International Journal of Engineering & Technology, 7 (3.18) (2018) 4-6

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Technical paper

Cochran's Q with Pairwise McNemar for Dichotomous Multiple Responses Data: a Practical Approach

Donald Stephen[1]*, Shahren Ahmad Zaidi Adruce[1]

[1] The Institute of Borneo Studies, Universiti Malaysia Sarawak
**Corresponding author E-mail: donaldstephen89@gmail.com*

Abstract

When utilizing single-response questions for a survey, researchers often overlook the possibility that an item can have a smorgasbord of viable answers. It results in the loss of information as it forces the respondents to select a best-of-fit option. A multiple-responses question allows the respondent to select any number of answers from a set of preformatted options. The ability to capture a flexible number of responses allows collectively exhaustive concepts to manifest for deductive verification. This paper explores the practical use of Cochran's Q test and pairwise McNemar test to examine the proportion of responses derived from the results of Multiple Responses Analysis (MRA). This includes Cochran's Q operation on MRA data table using a simulated data set. Cochran's Q test detects if there is a difference in the proportion of multiple concepts. In the case of a significant result, it would require a post hoc analysis to pinpoint the exact difference in pairwise proportions. This pairwise difference can be detected by utilizing pairwise McNemar test with Bonferroni Correction. This paper serves as a reference for researchers and practitioners who need to examine the proportion of collectively exhaustive concepts collected from a multiple responses item.

Keywords: *Cochran's Q, dichotomous multiple responses data, McNemar, Multiple Responses Analysis, proportion*

1. Introduction

Due to the complexity of MRA data structure, analyzing and tabulating MRA data can be an arduous process [1]. Lack of references in dealing with proportion difference renders MRA data preparation and analysis even more operose. Despite the laborious operation, MRA is essential in many facets of rating scales. The significance is more apparent when dealing with the situation when collectively exhaustive data is possible. In this situation, the use of single response question will result in loss of information as it forces the respondents to disregard other responses by selecting a best-of-fit answer [2].

Although the MRA data can be tabulated, researchers often face the difficulty to justify whether pairwise proportions are statistically different. To solve this problem, this paper shows the operation of Cochran's Q on MRA data. In 1950, Cochran presented a statistical test, the Q test to examine differences in proportions in c paired samples involving a large number of option (row), k [3]. For this purpose, MRA data need to be coded to mimic a data that is suitable for Cochran's Q. Data required for Cochran's Q is dichotomous - for example right or wrong, success or failure, and others [3-5]. In the case of MRA, an option can be selected or not selected. [6] described the use of Cochran's Q test to analyze multiple responses data but a more practical guide is needed to facilitate the process.

This paper explains the coding process of multiple responses data, which is crucial in the implementation of Cochran's Q. This paper also describes the treatment of missing data in MRA, which was often the missing element in existing guides. When the number of criteria tested (dependent samples) is exactly 2, Cochran's Q test and McNemar test will yield an equivalent result. Thus, McNemar test is used as the post hoc test when the null hypothesis is rejected [6]. McNemar test is a popular method to analyse significant difference of paired binary data [7-8]. It determines marginal homogeneity of row and column in 2x2 contingency table [8-9].

2. Method

To demonstrate the calculation, an example of a multiple responses item is formulated. This item requires the participants to select any criteria that they think are important in designing a phone:

☐	Price (A)	☐	Practicality (D)
☐	Colour (B)	☐	Design (E)
☐	Materials (C)	☐	Technical specification (F)

In this example, respondents have 6 options to which they can select any number of criteria that they think are important in designing a phone. If they think none of the criteria is important, they may choose not to select any. In this case, an empty response should not be regarded as a missing data. This should be one of the imperative assumptions when conducting the survey involving multiple responses items.

The dichotomous method of MRA data coding is implemented prior to the application of Cochran's Q calculation. The dichotomous method of storing MRA data separates an MRA item into several polar questions based on the number of options in the manner of "Did the participant select Price as one of the important criteria?", "Did the participant select Colour as one of the important criteria?" and so on. Every subset of MRA is viewed as a separate item that utilizes dichotomous dummy coding [10]. If an option was selected, it will be assigned a code 1. If it is not selected, it will be assigned as 0. This will ensure each option is represented by a variable. Data obtained from dichotomous multiple responses data is paired dichotomous data with the independence of observation. The example of an MRA data table is as in Table 1.

