

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

Donald Stephen<sup>1\*</sup>, Shahren Ahmad Zaidi Adruce<sup>1</sup>

<sup>1</sup> The Institute of Borneo Studies, Universiti Malaysia Sarawak

\*Corresponding author E-mail: [donaldstephen89@gmail.com](mailto: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:

<input type="checkbox"/>	Price (A)	<input type="checkbox"/>	Practicality (D)
<input type="checkbox"/>	Colour (B)	<input type="checkbox"/>	Design (E)
<input type="checkbox"/>	Materials (C)	<input type="checkbox"/>	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 select-