STAT.Mahammad.Mahmoud@ MSc.2024.Non Parametric Statistics

by Mahammad Mahmoud

Submission date: 15-Jan-2024 07:40PM (UTC+0200)
Submission ID: 2271451114
File name: STAT.Mahammad.Mahmoud_MSc.2024.Non_Parametric_Statistics.docx (130.34K)
Word count: 2970
Character count: 17143

Salahaddin University-Erbil

College Of Administration and Economics

High Education: Master

Department Of Statistics and Information

Subject: Non-Parametric

Semester: First

A Review Article about

Cochran's Q test

By

Mahammad Mahmoud Bazid

mohammadsherwany@gmail.com

Supervised by

Asst. Prof. Dr. Nazeera Sedeek Kareem

Academic year
2023-2024

Cochran's Q Test

1-Abstract

Motivation: Nonparametric statistics is a field of statistics that deals with data that may not fulfill the assumptions of parametric statistics. Nonparametric statistics, on the other hand, do not make significant assumptions about the underlying population distribution and sometimes depend on fewer or no distributional assumptions. Nonparametric tests are distribution-free, meaning they don't presume a particular probability distribution for the population. They are typically employed when the data are ordinal, categorical, or not normally distributed. Nonparametric tests make fewer assumptions about the data, making them more resilient in the case of outliers or skewed distributions. They are appropriate in instances While the hypotheses of parametric evaluation are broken or when working with data that cannot be assessed on an interval or ratio scale.

Description: The paper presents an enhanced version of Q test that utilize statistical methods to tackle instability in real-world scenarios. The CQ test examines the constancy of percentages of a certain outcome over numerous groups, taking into account both deterministic and undetermined factors.

Results: The test is implemented by using data from operations to analyze the machines' capabilities on different days of the workweek. The comparative analysis demonstrates the advantages of the proposed test in contrast to Cochran's Q test, providing valuable insights into the level of uncertainty and enhancing decision-making in ambiguous situations.

Conclusion: This study presents an alternative version of the Cochran test, using statistical methods to address ambiguity in practical situations. Cochran's Q test effectively assesses the reliability of result percentages across diverse groups, accounting for deterministic and undetermined factors. When this method is used to assess the capacities of machines, by analyzing production data collected on different days of the week, it reveals its superiority compared to Cochran's Q test. This benefit is visible in the insights it gives on the degree of indeterminacy, thereby boosting decision-making in circumstances

distinguished by uncertainty. The simulation study further underscores the critical impact of uncertainty in affecting test statistics and decision outcomes, demonstrating the usefulness of the recommended approach in representing practical problems volatility. Cochran's Q test presents a sophisticated and realistic strategy for addressing the uncertainties inherent in various datasets, proving it useful in common decision-making scenarios.

Keywords: Nonparametric statistics, Cochran's Q test statistics,

2-Introduction

CQ Test is a Nonparametric statistical test used to analyze dependent categorical data or repeated measurements. It is named after the British statistician William G. Cochran. Renowned statistician William G. Cochran made important advances in the science of statistics. He introduced the Cochran Q test as a component of his seminal work. Thus, 1950 is credited with the discovery and introduction of Q test. this test was first introduced and has since grown to be a vital statistical tool, especially when examining correlated categorical data or repeated measurements. Numerous disciplines, including psychology, education, social sciences, and medicine, make extensive use of the test. It is especially effective for evaluating data in which each observation is matched with another observation across multiple groups or situations Q test, a statistical tool, analyses the coherence of percentages across different categories in a binomial sample.

2.1-Purpose of Cochran's Q test:

The main purpose of CQ test is to evaluate if there is a significant difference in the distribution of a categorical variable across many related groups. It is especially advantageous in scenarios where the dependent variable is binary or categorical, and the data is regularly gathered from the same individuals or institutions.

2.2- When Use CQ Test:

The Cochran's Q test is suitable in the following circumstances:-

1-The dependent variable exhibits categorical characteristics.

2-The data are acquired from linked groups or several measurements.

3-The variable that is important has 2 categories (binary) or more.

