An Alternative Approach to Cochran Q Test for Dichotomous Data

Abstract
This paper proposes, develops and presents a statistical method for the analysis of sample data where these data are dichotomous responses assuming only two possible mutually exclusive values such as 1, representing positive response say, and 0 representing negative response say. The Chi-square test statistic based on the Chi-square test for independence is developed as an alternative to the usual Cochran Q test for dichotomous data to test the null hypothesis of equal positive response rate by subjects to a set of test or treatment. The proposed method is illustrated with some sample data and shown to be at least as powerful as the usual Cochran Q test when applied to the same sample observations.

Keywords: Chi-square; Dichotomous data; Cochran Q test; Subjects; Responses; Success

Introduction
Sometimes a researcher may perform an experiment involving repeated observations or blocks, in which the variable of interest is dichotomous, meaning that it can assume only one of two possible mutually exclusive values. One of these two possible values is considered a ‘success’, positive response, present, well, etc. This is often coded as ‘1’, while the other value may be considered a failure, negative response, absent, bad, etc, often coded a ‘0’ [1].

Research interest would then be to determine the proportions of subjects responding positive if the sampled blocks of subjects are the same across all treatments or tests. In this situation Cochran Q test [2-4] for the dichotomous data may be applied.

To adjust, and make allowance for some situations in which test or trial outcomes are not just dichotomous assuming only two mutually exclusive options such as 1 or 0, but when there may be some intermediate and third outcome such as unknown, indeterminate, non-definitive, etc, that may be code with say a minus sign(-). Oyeka, et al. [5] introduced a third category of response and developed a Chi-square test statistic for independence to test the null hypothesis of equal positive response rates under various treatments or test.

We will here however construct and alternative test statistic similar to Cochran Q test assuming that there are only two possible response outcomes or options that may be coded as either 1 or 0.

Methods
The proposed method would then be compared with the usual Cochran Q test as shown below.

The proposed method
Sometimes a researcher may be interested in comparing responses of blocks of subjects to a set of treatments in a diagnostic screened test or clinical trials. Specifically, suppose a researcher has collected a random sample of n block of subjects matched on some demographic characteristics such as age, sex, body weight, etc, where each block of subjects contain some c matched subjects and interest of the researcher is to administer each of these c subjects randomly one of c treatments. The subject responses to each of the treatments are all dichotomous assuming only one of two possible and mutually exclusive response options such as positive, or negative, present or absent, good or bad; success or failure, dead or alive, etc.

Let \( x_{ij} \) be the response by a randomly selected subject from the \( i \)th block of subjects administered treatment \( T_j \) for \( i=1,2,...,n \) and \( j=1,2,...,c \) where each \( x_{ij} \) is dichotomous, either positive or negative, present or absent etc. To develop a test statistic to help determine whether on the average subject responses are the same for all treatments or conditions we may let

\[
\begin{align*}
\pi_+ & = \text{Pr}(x_{ij} = 1) \\
\pi_- & = \text{Pr}(x_{ij} = 0)
\end{align*}
\]

for \( i=1,2,...,n \) and \( j=1,2,...,c \).

Let the total number of positive responses, that is total number of 1’s by subjects administered treatment \( T_j \) and

\[
\begin{align*}
\pi_j^+ & = \text{Pr}(u_j = 1) \\
\pi_j^- & = \text{Pr}(u_j = 0) \\
W_j & = \sum_{i=1}^{n} u_{ij}
\end{align*}
\]

for \( i=1,2,...,n \) and \( j=1,2,...,c \).

Be the number of positive responses, that is total number of 1’s by subjects administered treatment \( T_j \) and

\[
\begin{align*}
W_j & = \sum_{i=1}^{n} u_{ij}
\end{align*}
\]

for \( i=1,2,...,n \) and \( j=1,2,...,c \).