Table 1: Example of MRA data using dichotomous coding

Partic-ipant	Criteria					
	A	B	C	D	E	F
1	1	1	1	1	1	1
2	0	0	0	0	1	0
3	0	0	0	0	0	0
4	0	0	0	0	1	0
5	0	0	0	0	1	0
6	0	0	0	1	1	0

1 - Selected; 0 - Not selected

Researchers must note that although the option on MRA item is collectively exhaustive, the dichotomous coding of Cochran Q must be mutually exclusive. This method of coding multiple responses data is suitable for Cochran's Q because rows containing only 1's (as for participant 1) or only 0's (as for participant 3), will not affect the value of Q, unlike the F test [4]. This attribute of Q test enables it to analyze multiple responses data. However, secondary resources often unaware of this attribute. This is a common mistake while conducting Cochran's Q test [4].

To demonstrate the use of Cochran's Q, a simulated data is generated. A sample size of more than 16 is appropriate for Cochran's Q test [11]. The simulated data contains 68 respondents answering to a question. Data for simulated data is shown in Table 2. Ideally, the sample of n subjects should be randomly selected. However, in the practical scenario, obtaining a random sample is not always feasible.

Table 2: Relative people and responses percentage of simulated data

Criteria	Count	Relative person %	Relative Responses %
Price (A)	15	22.7%	13.2%
Colour (B)	7	10.6%	6.1%
Materials (C)	4	6.1%	3.5%
Practicality (D)	27	40.9%	23.7%
Design (E)	40	60.6%	35.1%
Technical (F)	21	31.8%	18.4%

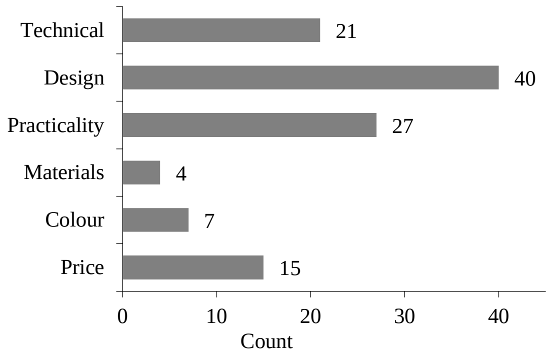

Fig. 1: Selected criteria from MRA

From Figure 1, Practicality, Design and Technical aspects regarded the highest number of endorsements. Therefore, these 3 criteria are extracted to be analyzed using Cochran's Q.

Cochran's Q in this context can be computed as: Suppose k is the number of criteria tested where k≥2 involving N participants. In this case, k=3 is selected based on the result of MRA

H_0: The proportions of responses are equal across all criteria

$H_0: \pi_1 = \pi_2 = \cdots = \pi_k$

H_1: The proportions of responses are not equal across all criteria.

$H_1: \pi_a \neq \pi_b$ for at least one pair π_a and π_b, with a ≠ b and 1 ≤ a, b≤ k.

Table 3: Calculation on MRA data table

D	E	F	$\sum_{j=1}^{k} X_{ij}$	$\left(\sum_{j=1}^{k} X_{ij}\right)^2$
X_{11}	X_{12}	X_{13}	$X_{11}+X_{12}+X_{13}$	$(X_{11}+X_{12}+X_{13})^2$
X_{21}	X_{22}	X_{23}	$X_{21}+X_{22}+X_{23}$	$(X_{21}+X_{22}+X_{23})^2$
X_{31}	X_{32}	X_{33}	$X_{31}+X_{32}+X_{33}$	$(X_{31}+X_{32}+X_{33})^2$
⋮	⋮	⋮	⋮	⋮
$\sum_{i=1}^{N} X_{ij}$	$\sum_{i=1}^{N} X_{ij}$	$\sum_{i=1}^{N} X_{ij}$	Sum: $\sum_{i=1}^{N}\left(\sum_{j=1}^{k} X_{ij}\right)$	Sum: $\sum_{i=1}^{N}\left(\sum_{j=1}^{k} X_{ij}\right)^2$