2.3- Methods of Cochran's Q Test:

1-Data Collection:

Collect categorical data from similar groups. Every group or individual must have measures in similar settings or at various points in time.

2-Formulate Hypotheses:

The null hypothesis (H0) often asserts that there are no significant variations in proportions among groups.

The alternative hypothesis (H1) argues that there is a considerable difference.

3-Compute CQ Statistic:

The test statistic, Cochran's Q, is obtained using the formula:

$$Q = \frac{(K-1) * (K * sum\ of\ squared\ differencess)}{sum\ of\ squared\ ranks}$$

Where **K** is the number of groups.

4- degrees of freedom:

(DF) = K – 1

5- Comparison with Critical Value:

The derived Q statistic is contrasted with a critical value from the chi-square distribution to evaluate statistical significance.

6-Dicision rule:

If a calculated Q statistic greater than the critical value, the null hypothesis is rejected, suggesting substantial variations in percentages among groups.

7- Post Hoc Tests (if necessary): If CQ is significant, post hoc tests (e.g., McNemar's test) may be undertaken to establish which individual groups vary considerably.

3- Review

In 2007 (Keneth J. Berry, Janis e. Johnston, and Paul W. Mielke, jr.) published an article about an alternative measure of effect size for Cochran's Q test for related proportions. Assessments of effect size are becoming significant in the field of psychology. In this work, a chance-corrected estimate of effect size is offered for Cochran's Q test for correlated percentages. The bipolar variable has various uses in psychological research and evaluation. Generally, a value of one is provided for every question that lets a subject respond well, while a value of zero is granted to any incorrect reply. A typical example is when volunteers are placed into an experimental situation, observed as to whether or not some certain response is created, and graded appropriately (Lunney, 1970). Cochran's Q test for related percentages (1950) is commonly used in psychological research to determine whether more than three matched groups of percentages differ with others. The matching may be based on elements of the numerous topics or related subjects with varying settings. For two matched sets of proportions, Cochran's Q test reduces to McNemar's test of change for two correlated sets (in 1947) (Kenneth J. Berry, Janis E. Johnston, and Paul W. Mielke, jr,2007).

In 2008 (Biggerstaff and Jackson) this article has determined the precise the range of CQ coefficient and have illustrated how this could be exploited about every one of its various purposes in a systematic review. The hypothesis that was produced is being demonstrated in both real-world and more possible situations: in especially, the distribution of samples delivers penteration into the reliability of the typical test for an existence of variability and the distributions of samples of demonstrating the effects of this, and it may also be utilized to provide estimation intervals for t2. The technique is therefore of interest to those who apply standard meta-analysis procedures, regardless of their choice for the different applications of Q. The exact distributional result for Q and the related discoveries extend naturally to ordinary meta-regression under the normal model, by a simple change to the matrix formalism utilized here, and this may form the focus of future investigation. The presence and impact of heterogeneity

in “the standard one-way random effects model in the meta-analysis are typically explored using the Q statistic provided to Cochran. We establish the exact distribution of this statistic under the assumptions of the random effects model, and also propose two moment-based approaches and an approximate saddle point estimate for Q. The strategy is presented by analyzing a recent simulation work concerning the heterogeneity metrics and applying all the recommended techniques to four published meta-analyses” (Biggerstaff and Jackson,2008).

In 2010 (Tiago V. Pereira and Nikolaos A. Patsopoulos) This work provides a thorough analysis of the interpretation of Cochran's Q test in the context of meta-analyses. It specifically highlights the impact of power and pre-existing assumptions on heterogeneity. By examining 1011 meta-analyses of clinical trials with binary outcomes, the research attention on the importance of the Q test in detecting differences across studies (T^2) and the significance of previous perspectives on heterogeneity. The reference to Bayes theorem highlights the common occurrence of limited ability to identify regular variability, since non-significant Q tests typically do not change established ideas, but strong discoveries substantially increase the probability of variability. (Pereira and Nikolaos A. Patsopoulos,2010).

