NCSS Statistical Software NCSS.com

Chapter 521

Cochran's Q Test

Introduction

This procedure computes the non-parametric Cochran's Q test for related categories where the response is binary. Cochran's Q is used for testing k = 2 or more matched sets, where a binary response (e.g. 0 or 1) is recorded from each category within each subject. Cochran's Q tests the null hypothesis that the proportion of "successes" is the same in all groups versus the alternative that the proportion is different in at least one of the groups.

Cochran's Q test is an extension of the McNemar test to a situation where there are more than two matched samples. When Cochran's Q test is computed with only k = 2 groups, the results are equivalent to those obtained from the McNemar test (without continuity correction). Cochran's Q is also considered to be a special case of the non-parametric Friedman test, which is similar to repeated measures ANOVA and is used to detect differences in multiple matched sets with numeric responses. When the responses are binary, the Friedman test becomes Cochran's Q test.

This procedure also computes two-sided, pairwise multiple comparison tests that allow you to determine which of the individual groups are different if the null hypothesis in Cochran's Q test is rejected. The individual alpha level is adjusted using the Bonferroni method to control the overall experiment-wise error rate.

This procedure is based on the results and formulas given in chapter 26 of Sheskin (2011). We refer you to that chapter for additional information about Cochran's Q test.

Experimental Design

A typical design for this scenario involves N individuals where a binary measurement (e.g., 0 or 1) is made on each individual for each of k categories, where $k \ge 2$. Typical data might appear as

Subject	Condition 1	Condition 2	Condition 3
1	0	1	0
2	1	1	0
3	1	1	1
4	0	0	0
5	1	0	0
6	0	1	1
7	0	0	0
8	1	1	0
9	0	1	0
10	1	1	1
11	0	1	0
12	1	1	0

where, in this case, each subject responds to 3 different conditions with either a Yes (1) or No (0). In **NCSS**, the responses may be coded as either text values (e.g., Yes, No) or numeric values (e.g., 0, 1).

Technical Details

Suppose we have k binary measurements on each of N subjects (where the "subject" may be a set of matched individuals). Let $Y_{i,j}$ be the binary response from subject i in category j (i = 1 to N, j = 1 to k), with success = 1 and failure = 0. Let the proportions, $\pi_1, \pi_2, \ldots, \pi_k$, represent the proportion of "successes" in each of the k groups.

Cochran's Q is used to test the null hypothesis

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

versus the alternative

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

In **NCSS**, these proportions, π_a and π_b , are displayed as percentages.

Assumptions

The Cochran's Q test and associated multiple comparisons require the following assumptions:

- 1. Responses are binary and from *k* matched samples.
- 2. The subjects are independent of one another and were selected at random from a larger population.
- 3. The sample size is sufficiently "large". (As a rule of thumb, the number of subjects for which the responses are not all 0's or 1's, n, should be ≥ 4 and nk should be ≥ 24 . This assumption is not required for the exact binomial McNemar test.)

Cochran's Q Test Statistic

For binary responses, $Y_{i,i}$, in k matched groups from N subjects, the Cochran's Q test statistic is computed as

$$Q = \frac{(k-1)[kC - T^2]}{kT - R}$$

where

$$C = \sum_{j=1}^{k} \left(\sum_{i=1}^{N} Y_{i,j} \right)^2$$

$$T = \sum_{i=1}^{N} \left(\sum_{j=1}^{k} Y_{i,j} \right)$$

$$R = \sum_{i=1}^{N} \left(\sum_{j=1}^{k} Y_{i,j} \right)^2$$

For "large" samples, the test statistic, Q, is distributed as chi-square with k-1 degrees of freedom. As in the McNemar test, only subjects who do not have the same response in all categories contribute to the overall Q statistic.

The p-value for the test is computed as

P-Value =
$$Pr(Q > \chi^2_{1-\alpha,k-1})$$

where $\chi^2_{1-\alpha,k-1}$ is the value of the $(1-\alpha)$ quantile of the chi-square distribution with k-1 degrees of freedom.

Multiple Comparisons

When the null hypothesis of success proportion equality is rejected by Cochran's Q test, you can proceed to determine which of the groups are different by computing multiple pairwise comparisons.

Pairwise tests between groups "a" and "b" test the null hypothesis

$$H_0: \pi_a = \pi_b$$

versus the alternative

$$H_A$$
: $\pi_a \neq \pi_b$

In **NCSS**, these proportions, π_a and π_b , are displayed as percentages.

