http://www.pjbs.org

PJBS

ISSN 1028-8880

Pakistan Journal of Biological Sciences

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Pakistan Journal of Biological Sciences 9 (7): 1324-1327, 2006
ISSN 1028-8880
© 2006 Asian Network for Scientific Information

A Study on Power of Chi-square and G Statistics in Biology Sciences

[1]Ecevit Eyduran, [1]Taner Özdemir, [2]M. Kazım Kara, [3]Sıddık Keskın and [4]Bahattin Çak
[1]Biometry Genetics Unit, Department of Animal Science, Faculty of Agriculture,
University of Yüzüncü Yıl, 65080 Van, Turkey
[2]Biometry Genetics Unit, Department of Animal Science, Faculty of Agriculture,
University of Adnan Menderes, Aydın, Turkey
[3]Biostatistics Unit, Faculty of Medicine,
[4]Department of Animal Science, Faculty of Veterinary Medicine,
University of Yüzüncü Yıl, 65080 Van, Turkey

Abstract: The objective of this study was to examined Chi-Square and G test statistics in place of enough sample size, contingency coefficient and power of test for different four contingency tables (data set) regarding biology sciences. Besides, this study was to determine whether sample sizes of various four samples in biology sciences were sufficient. The reliability of two statistics related to Sample size, contingency coefficient and power of test. Power analysis for Chi-Square and G test statistics were performed using a special SAS macro According to results of power analysis, sample sizes of other sets of data except the third data set were determined to be sufficient because power values for both statistics were more than 88%. With respect to power analysis, G statistics for the initial two data sets were more advantageous than other as power value of G statistics were larger than that of other. In the last data set, as sample size were 1607 and power values for both statistics were 100%, both were asymptotically equivalent each other. As power values of the third data set for Chi-Square and G test statistics were approximately 46.77 and 58.16%, respectively, sample size with 20 for both were determined to be insufficient. When we artificially increased 30 to 200 by 10, sufficient sample size for third data should be 50 so as to provide power values of 80% with respect to results of SAS special macro. As a result, this study emphasized that researchers should have taken into sample sizes and power of test account except for probability of Type Error I in contingency tables in order to determine the best one of both statistics.

Key words: Contingency table, Chi-square statistic, G statistic, power of test

INTRODUCTION

Chi-Square and G (which is also called as Likelihood Ratio Chi-Square) statistics have been widely used to test independence or goodness of fit in contingency tables (row by column) (Sokal and Rohlf, 1981; Duzguneo *et al*., 1983; Everitt, 1992; Agresti, 2002; Özdemir *et al*., 2004; Eyduran and Özdemir, 2005; Eyduran *et al*., 2005). The number of variables for two statistics should be two (e.g. x and y) which have either grouped or categorical measurements (Sokal and Rohlf, 1981; Duzguneo *et al*., 1983; Agresti, 2002; Eyduran and Özdemir, 2005; Eyduran *et al*., 2005). G statistic of these two statistics has the Chi-squared approximation when $n \rightarrow +\infty$ is (Duzguneo *et al*., 1983; Agresti, 2002; Eyduran and Özdemir, 2005). Two Statistics gives an idea about whether (or not) an association between two variables is. However, when more than 20% of the cells have expected counts less than 5, Chi-Square Statistic may not be a valid statistic. In other words, Statistics package programs for this case, one of which, SAS program, gives warning Chi-square statistic may not be a valid statistic (SAS, 1998). It was stated, therefore, many authors preferred G statistic to Chi-square statistic (Sokal and Rohlf, 1981; Duzguneo *et al*., 1983; Everitt, 1992; Özdemir *et al*., 2004; Eyduran and Özdemir, 2005; Eyduran *et al*., 2005). Besides, it was reported that for contingency tables, G statistic could be more favorable approach than Chi-square where n/rc was less than five (Agresti, 2002; Eyduran and Özdemir, 2005). In addition, it was suggested that G test was more suitable than other when observed counts in contingency table were less than five (Sokal and Rohlf, 1981; Duzguneo *et al*., 1983; Everitt, 1992; Özdemir *et al*., 2004; Eyduran and Özdemir, 2005). However, it was suggested that values of two statistics were similar to

Corresponding Author: Ecevit Eyduran, Biometry Genetics Unit, Department of Animal Science, University of Yüzüncü Yıl, 65080 Van-Van-Turkey Tel: (+90) 432 225 10 24-2648

each other when sample size was sufficient or much (Sokal and Rohlf, 1981; Özdemir *et al*., 2004; Eyduran and Özdemir, 2005).