In 2015 (Elena Kulinskaya and Michael B Dollinger) published an article about the study investigating the distribution of Cochran's Q when analyzing log odds ratios between two arms in individual studies. It reveals that the distribution of Q deviates from a chi-square distribution, particularly when binomial probabilities in the arms are far from 0.5. The convergence to the correct chi-square distribution is slow with increasing study sizes. Formulas for estimating moments suggest a gamma distribution provides a good fit. The $Q\gamma$ test, derived from this distribution, competes well with the Breslow-Day test for homogeneity. However, in routine testing, the simpler Breslow-Day test is recommended. In sparse data situations, where the Breslow-Day test struggles, Qγ remains well-defined and is recommended. The study emphasizes the impact of the effect of interest on the non-asymptotic distribution of Q, with potential

improvements in estimators for small heterogeneity values being a subject for future research (Elena Kulinskaya and Michael B Dollinger,2015).

In 2017 (thierry fahmy and arnaud belletoile) published a paper about Algorithm 983: Fast Computation of the Non-Asymptotic Cochran's Q Statistic for Heterogeneity Detection We have proposed a rapid execution of Cochran's accurate Q test for heterogeneity identification. According to the absence of the sample and the number of individuals, our approach also works well for several treatments as big as 20. It substantially widens the list of applications for which the precise form of Cochran's Q test may be used which was previously confined to relatively tiny data. Our consumers may thus profit from the correct p-value even for reasonably large data sets while not surpassing a few minutes of processing time. In terms of calculation time, our technique has been demonstrated to outperform by many orders of magnitude (up to 106 times quicker) a commercially available solution that serves as a reference in the domain. After a quick discussion of Cochran's Q test and the reason for its precise version, we describe our technique and demonstrate how it is applied (Fahmy and Bellétoile, 2017).

In 2018 (Donald Stephen and Shaheen Ahmad Zaidi Adruce) study covered how to take advantage of Cochran's Q test and paired McNemar test to analyze the percentage of replies generated from the findings of Multiple replies Analysis (MRA). This incorporates Cochran's Q procedure on MRA tables of data using a simulated data set. Cochran's Q test discovers whether there is a variation in the percentage of numerous ideas. In the context of an important finding, it would need a post hoc analysis to determine the particular variation in pairwise ratios. This study provides a guide for educators and professionals who need to investigate the percentage of collectively exhaustive notions acquired from a multiple replies question. In the future, studies should study numerous methodologies and perform a comparison analysis to discover the best-suited way in a different context of usage (Donald Stephen and Shaheen Ahmad Zaidi Adruce,2018).

In 2021 (I. J. David1 and M. U. Adehi) this study evaluated the incidence of geohelminth eggs in 184 samples of soil of 16 playgrounds in Abuja city, Nigeria. The Cochran's Q(CQT) test was employed to assess if there

had been significant variation in the ratios of the examined egg kinds in the soil samples, and it was discovered which of the egg kinds had considerably bigger ratios using the minimum required difference (MRD) approach. It was revealed that these geohelminth yolks exist in considerably varying amounts in the 184 samples of soil studied. Finally, Taenia and Coccocidia eggs are very prevalent, with percentages substantially bigger than the ratios of each of the other geohelminth eggs in the analyzed soil samples, depending upon the MRD mean comparative test (I. J. David1 and M. U. Adehi,2021).

In 2022 (Lee Mason and Maria Otero) The research examines Cochran's Q Test use in measuring stimulus over-selectivity within the verbal repertoire of children with ASD. It underlines the relevance of applying statistical tools, notably Cochran's Q test, for quantitatively analyzing stimuli under selection. The findings of the test have therapeutic relevance in tailoring behavior-analytic therapies. Significant discrepancies in the spoken repertory imply the need for enhanced therapy, and those with greater Cochran's Q values may need more intense care. The report argues that Cochran's Q test may be used as a continuous tool to track developmental changes over time and analyze the effect of different treatments in single-subject study designs. Additionally, Cochran's Q is presented for assessing over-selectivity in behavioral data and showing constrained stimulus control over selection-based responses. The test's potential applicability in researching derivational stimulus control is also highlighted, especially in examining differences within percentages of distinct dependent populations, such as reflexive, symmetric, and transitive relations. Overall, Cochran's Q test is given as a beneficial tool for analyzing and individualizing treatments in the language behavior of persons with Autism.(Mason et al., 2022).