Let

\[
\begin{align*}
\pi_j^+ & = \text{Pr}(u_j = 1) \\
\pi_j^- & = \text{Pr}(u_j = 0)
\end{align*}
\]

and

\[
W_j = \sum_{i=1}^{n} u_{ij}
\]
An Alternative Approach to Cochran Q Test for Dichotomous Data

An illustration of proposed alternative to Cochran Q test: To illustrate the proposed method we apply equation 1 to obtain values of $u_{ij}$ for the data of Table 1 for $i=1,2,...,15; j=1,2,...,4$. The summary values of $u_{ij}$ that is of $f^+_j$ and $f^-_j$ with the corresponding sample proportions $p_j$ of 1's are shown at the bottom of Table 1, for $j=1,2,3,4$. Using the values of $p_j$ in equation 13 we obtain the Chi-square test statistic for the null hypothesis of no difference in possible response rates, that is the proportion of subjects, or patients improving under the four drugs as

$$
\chi^2 = \frac{n \sum (p_j - \pi^j)^2}{\pi^j(1-\pi^j)}
$$

(12)

Which under the null hypothesis $H_0$ of equal positive response rate has approximately the $\chi^2$ distribution with $c-1$ degrees of freedom for sufficiently large $n$ and $c$. Equation (12) when expressed in terms of sampled proportion becomes

$$
\chi^2 = \frac{n \sum (p_j^+ - p_0^+)^2}{p^+ (1-p^+)}
$$

(13)

Illustrative example: The effects of four drug presentations on patients are to be studied. Interest is to determine whether or not the four drugs equally improved patients’ condition. Sixty patients are selected and grouped into 15 blocks so that the four patients in each block are approximately identical in age, initial condition, sex, etc. Patients in each block are random only selected for treatment with only one of the four experimental drugs. After the specified medication period, the patients are classified as either improved (success) or not improved (failure) under a given drug and coded with a 1 or 0 respectively. The results are shown in Table 1. Can it be concluded on the basis of these data that patients improve equally on all the four drugs?

An illustration of proposed alternative to Cochran Q test: To illustrate the proposed method we apply equation 1 to obtain values of $u_{ij}$ for the data of Table 1 for $i=1,2,...,15; j=1,2,...,4$. The summary values of $u_{ij}$ that is of $f^+_j$ and $f^-_j$ with the corresponding sample proportions $p_j$ of 1's are shown at the bottom of Table 1, for $j=1,2,3,4$. Using the values of $p_j$ in equation 13 we obtain the Chi-square test statistic for the null hypothesis of no difference in possible response rates, that is the proportion of subjects, or patients improving under the four drugs as

$$
\chi^2 = \frac{n \sum (p_j - \pi^j)^2}{\pi^j(1-\pi^j)}
$$

(12)

Which under the null hypothesis $H_0$ of equal positive response rate has approximately the $\chi^2$ distribution with $c-1$ degrees of freedom for sufficiently large $n$ and $c$. Equation (12) when expressed in terms of sampled proportion becomes

$$
\chi^2 = \frac{n \sum (p_j^+ - p_0^+)^2}{p^+ (1-p^+)}
$$

(13)

Illustrative example: The effects of four drug presentations on patients are to be studied. Interest is to determine whether or not the four drugs equally improved patients’ condition. Sixty patients are selected and grouped into 15 blocks so that the four patients in each block are approximately identical in age, initial condition, sex, etc. Patients in each block are random only selected for treatment with only one of the four experimental drugs. After the specified medication period, the patients are classified as either improved (success) or not improved (failure) under a given drug and coded with a 1 or 0 respectively. The results are shown in Table 1. Can it be concluded on the basis of these data that patients improve equally on all the four drugs?

The effects of four drug presentations on patients are to be studied. Interest is to determine whether or not the four drugs equally improved patients’ condition. Sixty patients are selected and grouped into 15 blocks so that the four patients in each block are approximately identical in age, initial condition, sex, etc. Patients in each block are random only selected for treatment with only one of the four experimental drugs. After the specified medication period, the patients are classified as either improved (success) or not improved (failure) under a given drug and coded with a 1 or 0 respectively. The results are shown in Table 1. Can it be concluded on the basis of these data that patients improve equally on all the four drugs?
Which with $4-1=3$ degrees of freedom is also not statistically significant ($Z_{0.995}^2=11.35$). However the Chi-square value of 6.073 obtained using the usual Cochran Q test is slightly less than the corresponding Chi-square value of 6.300 obtained using the proposed method. Hence the usual Cochran Q test statistic is likely to lead to an acceptance of false null hypothesis (Type II error) more frequently and is therefore likely to be less powerful than the present method at least for the present data.