The reliability of two statistics associated to sample size, contingency coefficient and power of test. In this context, three criteria are different for each contingency table, which can be composed the various numbers of observations in all cells (Eyduran and Özdemir, 2005). Besides, in any contingency table, one can concern whether sample size, n, for gaining power with at least 80% is sufficient except for probability of type error I. Besides, the best choice between both statistics was based on their power values.

The aim of this study was to examine Chi-Square and G test statistics in place of enough sample size, contingency coefficient and power of test for different contingency tables regarding biology sciences, which were mentioned in materials and methods section, by means of a special macro downloaded from SAS library (http://ftp.sas.com/techsup/download/stat/powerrxc.html). In other hand, by using the special SAS macro, the goal of present paper was to test or determine whether sample sizes for different four sets of sample data on biological sciences were sufficient.

MATERIALS AND METHODS

Materials: As a research material, different data sets were used. Various data sets in relation to biology sciences are given in form of contingency tables as given in Table 1.

Table 1: Contingency tables in data sets regarding biological sciences

Data Set 1: Contingency table of psychological case by sex		
N = 107	Sex (X)	
Psychological problem (Y)	Female (1)	Male (0)
Presence (1)	22	54
Absence (0)	0	31
Data set 2: Contingency table of various experiments by survivor counts		
N=121	Survival case (X)	
Experiments (Y)	Dead (1)	Alive (0)
Bacteria and antiserum(1)	13	54
Bacteria (0)	25	29
Data Set 3: Contingency table of dirty fleece by transferrin types		
N=20	Transferrin Types	
Fleece yield (Y)	A type (1)	AB (0)
Low (1)	7	3
High (0)	10	0
Data set 4: Contingency table of contraceptives use by desiring children		
N=1607	Contraceptives use (X)	
Desiring children (Y)	Using (1)	Not using (0)
Desiring (1)	219	753
No desiring (0)	288	347

Data set 1 was composed of questionnaire records on psychological cases and sex of 107 refugees in 2001 (Özdemir, 2001). The numbers of female and male were 22 and 85, respectively as well as the numbers of refugees being occurrence and absences of any psychological problem were 76 and 31, respectively. As a psychological sample, a question of interest was whether relationship between psychological cases and sex was significant.

Data set 2 was consisted of results of an immunological study that the effects of different two experiments on survivor of 121 mice were examined. One hundred and twenty one mice were separated into two groups, 57 that given standard dose of pathogenic bacteria followed by an antiserum and a control group of 54 received only the bacteria. After satisfactory time elapsed for incubation period, 38 dead mice and 73 survivors were counted. A question of interest was whether association between various experiments and sex is considerable, that is, whether the antiserum had protected mice (Sokal and Rohlf, 1981).

Data set 3 was made up of records on levels of dirty fleece yield and transferrin types of 20 Norduz Goat in 2001 (Eyduran *et al*., 2005). The number of transferrin type A and AB were 17 and 3 sheep, respectively as well as the number of low and high levels of fleece yield were both 10 sheep. In this sample on animal science, a question of interest was whether relationship between different transferrin types and levels of dirty fleece yield was important.

Data set 4 was comprised of results of questionnaire performed on 1607 women (Eyduran *et al*., 2005). A question of interest for the data set was the relationship between contraceptives uses and desiring children was significant.

Different data sets concerning biological sciences in Table 1 were analyzed using a special SAS macro (http://ftp.sas.com/techsup/download/stat/powerrxc.html).

METHODS

Chi-Square 1 and G statistics 2 are used for testing goodness of fit in contingency tables (Everitt, 1992; Agresti, 2002; Eyduran and Özdemir, 2005) and the notation of two statistics are given below:

$$X^2 = \sum \frac{(f - f_i)^2}{f_i} \quad (1)$$

$$G = 2\sum f.\mathrm{In}[f / f_i] \quad (2)$$

Where, f, observed frequency and f_i, expected frequency.