NCSS provides two methods as described in chapter 26 of Sheskin (2011). The first method identifies the minimum required difference, *MRD*, needed to declare a pair of experimental conditions as significantly different. The second method simply employs pairwise McNemar tests among groups to find significant differences.

Both multiple comparison methods use the Bonferroni alpha adjustment to control the overall experiment-wise error of the tests. The adjustment simply divides the overall required alpha, α , by the number of pairwise tests, c, where

$$c = \frac{k(k-1)}{2}.$$

The alpha-level for each individual test, α_{adi} , is

$$\alpha_{adj} = \frac{a}{c}.$$

For example, if the desired overall alpha were 0.05 and three groups were being compared, then the individual alpha level for each test would be 0.05/[3(2)/2] = 0.05/3 = 0.0167.

Minimum Required Difference

For sufficiently large sample sizes (i.e., $n \ge 4$ and $nk \ge 24$, where n is the number of subjects for which the responses are not all 0's or 1's), the minimum required difference in proportions for any pair of k experimental groups to be declared different is

$$MRD = z_{adj} \sqrt{2 \left[\frac{kT - R}{N^2 k(k-1)} \right]}$$

where N, T, and R are defined as in Cochran's Q statistic, with

$$T = \sum_{i=1}^{N} \left(\sum_{j=1}^{k} Y_{i,j} \right)$$

$$R = \sum_{i=1}^{N} \left(\sum_{j=1}^{k} Y_{i,j} \right)^2$$

and

 z_{adj} is the value of the $(1 - \alpha_{adj}/2)$ quantile from the standard normal distribution.

Two groups are declared to be significantly different with protected overall alpha, α , if their absolute difference in proportions is greater than *MRD*, that is if

$$|\pi_a - \pi_b| > MRD$$

McNemar Tests

The McNemar test statistic for each pair of groups is computed as

$$M = \frac{(n_1 - n_2)^2}{n_1 + n_2}$$

where

 n_1 = number of subjects where group "a" response = 0 and group "b" response = 1.

 n_2 = number of subjects where group "a" response = 1 and group "b" response = 0.

NCSS Statistical Software NCSS.com

Cochran's Q Test

Large-Sample (Asymptotic)

For sufficiently large sample sizes (i.e., $n \ge 4$ and $nk \ge 24$, where n is the number of subjects for which the responses are not all 0's or 1's), the test statistic, M, is asymptotically distributed as chi-square with 1 degree of freedom. The p-value for the individual test with protected overall alpha, α , is computed as

P-Value =
$$\Pr\left(M > \chi^2_{1-\alpha_{adj},1}\right)$$

where $\chi^2_{1-\alpha_{adj},1}$ is the value of the $\left(1-\alpha_{adj}\right)$ quantile of the chi-square distribution with 1 degree of freedom.

Exact Test

Exact p-values for the McNemar test can also be computed by enumerating and summing individual binomial probabilities of results more extreme than the observed. The exact test results are more accurate than the asymptotic test results because there is no approximation.

A Note on the Power of the Two Multiple Comparison Tests

The Minimum Required Difference multiple comparison method uses all of the available information in the data in its calculations, but the multiple McNemar tests comparison method uses just the values from subjects who have different responses for the two categories, not all of the data. For this reason, some argue that the Minimum Required Difference method is more powerful for finding differences than using multiple McNemar tests (see Note #9 on page 1135 of Sheskin (2011)).

Data Structure

The data may be entered in two formats, as shown in the examples below. The examples give the binary responses of 12 subjects to each of 3 experimental conditions.

The first format, shown in the first table below, puts the responses for each group in separate columns; that is, each column contains all responses for each condition. Each row corresponds to a single subject. This format allows for the use of an additional optional frequency variable for summarized data.

The second format, shown in the second table below, arranges the data so that all responses are entered in a single column. A grouping variable contains an index that gives the group (Condition 1, 2, or 3) to which each row of data belongs. The subject variable specifies the individual to which each response belongs. This second format allows you to specify multiple response variables; a separate analysis is carried out for each response variable.

Multiple Response Variables

Subject	Condition 1	Condition 2	Condition 3
1	0	1	0
2	1	1	0
3	1	1	1
4	0	0	0
5	1	0	0
6	0	1	1
7	0	0	0
8	1	1	0
9	0	1	0
10	0	1	0
11	0	1	0
12	0	1	0

Response Variable, Grouping Variable, and a Subject Variable

Subject	Condition	Response
1	1	0
1	2	1
1	3	0
2	1	1
2	2	1
2	3	0
3	1	1
3	2	1
3	3	1
4	1	0
4	2	0
4	3	0
5	1	1
5	2	0
5	3	0
	•	•
	•	•
	•	•

NCSS Statistical Software NCSS.com

Cochran's Q Test

Example 1 – Multiple Response Variables

This section presents an example of how to run an analysis on hypothetical data where the responses are stored in separate columns, one for each category. In this example, a series of physical exams were given to 12 subjects. If the subject passed the exam, a "1" was recorded, otherwise a "0" was recorded for failure. They wish to determine if there is a difference among the three exams.