Table 1: Patients’ Response to four Drug Preparations (1=improved, 0=Not improved).

| Patients(Blocks) | Drug 1 | Drug 2 | Drug 3 | Drug 4 | Total |
|------------------|--------|--------|--------|--------|-------|
| 1                | 1      | 1      | 0      | 0      | 2     |
| 2                | 1      | 1      | 0      | 1      | 2     |
| 3                | 1      | 0      | 0      | 0      | 1     |
| 4                | 1      | 1      | 1      | 1      | 4     |
| 5                | 1      | 1      | 0      | 1      | 3     |
| 6                | 0      | 1      | 0      | 0      | 1     |
| 7                | 0      | 1      | 1      | 0      | 2     |
| 8                | 1      | 1      | 1      | 0      | 3     |
| 9                | 0      | 0      | 1      | 0      | 1     |
| 10               | 0      | 0      | 0      | 0      | 0     |
| 11               | 0      | 0      | 0      | 0      | 0     |
| 12               | 1      | 1      | 0      | 1      | 3     |
| 13               | 1      | 0      | 0      | 1      | 2     |
| 14               | 0      | 1      | 1      | 0      | 2     |
| 15               | 1      | 1      | 0      | 0      | 2     |
| Total(n)         | 15     | 15     | 15     | 15     | 60(n,c) |

\[ f_j^+ = \frac{n_j - f_j^-}{f_j^-} \]

\[ f_j^- = \frac{n_j - f_j^+}{f_j^+} \]

\[ p_j = \frac{f_j^+}{f_j^-} \]

| Patients(Blocks) | Drug 1 | Drug 2 | Drug 3 | Drug 4 | Total |
|------------------|--------|--------|--------|--------|-------|
|                  |        |        |        |        | 60(n,c) |
|                  |        |        |        |        | 29(=f^+ ) |
|                  |        |        |        |        | 31(=f^- ) |
|                  |        |        |        |        | 0.48(=p  ) |

| Patients(Blocks) | Drug 1 | Drug 2 | Drug 3 | Drug 4 | Total |
|------------------|--------|--------|--------|--------|-------|
|                  |        |        |        |        | 0.62 |
|                  |        |        |        |        | 0.67 |
|                  |        |        |        |        | 0.33 |
|                  |        |        |        |        | 0.33 |
Summary and Conclusion

We have discussed and presented above an alternative modified and probably easier method for the analysis of sample data that may be appropriate for use with the usual Cochran Q test.

A test statistic based on the Chi-square test for independence is developed for testing the null hypothesis that subjects or blocks of matched subjects on the average do not differ in proportions responding positive when administered a number of tests or treatments in a diagnostic screening test or clinical trial.

The proposed test method is illustrated with some sample data and shown to be at least as powerful as the usual Cochran Q test when applied to data of equal sizes. This is because the Chi-square value of 6.073 obtained using the usual Cochran Q test is slightly less than the corresponding Chi-square value of 6.300 obtained using the proposed method. Hence the usual Cochran Q test statistic is likely to lead to an acceptance of false null hypothesis (Type II error) more frequently and is therefore likely to be less powerful than the present method at least for the present data.

References

1. Oyeka ICA, Nnanatu CC (2014) Pairwise Comparison in Repeated Measures. Journal of Modern Applied Statistical Methods 13(2): Article 8.
2. Oyeka CA (2010) An Introduction to Applied Statistical Methods. (8th edn), Nobern Avocation Publishing Company, Enugu, Nigeria.
3. Cochran WG (1950) The comparison of percentages in matched samples. Biometrika 37(3-4): 256-266.
4. Freund JR. (1992) Mathematical Statistics. (5th edn), Prentice-Hall Internal Editions, New York, USA.
5. Oyeka CA, Utazi CE, Nwosu CR, Uwawunkonye GE, Ikpegbu PA, et al. (2010) A Statistical Comparison of Test Scores: A Non-Parametric Approach. Journal of Mathematical Sciences 21(1): 77-87.