Power estimation for Chi-Square and G statistics: Theoretical details regarding special SAS macro were explained by Agresti (2002) as follows:

Assume that H_0 is equivalent to model M for a contingency table. Let π_i denote the true probability in ith cell and Let π_i (M) denote the value to which the Maximum Likelihood (ML) estimate π_i for model M converges, where $\Sigma\pi_i = \Sigma\pi_i(M) = 1$. For multinomial sample of size n, the non-centrality parameter for Chi-Square 3 can be expressed as follows:

$$\lambda = n\sum_i \frac{[\pi_i - \pi_i(M)]^2}{\pi_i(M)} \quad (3)$$

Expression 3 is the same form as Chi-Square statistics, with in place of the sample proportion p_i and $\pi_i(M)$ in place of π_i. The non-centrality parameter for Likelihood Ratio Chi-Square Statistics 4 can be expressed below:

$$\lambda = 2n\sum_i \pi_i \log\frac{\pi_i}{\pi_i(M)} \quad (4)$$

RESULTS AND DISCUSSION

The values, power values and contingency coefficient of G and Chi-Square Statistics for different four sets of data (or contingency tables) on the subject of biology sciences mentioned above are given in Table 2.

Examined all values of two statistics for the first contingency table (Table 2), there was a close relationship between psychological cases and sex ($p<0.01$). It could be said, therefore, sample size (n = 107) for the first contingency table was sufficient. Because, the power values of G and Chi-Square statistics for the first contingency table were approximately 98.6% and 91.940%, both of which, were much high-level. In other words, it could be suggested that results of the first contingency table were much more reliable. However, for the first contingency table, power value of G statistics were little larger than other and it could be suggested that G statistics were more advantageous than other when one cell of the first contingency table in Table 1 were zero (less than five). As examined in Table 1, the finding, results of power analysis, was in agreement with those reported by Sokal and Rohlf (1981), Duzguneo *et al.* (1983), Everitt (1992), Özdemir *et al.* (2004) and Eyduran and Özdemir (2005).

By examining for the second contingency table in Table 2, all values of the second contingency table were close on each other. The findings on being same as the values of statistics were consistent with those reported by Sokal and Rohlf (1981), Özdemir *et al.* (2004), Eyduran and Özdemir (2005). There was a significant association between various experiments and sex ($p<0.01$). Besides, power values for G and Chi-square statistics of the second contingency table were both also approximately 89%, which had high-reliability. In other words, it could be said that sample size of the second contingency table was sufficient because power values of two statistics was larger than 80%. Many authors were reported that the values of both statistics could be similar (Sokal and Rohlf, 1981; Everitt, 1992; Özdemir *et al.*, 2004; Eyduran and Özdemir, 2005). It could be suggested that both statistics were similar to each other as frequencies of the cells in the contingency table were more than five.

All values of fourth contingency table in Table 2 were close and similar to each other. As Sample size was 1607, power values of two statistics were both 100%. It could be suggested that an ideal data set was data set 4. If sample $n\rightarrow+\infty$, the reliability of study increases.

As to the fourth contingency tables in Table 2, the relationship between contraceptives use and desiring children was much more significant ($p<0.001$).

Considered on the third contingency tables in Table 2, sample size could be said to be insufficient. As shown in Table 1, G Statistic was more reliable than other because two of observed frequencies in the third contingency table were less than five (Sokal and Rohlf, 1981; Duzguneo *et al.*, 1983; Everitt, 1992; Özdemir *et al.*, 2004; Eyduran and Özdemir, 2005). Moreover, SAS program for this contingency table gave warning 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test (SAS, 1998). This was meant, therefore, for the third contingency table, G statistic was more reliable than Chi-Square. The power values of third contingency table for G and Chi-Square statistics were moderate level (58.2%) and low-level (46.8%), respectively.

Although the third data set were sparse data, its contingency coefficient was higher than those of the other sets of data were. However, other data sets had high

Table 2: The values, power values and contingency coefficient of G and Chi-Square statistics in each data set alpha = 0.05

	G statistic value	G statistic probability	Chi-square statistic value	Chi-square statistic probability	G statistic power value	Chi-square power value	Contingency coefficient
Data set 1	17.2735	<0.0001	11.2963	0.0008	0.98596	0.91940	0.3249
Data set 2	10.1040	0.0015	10.0388	0.0015	0.88853	0.88656	0.2768
Data set 3[a]	4.6911	0.0303	3.5294	0.0603	0.58159	0.46767	0.3873
Data set 4	91.6744	<0.0001	92.6442	<0.0001	1.00000	1.00000	0.2335

[a] WARNING: 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test

Table 3: The power values obtained by artificially increasing sample size in data set 3 for alpha = 0.05

Sample size	Chi-square statistics power value	G-statistics power values	Sample size	Chi-Square statistics power value	G-statistics power values
30	0.63343	0.75575	120	0.99588	0.99959
40	0.75706	0.86500	130	0.99767	0.99982
50	0.84387	0.92849	140	0.99870	0.99992
60	0.90217	0.96339	150	0.99928	0.99996
70	0.93999	0.98178	160	0.99960	0.99998
80	0.96386	0.99115	170	0.99978	0.99999
90	0.97858	0.99579	180	0.99988	1.00000
100	0.98748	0.99803	190	0.99994	1.00000
110	0.99278	0.99909	200	0.99997	1.00000

power value. It could be suggested that this case could be resulted from being sufficient sample size (Table).