$$Q=\frac{(k-1)\left[k\left(\sum_{j=1}^{k}\left(\sum_{i=1}^{N} X_{ij}\right)^2\right)-\left[\sum_{i=1}^{N}\left(\sum_{j=1}^{k} X_{ij}\right)\right]^2\right]}{k\left[\sum_{i=1}^{N}\left(\sum_{j=1}^{k} X_{ij}\right)\right]-\left[\sum_{i=1}^{N}\left(\sum_{j=1}^{k} X_{ij}\right)^2\right]} \quad (1)$$

Equation (1) shows the formula for Cochran's Q test. Table 3 helps in the understanding of the computation method of Cochran's Q test. After obtaining Q value, p-value can be calculated as $p=\Pr\left(Q>X^2_{1-\alpha,k-1}\right)$, where $X^2_{1-\alpha,k-1}$ is the value of the (1-α) quantile of the chi-square distribution based on the degree of freedom (df), that is k-1. If Cochran's test is significant, post-hoc analysis will be executed by implementing multiple McNemar analysis with Bonferroni correction.

Table 4: Descriptive statistics of Practicality, Design and Technical

Criteria	N	Mean	SD	Min	Max	Count 0	Count 1
D	68	.3971	.49293	.00	1.00	41	27
E	68	.5882	.49581	.00	1.00	28	40
F	68	.3088	.46544	.00	1.00	47	21

As in Table 4 the three characteristics range from 0.309 to 0.589. Cochran's Q test indicates there is a significant difference between the proportion of responses on three criteria: Practicality, Design and Technical aspects $X^2(2) = 10.481, p = 0.005$. Null hypothesis is rejected because the proportion of the responses on the three characteristics is not equal, hence indicates the need to perform post-hoc test by implementing pairwise McNemar tests with Bonferroni Correction. Bonferroni correction is applied to reduce the possibility of type 1 error. With Bonferroni Correction, α is defined by 0.5/3=0.017.

Table 5: Cross tabulation- Design & Technical

Design	Technical	
	not selected	Selected
not selected	18	10
Selected	29	11

Table 6: Cross tabulation- Design & Practicality

Design	Practicality	
	not selected	Selected
not selected	16	12
Selected	25	15

Table 7: Cross tabulation- Technical & Practicality

Technical	Practicality	
	not selected	Selected
not selected	28	19
Selected	13	8

Table 8: McNemar analysis of Practicality, Design, and Technical aspects

Pair	Proportions		Difference (Unsigned)	McNemar Test (2-tailed)	For Discordant Cells:		
	P_A	P_B			A=1 B=0	A=0 B=1	Odds Ratio (larger/smaller)
Design (A) & Technical (B)	0.5882	0.3088	0.2794	0.003378	29	10	2.9
Design (A) & Practicality (B)	0.5882	0.3971	0.1911	0.047031	25	12	2.0833
Technical (A) & Practicality (B)	0.3088	0.3971	0.0883	0.377086	13	19	1.4615

Table 5, Table 6, and Table 7 show cross tabulation of three criteria. Three different pairs can be formed- Design & Technical, Design & Practicality, as well as Technical & Practicality. Pairwise McNemar is performed on each pair to examine the proportion. Table 8 shows the unsigned difference between proportions of the various pair, ranging from 0.0883 to 0.2794. Pairwise McNemar result shows that the difference is with Design and Technical pair with the p value of 0.003. With Bonferroni Correction, the new alpha is set as 0.017. p(0.003)<0.017. By examining the odd ratio, Technical and Practicality pair recorded 1.4615, which is close to the expected value of 1 for a non-significant result. However, Design and Practicality pair recorded odd ratio value of 2.0833. This value indicates that although the Design and Practicality pair shows a borderline non-significant result.

