before the classic presentation of the research results. This stage involves the interpretation of the test results based on statistical significance. However, just like statistical significance does not give any information about the content of the analyzed data, it may also be found significant just because of the largeness of the sample size in some cases. On the other hand, a statistically insignificant result may result from the smallness of the sample size or taking random variables from a sample that does not represent the related group. In other words, the results may have been found statistically significant or statistically insignificant wrongly. On the other hand, working with a very large sample may lead to loss of time, labor, and money. For that reason, appropriate sample size must be determined within the framework of research hypothesis and research purpose in the planning stage of the study.

In biological research, determination of sample size is an important part of scientific research process. In a study where the sample size is much larger than necessary, the researcher will achieve his purpose before the study ends, and so some experimental units will have been included in the study unnecessarily. On the other hand, when the sample size is much smaller than necessary, it will be less likely for the researcher to achieve his purpose. Accordingly, in a biological research, sample size must allow the addressed hypotheses to be tested reliably.

According to the result of the simulation made in the present study, the power of the test highly varies in different combinations when the least biologically significant differences change. In this study, an attempt was made to determine the most valid combinations in the specified scenarios to keep the power of the test at 80% at the least. When the number of categories is 2, determination of trend requires working with very large samples. When the number of categories is 3 or 4, the desired power can be obtained with smaller samples compared to the case where the number of categories is 2. When the number of categories is over 4, a substantial increase is needed in sample size to obtain the desired power. Change in marginal frequencies does not have much influence on sample size.

This study made an attempt show that the Cochran-Armitage trend test can be used for determining the sample size in different scenarios based on the number of categories in the studies about the power of the linear trend increasing or decreasing in proportions. In this regard, this study may guide researchers in future studies for determination of appropriate sample size, which is one of the main difficulties encountered in especially the disciplines dealing with living beings.

Yol et al. (2013) used sample sizes in trend of 3:1 proportion. In the experiment, sample size was stated to be 65 for purple plants and 22 for normal plants. In the table 1, the results of the simulation study show that sample size has to be approximately 800 for the power of the test to be 79.26%. Yol et al. (2013) took it as p=0.95 and found the results statistically insignificant. Accordingly, it can be said that the fact that the desired power of the test is not achieved and thus the sample size is inadequate may cause the chi-square test to yield insignificant results.

Akgün et al. (2011) reported the survival rates of seedlings treated with colchisin and the control by using 2000 seeds. The survival rates were 440 seedlings giving a 22% ratio with a power over 87% as followed in Table 1. In the table 1, the results of the simulation study indicate that sample size has to be approximately 800 so that the power of the test is 79.26%. Here sample size of 2000 increased the power over 87%.

Cuming et al. (2015) reached 27 and 43 subjects in relation to the parameters they focused on in the natural green color cotton fiber color and fiber quality QTL analysis. Based on the results of our simulation study, as is shown in the table 6, we can say that they have to reach 105 subjects. This may be guiding for determining the sample sizes used in the chi-square tests carried out in the field crops studies mentioned above based on the sample sizes obtained for the Cochran-Armitage trend test according to the simulation results.

LITERATURE CITED

Ahn, K., C. Haynes, W. Kim, R. St. Fleur, D. Gordon, and S. J. Finch. 2007. The effects of SNP genotyping errors on the power of the cochran-armitage linear trend test for case/control association studies. Annals of Human Genetics, 71(2):249–261. http://doi.org/10.1111/j.1469-1809.2006.00318.x

Akgun, I., M. Tosun, K. Haliloglu, M. Aydin. 2011. Development of autotetraploid perennial rye (Secale montanum Guss.) and selection for seed set, Turk J Field Crops, 16(1): 23-28.

Armitage P. 1955. Tests for linear trends in proportions and frequencies. Biometrics; 11:375–386.

Banks, K. E., D. H. Hunter and D. J. Wachal. 2005. Diazinon in surface waters before and after a federally-mandated ban. Science of the Total Environment, 350(1-3):86–93. http://doi.org/10.1016/j.scitotenv.2005.01.017

Chapman D.G. and J. Nam. 1968. Asymptotic power of chi-square tests for linear trends in proportions. Biometrics; 24:317–327.

CochranW.G. 1954. Some methods for strengthening common the tests. Biometrics; 10:417–451.

Cuming Semizer D., F. Altan, H. Akdemir, M. Tosun, A. Gurel and B. Tanyolac. 2015. Qtl analysis of fiber color and fiber quality in naturally green colored cotton (Gossypium hirsutum L.), Turk J Field Crops, 20(1), 49-58

Hintze, J. PASS 11. [Chapter 255 Cochran-Armitage Test for Trend in Proportions] Kaysville, Utah, USA: NCSS, LLC; 2011:2-5: 595:1-7

Kang, S. H., and J. W. Lee. 2007. The size of the Cochran-Armitage trend test in 2 ?? C contingency tables. Journal of Statistical Planning and Inference, 137(6):1851–1861. http://doi.org/10.1016/j.jspi.2006.03.009

Lachin, J. M. 2011. Power and sample size evaluation for the Cochran-Mantel-Haenszel mean score (Wilcoxon rank sum) test and the Cochran-Armitage test for trend. Statistics in Medicine, 30(25):3057–3066. http://doi.org/10.1002/sim.4330

Mehta, C. R., N. R.Patel, P. Senchaudhuri and N. Dec. 2007. Exact Power and Sample-Size Computations for the Cochran-Armitage Trend Test SHORTER COMMUNICATIONS EDITOR : Exact Power and Sample-Size Computations for the Cochran-Armitage Trend Test, 54(4):1615–1621.

Mortazavian S. M. M. and S. Azizi-Nia. 2014. Nonparametric Stability Analysis In Multi-Environment Trial Of Canola, Turk J Field Crops, 19(1):108-117

Nam J. A. 1987. Simple approximation for calculating sample sizes for detecting linear trend in proportions. Biometrics; 43:701–705.

Shen, H., Y. Hu, Y. Chen and T. Tung. 2014. Prevalence and Associated Metabolic Factors of Gallstone Disease in the Elderly Agricultural and Fishing Population of Taiwan.

Slager S.L. and D.J. Schaid. 2001. Case-Control Studies Of Genetic Markers: Power And Sample Size Approximation For Armitage's Test Of Trend. Human Heredity; 52:149–153. DOI: 10.1159/000053370.

Tekindal M.A., H. Bayrak, B. Ozkaya, Y. Yavuz. 2014. Second-order response surface method: factorial experiments an alternative method in the field of agronomy, Turk J Field Crops, 19(1):35-45

Yol E., F. Seymus and B. Uzun. 2013. Genetic Control Of Purple Plant Color In Sesame, Turk J Field Crops, 18(2), 229-232

Zheng, G., and J.L. Gastwirth. 2006. On estimation of the variance in Cochran-Armitage trend tests for genetic association using case-control studies. Statistics in Medicine, 25(18):3150–3159. http://doi.org/10.1002/sim.2250