In 2022 (G.Kumar and E.J. LalithKumar) published a paper about the application of Cochran's Q-test for evaluating the performance of an educational program in such instances we may apply CQ test from full study procedure. For such an analysis, one value of the attribute is recorded with a "1" and the other with a "0". Cochran (1950) created a test for variations whereby

he designated “1” for success and “0” for failure. The test process is termed Cochran’s Q-test. The test is a dichotomous counter-component of the Friedman Test. A review of the data we have noted that there exists a considerable difference in the viewpoint of the research scholars in respect of the success of the conduct of the seminar. The editor of this study would like to remark that Cochran’s Q-test has been effectively applied in numerous fields viz, behavioral sciences, business, educational, and social sciences for evaluating the importance of differences in numerical data (Kumar and LalithKumar).

In 2023 (Muhammad Aslam) published a paper about the suggested Cochran’s Q test increases decision-making in uncertain settings by adding the degree of indeterminacy. It gives a thorough framework, offering insights into the test statistic's range and accompanying uncertainty. Simulation research indicated that increasing uncertainty lowered the test statistic's range, reducing power and error rates. This underlines the necessity of recognizing ambiguity in decision-making. The Cochran’s Q test is beneficial in scenarios with ambiguous data, allowing decision-makers a nuanced knowledge and aiding educated choices. study may examine further statistical features of this suggested test (Muhammad,2023).

4- Conclusion

We reviewed several publications on Major Analysis and experimental element analysis in this report. The CQ test is a critical instrument for researchers investigating categorical data across interconnected groups, offering useful insights into the presence of large disparities in percentages. The non-parametric character of this method ensures its resilience when the hypotheses of parametric analysis aren't satisfied. Comprehending the fundamental ideas and methodologies of CQ test is crucial for drawing well-informed statistical judgements in the study of categorical data from related groups. The Cochran’s Q test is valuable in handling ambiguous data, providing decision-makers with nuanced insights, and supporting informed choices. Nevertheless, its

conclusions depend on the degree of indeterminacy and are most applicable to data characterized by imprecision, fuzziness, or interval observations. study covers how to take advantage of Cochran's Q test and paired McNemar test to analyze the percentage of replies generated from the findings of Multiple replies Analysis (MRA). This incorporates Cochran's Q procedure on MRA tables of data using a simulated data set. Cochran's Q test discovers whether there is a variation in the percentage of numerous ideas. Q test offers nothing to modify conventional beliefs on heterogeneity, but strong results considerably raise the chance of its occurrence. This underlines the necessity for careful interpretation of statistical heterogeneity in meta-analyses, including both the capacity to identify individual variances and previous assumptions about heterogeneity. Given the extensive usage of the Q test in multiple meta-analyses, the suggested technique may give significant information to analysts and readers. However, it is crucial to realize several limitations, such as the specificity of the posterior probability to a single outcome and the absence of consideration for findings of the Q test on other outcomes within the same studies, since heterogeneity may vary across various outcomes.

References

ASLAM, M. 2023. Cochran's Q test for analyzing categorical data under uncertainty. *Journal of Big Data,* 10**,** 147.

BERRY, K. J., JOHNSTON, J. E. & MIELKE JR, P. W. 2007. An alternative measure of effect size for Cochran's Q test for related proportions. *Perceptual and motor skills,* 104**,** 1236-1242.

BIGGERSTAFF, B. J. & JACKSON, D. 2008. The exact distribution of Cochran's heterogeneity statistic in one-way random effects meta-analysis. *Statistics in medicine,* 27**,** 6093-6110.

DAVID, I., ADEHI, M. & IKWUOCHE, P. 2021. Cochran's Q-Test on Soil Helminth Prevalence. *Biometrical Letters,* 58**,** 169-185.