Setup

To run this example, complete the following steps:

1 Open the PhysExam1 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **PhysExam1** and click **OK**.

2 Specify the Cochran's Q Test procedure options

- Find and open the **Cochran's Q Test** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the Example 1 settings file. To load
 these settings to the procedure window, click Open Example Settings File in the Help Center or File
 menu.

3 Run the procedure

Click the Run button to perform the calculations and generate the output.

Data Summary

Data Summary Rows Processed: 12 Rows with Missing Values: 0 Rows Used in the Analysis: 12 Responses: 2(0, 1)Groups (k): 3 (Exam1, Exam2, Exam3) Subjects or Blocks (N): Number with responses that are not all equal (n): 11 (nk = 33)** Large-sample (asymptotic) test results should be used only if n ≥ 4 and nk ≥ 24. Status: Conditions are met.

The Data Summary report gives a summary description of the data used in the analysis. The summary indicates that the large-sample conditions ($n \ge 4$ and $nk \ge 24$) are met. One subject passed every exam, so his/her responses will not contribute to the Cochran's Q test statistic.

Combined Table

		Resp	onse	
Variable		0	1	Total
Exam1	Count	6	6	12
	% within Group	50	50	100
Exam2	Count	2	10	12
	% within Group	16.67	83.33	100
Exam3	Count	9	3	12
	% within Group	75	25	100
Total	Count	47	40	20
Total	Count % within Group	17 47.22	19 52.78	36 100

This report gives the counts and percentages of each response within each category. Exam 2 has the highest pass rate with 83.33%, while Exam 3 has the lowest pass rate with only 25%.

Cochran's Q Test

Cochran's Q Test

(Exam1 by Exam2 by Exam3)

H0: The proportions of [Response = "1"] in all groups are equal.

H1: The proportion of [Response = "1"] in at least one group is different.

Test	Туре	Chi-Square Statistic Q	DF	Asymptotic P-Value	Reject H0 at α = 0.05?
Cochran's Q	2-Sided	6.7273	2	0.03461	Yes

Cochran's Q test has a p-value of 0.03461, indicating that the success rate for at least one group is different from the others. Because this result is significant, we can proceed to consider the multiple comparison tests.

Multiple Comparisons using Minimum Required Absolute Difference

Multiple Comparisons using Minimum Required Absolute Difference

H0: $\pi_i = \pi_j$ (The proportions of [Response = "1"] in the two groups are equal.) H1: $\pi_i \neq \pi_j$ (The proportions of [Response = "1"] in the two groups are not equal.) Number of Comparisons (c): 3

	Proportions		Absolute	Absolute Difference		
				Minimum	Reject H0 with Overall	
Comparison*	π; (%)	π _j (%)	π _i - π _j	Required	$\alpha = 0.05?†$	
Exam1 vs. Exam2	50.00	83.33	33.33	54.02	No	
Exam1 vs. Exam3	50.00	25.00	25.00	54.02	No	
Exam2 vs. Exam3	83.33	25.00	58.33	54.02	Yes	

^{*} These tests should only be considered if the null hypothesis of equality was rejected by Cochran's Q Test.

These multiple comparison results indicate that exams 2 and 3 are significantly different from each other with an absolute difference of 58.33%.

Multiple Comparisons using the McNemar Test

Multiple Comparisons using the McNemar Test

H0: $\pi_i = \pi_j$ (The proportions of [Response = "1"] in the two groups are equal.) H1: $\pi_i \neq \pi_j$ (The proportions of [Response = "1"] in the two groups are not equal.) Number of Comparisons (c): 3

	Propo	ortions		Chi-Square Statistic			Reject H0 with Overall
Comparison*	π _i (%)	π _j (%)	Test	M	DF	P-Value	$\alpha = 0.05?\dagger$
Exam1 vs. Exam2	50.00	83.33	Asymptotic Binomial Exact	2.0000	1	0.15730 0.28906	No No
Exam1 vs. Exam3	50.00	25.00	Asymptotic Binomial Exact	1.8000	1	0.17971 0.37500	No No
Exam2 vs. Exam3	83.33	25.00	Asymptotic Binomial Exact	5.4444	1	0.01963 0.03906	No No