On condition that contingency coefficients for the third contingency table was fixed when we artificially increased 30 to 200 by 10 by using special SAS macro mentioned above in order to determine sufficient sample size for the third contingency table, minimum sample sizes should be 50 as two statistics were reliable. However, if sample size were 200, the power values of Chi-square and G statistics would be achieved to 99.997 and 100%.

CONCLUSIONS

The aim of this study was to examine Chi-Square and G test statistics in place of contingency coefficient and power of test and to determine sufficient sample size for various contingency tables regarding biology sciences by using a special macro.

As the power values of the first, second and fourth contingency tables in Table 2 were larger than 88%, it could be said that sample size, n, (107, 121 and 1607, respectively) of those contingency tables were more sufficient, and it could be suggested, therefore, G and Chi-Square Statistics results of those contingency tables were more much reliable. However, sample size of the third contingency table was insufficient. Because, the power values of third contingency table for G and Chi-Square statistics were moderate level (58.2%) and low-level (46.8%). Provided that contingency coefficient of two statistics in third contingency table was fixed, when we raised 20 to 200 by 10 in order to find out sufficient sample size for the contingency table (at alpha = 0.05 level), minimum sufficient sample size for it should be 50 with respect to Table 3. In other words, to obtain a reliable result, sample size for G and Chi-Square statistics in the third contingency table should be minimum 50.

If researchers were encountered in warning 50% of the cells have expected counts less than 5. Chi-Square may not be a valid test as being the third contingency table, they could prefer G statistics to other (SAS, 1998).

As a result, it was concluded that the reliability of G and Chi-Square statistics could be changed to sample size, frequencies combination in the cells of contingency table. Besides, This research emphasized that researchers should have taken into sample sizes and power of test account it order to determine the best selection between both statistics rather than probabilities of Type Error I for both were statistically significant.

ACKNOWLEDGMENT

We are thankful to Prof. Dr. Cemil TUNÇ due to his support.

REFERENCES

Agresti, A., 2002. Categorical Data Analysis. 2nd Edn, Wiley, New York.

Duzguneo, O., T. Kesici and F. Gurbuz, 1983. Statistics Methods I. 1st Edn., University of Ankara Publishings of Agriculture Faculty, Ankara, Turkey, Pages: 229.

Everitt, B.S., 1992. The Analysis of Contingency Tables. 2nd Edn., Chapman and Hall. London.

Eyduran, E., T. Özdemir and M. Küçük, 2005. Chi-Square and G Test in Animal Science. The Journal of Faculty of Vet. Med. University of Yüzüncü Yıl, pp: 1-3. Van-Turkey.

Eyduran, E. and T. Özdemir, 2005. Examining Chi-Square, Likelihood Ratio Chi-Square and Independent Ratios in 2x2 Tables: Power of Test. International Congress on Information Technology in Agriculture, Food and Environment-itafe'05. Proceedings, pp: 720-723, Adana-Turkey

http://ftp. sas.com/techsup/download/stat/powerrxc. html.

Özdemir, T., 2001. The Problems of Asylum Seeker Waiting for Refugees or Accepted as Refugees by Union National High Commissary Refugees in Van, in Turkey. Master Thesis. Van-Turkey.

Özdemir, T., E. Eyduran, B. Çak and M.K. Kara, 2004. Comparison of Chi-Square and G Test in Analyzing Categorized Data in Animal Science. University of Süleyman Demirel, IV. National Animal Science Congress, Proceedings, pp: 657-660, Isparta-Turkey.

SAS., 1998. SAS/STAT User's Guide. Version 6, SAS Institute Inc., Cary, North Carolina, USA.

Sokal, R.R. and F.J. Rohlf, 1981. Biometry. The Principles and Practice of Statistics in Biological Research. W.H. Freeman and Company. New York.