FAHMY, T. & BELLÉTOILE, A. 2017. Algorithm 983: Fast Computation of the Non-Asymptotic Cochran's Q Statistic for Heterogeneity Detection. *ACM Transactions on Mathematical Software (TOMS),* 44**,** 1-12.

KULINSKAYA, E. & DOLLINGER, M. B. 2015. An accurate test for homogeneity of odds ratios based on Cochran's Q-statistic. *BMC medical research methodology,* 15**,** 1-19.

KUMAR, G. & LALITHKUMAR, E. APPLICATION OF COCHRAN'S Q-TEST FOR EVALUATING THE PERFORMANCE OF AN EDUCATIONAL PROGRAM.

MASON, L., OTERO, M. & ANDREWS, A. 2022. Cochran's Q test of stimulus overselectivity within the verbal repertoire of children with autism. *Perspectives on behavior science***,** 1-21.

PEREIRA, T. V., PATSOPOULOS, N. A., SALANTI, G. & IOANNIDIS, J. P. 2010. Critical interpretation of Cochran's Q test depends on power and prior assumptions about heterogeneity. *Research Synthesis Methods,* 1**,** 149-161.

STEPHEN, D. & SAZ, A. 2018. Cochran's Q with pairwise McNemar for dichotomous multiple responses data: A practical approach. *Int J Eng Technol,* 7**,** 4-6.

STAT.Mahammad.Mahmoud@MSc.2024.Non Parametric Statistics

ORIGINALITY REPORT

29%	18%	28%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	journalofbigdata.springeropen.com Internet Source	5%
2	www.researchgate.net Internet Source	4%
3	Thierry Fahmy, Arnaud Bellétoile. "Algorithm 983", ACM Transactions on Mathematical Software, 2017 Publication	3%
4	I. J. David, M. U. Adehi, P. O. Ikwuoche. "Cochran's Q-Test on Soil Helminth Prevalence", Biometrical Letters, 2021 Publication	3%
5	KENNETH J. BERRY. "AN ALTERNATIVE MEASURE OF EFFECT SIZE FOR COCHRAN'S Q TEST FOR RELATED PROPORTIONS", Perceptual and Motor Skills, 2007 Publication	3%
6	Brad J. Biggerstaff. "The exact distribution of Cochran's heterogeneity statistic in one-way	3%

random effects meta-analysis", Statistics in Medicine, 12/20/2008
Publication

7	Elena Kulinskaya, Michael B Dollinger. "An accurate test for homogeneity of odds ratios based on Cochran's Q-statistic", BMC Medical Research Methodology, 2015 Publication	2%
8	wikimili.com Internet Source	1%
9	Simon Voelker, Sebastian Hueber, Christian Corsten, Christian Remy. "HeadReach: Using Head Tracking to Increase Reachability on Mobile Touch Devices", Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020 Publication	1%
10	link.springer.com Internet Source	1%
11	journals.sagepub.com Internet Source	1%
12	Tiago V. Pereira, Nikolaos A. Patsopoulos, Georgia Salanti, John P. A. Ioannidis. "Critical interpretation of Cochran's Q test depends on power and prior assumptions about heterogeneity", Research Synthesis Methods, 2010 Publication	1%

13	onlinelibrary.wiley.com Internet Source	1 %
14	Submitted to Navitas Professional Institute Pty Ltd Student Paper	<1 %
15	www.netlib.org Internet Source	<1 %
16	deepai.org Internet Source	<1 %
17	ir.unimas.my Internet Source	<1 %
18	scirp.org Internet Source	<1 %
19	Hoaglin, David C.. "Misunderstandings about Q and ‘Cochran's Q test' in meta-analysis", Statistics in Medicine, 2015. Publication	<1 %
20	Lee Mason, Maria Otero, Alonzo Andrews. "Cochran’s Q Test of Stimulus Overselectivity within the Verbal Repertoire of Children with Autism", Perspectives on Behavior Science, 2021 Publication	<1 %
21	Wolfgang Viechtbauer. "Confidence intervals for the amount of heterogeneity in meta-analysis", Statistics in Medicine, 2007	<1 %

Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		