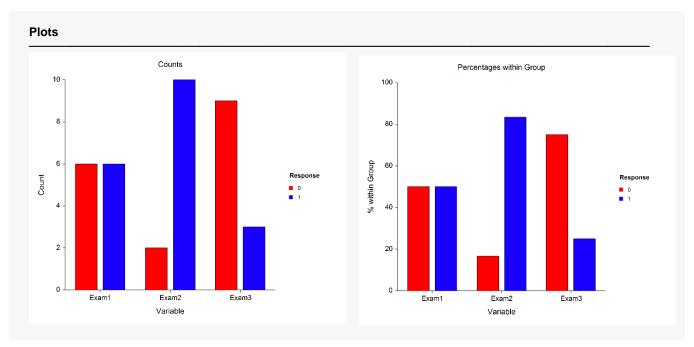
^{*} These tests should only be considered if the null hypothesis of equality was rejected by Cochran's Q Test.

These multiple comparison results indicate that there are no pairs significantly different from one another. It's interesting to note here that no pairs were found to be different even though the overall Cochran's Q test found a significant difference and the multiple comparisons test using minimum required absolute difference found a difference between exams 2 and 3. This is likely because McNemar test has lower overall power because only discordant pairs are used in the computation of the test statistic. The other multiple comparison procedure uses all the data and, thus, has more power.

[†] Individual Comparison Alpha = (Overall Alpha)/c = 0.05/3 = 0.01667.

[†] Individual Comparison Alpha = (Overall Alpha)/c = 0.05/3 = 0.01667.

Plots



This section provides a graphical representation of the counts and percentages for each response within each category.

Example 2 – Multiple Response Variables with a Frequency Variable

Continuing from Example 1, we'll now show you how to analyze data that has been tabulated (summarized) using a frequency variable. The data for this example is exactly the same as that from Example 1 except that it has been summarized. The variable "Count" indicates how many subjects correspond to each series of responses.

Setup

To run this example, complete the following steps:

1 Open the PhysExam2 example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **PhysExam2** and click **OK**.

2 Specify the Cochran's Q Test procedure options

- Find and open the **Cochran's Q Test** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

nput Type	Multiple Response Variables, One Variable per Group Category
Response Variables	Fyam1-Fyam3

3 Run the procedure

• Click the **Run** button to perform the calculations and generate the output.

Output

Rows Processed:	8
Rows with Missing Values:	0
Rows Used in the Analysis:	8
Responses:	2 (0, 1)
Groups (k):	3 (Exam1, Exam2, Exam3)
Subjects or Blocks (N): • Number with responses	12
that are not all equal (n):	11 (nk = 33)*

Cochran's Q Test

(Exam1 by Exam2 by Exam3)

H0: The proportions of [Response = "1"] in all groups are equal.

H1: The proportion of [Response = "1"] in at least one group is different.

Test	Туре	Chi-Square Statistic Q	DF	Asymptotic P-Value	Reject H0 at α = 0.05?
Cochran's Q	2-Sided	6.7273	2	0.03461	Yes

The Data Summary report indicates that only 8 rows were processed this time (12 rows were processed in Example 1), but the number of subjects, 12, is the same as in Example 1 because some of the rows represent more than one individual. The Cochran's Q test results are the same as in Example 1.

Example 3 - Responses All in a Single Column

This section presents an example of how to run an analysis on hypothetical data where the responses are stored in a single column with a separate subject and grouping variable. In this example, responses from 20 individuals are recorded for each of 4 pain relief drugs (A, B, C, and D). One hour after the administration of each drug, each subject was asked whether the medication was effective for them in controlling pain. Responses were recorded as either "Yes" or "No". Drugs were administered in random order, each after a washout period.

Setup

To run this example, complete the following steps:

1 Open the PainDrug example dataset

- From the File menu of the NCSS Data window, select Open Example Data.
- Select PainDrug and click OK.

2 Specify the Cochran's Q Test procedure options

- Find and open the **Cochran's Q Test** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the Example 3 settings file. To load
 these settings to the procedure window, click Open Example Settings File in the Help Center or File
 menu.

Input Type	Response Variable(s), a Grouping Variable, and a Subject
	Variable
Response Variable(s)	Response
Grouping Variable	Drug
Subject Variable	Subject
Reports Tab	
Data Summary Report	Unchecked
Show Combined Table	Checked
Counts	Checked
Percentages within Groups	Checked
Percentages within Responses	Unchecked
Overall Total Percentages	Unchecked
Show Individual Tables	Unchecked
Cochran's Q Test	Checked
Multiple Comparisons using	Unchecked
Minimum Required Absolute Difference	
Multiple Comparisons using the	Unchecked
McNemar Test	
Plots Tab	

Report Options (in the Toolbar)	
Variable Labels	Column Labels

3 Run the procedure

• Click the **Run** button to perform the calculations and generate the output.