Consider another example whereby a post hoc analysis is not needed. In Table 9 (another set of simulated data), the proportions are examined to detect if there is a significant difference between any pair. The proportions of responses for Practicality, Design, and Technical range from 0.309 to 0.353. Cochran's Q test shows there is no significant difference between proportion of responses on three criteria, $X^2(2)=0.259$, $p=0.878$. In this case, it does not require a post hoc analysis as pairwise comparison using McNemar test would also yield insignificant result. Therefore, it can be concluded that all three criteria were regarded as equally important in designing a phone.

Table 9: Descriptive statistics of Practicality, Design and Technical (simulated data 2)

Criteria	N	Mean	SD	Min	Max	Count 0	Count 1
D	68	0.3529	0.48144	.00	1.00	44	24
E	68	0.3088	0.46544	.00	1.00	47	21
F	68	0.3235	0.47130	.00	1.00	46	22

3. Summary

From the example of multiple responses item: "select any criteria that you think are important in designing a phone", there are 6 options available to which participants can select any number of options. They can also choose not to select any option if they think none of the criteria apply to the question. To examine the proportion of responses, researchers should code the responses of multiple responses data into dichotomous dummy coding (1- selected; 0- not selected). This allows for analysis with Cochran's Q test. From the result of MRA, Practicality, Design, and Technical aspects were found to have the highest proportion. Cochran's Q test indicated that there is a difference among the proportions. In this case, pairwise McNemar with Bonferroni Correction is adapted to examine the difference between each pair. The difference in proportion was found in Design and Technical pair. From this result, it can be concluded that Design can be regarded as the most prominent criteria in designing a phone.

4. Conclusion

Multiple Responses Analysis (MRA) is performed to obtain relative frequency to make a deductive verification about the distribution of responses on all options. Data transformation of multiple responses data is necessary due to the complexity of data structure. This stage is crucial to ease tabulating process and data analysis. Although this paper describes the application of Cochran's Q and pairwise McNemar, it does not disregard the application of other suitable methods to deal with the complexity of multiple responses data. Future studies should examine various methods and make a comparative analysis to find the most suitable approach in a different context of use.

Acknowledgement

Authors would like to acknowledge The Ministry of Higher Education, Malaysia for providing financial aid through Mybrain15 sponsorship.

References

[1] Jann B (2005), Tabulation of multiple responses. *Stata Journal* 5(1), 92–122.
[2] Plašil M, Vlach P (2006), Analysis of multiple-response data. In *9th International Scientific Conference on Applications of Mathematics and Statistics in Economics*. Trutnov.
[3] Cochran WG (1950), The comparison of percentages in matched samples. *Biometrika* 37(3), 256–266. http://doi.org/10.1093/biomet/72.2.241
[4] Tate MW, Brown SM (1970), Note on the Cochran Q test. *Journal of the American Statistical Association* 65(329), 155–160. http://doi.org/10.1080/01621459.1970. 10481069
[5] Yarnold PR (2015), UniODA vs Cochran's Q test: Comparing success of alternatives. *Optimal Data Analysis* 4(7), 116–117.
[6] Sheskin DJ (2007), *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press, Boca Raton. http://doi.org/10.1201/9781420036268
[7] Fagerland MW, Lydersen S, Laake P (2013), The McNemar test for binary matched-pairs data: Mid-p and asymptotic are better than exact conditional. *BioMed Central Medical Research Methodology* 13(91), 1–8. http://doi.org/10.1186/1471-2288-13-91
[8] Yarnord PR (2015), UniODA vs. McNemar's test for correlated proportions: Diagnosis of disease before vs after treatment. *Optimal Data Analysis* 4, 24–26. http://doi.org/10.1007/BF02295996
[9] Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PMM (2015), Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *Journal of Clinical Epidemiology* 68(3), 299–306. http://doi.org/10.1016/j.jclinepi.2014.09.00
[10] Santos JR (2000), Getting the most out of multiple response questions. *Journal of Extension* 38(3), 3–8.
[11] Lewis-Beck M, Bryman AE, Liao TF (2003), The SAGE encyclopedia of social science research methods. SAGE Publications, Thousand Oaks.