Output

Combined Table

		Resp	onse	
Drug		No	Yes	Total
A	Count	7	13	20
	% within Group	35	65	100
В	Count	11	9	20
	% within Group	55	45	100
С	Count	7	13	20
	% within Group	35	65	100
D	Count	14	6	20
	% within Group	70	30	100
T-1-1	0	00	4.4	00
Total	Count % within Group	39 48.75	41 51.25	80 100

Cochran's Q Test

(A by B by C by D)

H0: The proportions of [Response = "Yes"] in all groups are equal.
H1: The proportion of [Response = "Yes"] in at least one group is different.

Test	Туре	Chi-Square Statistic Q	DF	Asymptotic P-Value	Reject H0 at α = 0.05?
Cochran's Q	2-Sided	6.2239	3	0.10121	No

Cochran's Q test indicates that there is not a significant difference among the 4 medications. There is no reason to look at the multiple comparison tests at this point.

Example 4 – Validation of Cochran's Q and Multiple Comparison Tests using Sheskin (2011)

Sheskin (2011) presents an example of computing Cochran's Q test and the associated multiple comparisons in chapter 26, starting on page 1120. The data for the example consists of responses from 12 female subjects about whether or not they would purchase an automobile manufactured by three different companies: Chenesco, Howasaki, and Gemini. The data for this validation example are contained in the dataset called "Sheskin".

Sheskin (2011) computes the group proportions for Response = "1" in Chenesco, Howasaki, and Gemini as 0.25 (25%), 0.75 (75%), and 0.25 (25%), respectively. They compute a Cochran's Q value of 8.0. They do not compute the p-value directly, but state that it is between 0.01 and 0.05. For the multiple comparison test using the minimum required absolute difference with an overall alpha level of 0.05, they find the minimum required difference to be 0.49 (or 49%). They conclude that Howasaki is different from both Chenesco and Gemini, each with differences of 50%, which are both greater than 49%. Sheskin further computes the binomial exact test p-values for the multiple McNemar test comparisons as 0.0312 for Chenesco vs. Howasaki and 0.0704 for Howasaki vs. Gemini. Both are greater than the Bonferroni-adjusted alpha level of 0.0167 so both tests fail to reject the null hypothesis.

The results from **NCSS** match all of these results, with slight difference due to rounding.

Setup

To run this example, complete the following steps:

1 Open the Sheskin example dataset

- From the File menu of the NCSS Data window, select Open Example Data.
- Select **Sheskin** and click **OK**.

2 Specify the Cochran's Q Test procedure options

- Find and open the **Cochran's Q Test** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the Example 4 settings file. To load
 these settings to the procedure window, click Open Example Settings File in the Help Center or File
 menu.

Input Type	Multiple Response Variables, One Variable per Group Category
Response Variables	Chenesco-Gemini

3 Run the procedure

• Click the **Run** button to perform the calculations and generate the output.

Output

Combined Table

		Response		Total
Variable		0	1	
Chenesco	Count	9	3	12
	% within Group	75	25	100
Howasaki	Count	3	9	12
	% within Group	25	75	100
Gemini	Count	9	3	12
	% within Group	75	25	100
Total	Count	21	15	36
	% within Group	58.33	41.67	100

Cochran's Q Test

(Chenesco by Howasaki by Gemini)

H0: The proportions of [Response = "1"] in all groups are equal.

H1: The proportion of [Response = "1"] in at least one group is different.

Test	Туре	Chi-Square Statistic Q	DF	Asymptotic P-Value	Reject H0 at α = 0.05?
Cochran's Q	2-Sided	8	2	0.01832	Yes

Multiple Comparisons using Minimum Required Absolute Difference

H0: $\pi_i = \pi_j$ (The proportions of [Response = "1"] in the two groups are equal.) H1: $\pi_i \neq \pi_j$ (The proportions of [Response = "1"] in the two groups are not equal.)

Number of Comparisons (c): 3

	Prope	ortions	Absolute	Reject H0	
Comparison*	π _i (%)	π _j (%)	π _i - π _j	Minimum Required	with Overall α = 0.05?†
Chenesco vs. Howasaki	25	75	50	48.87	Yes
Chenesco vs. Gemini	25	25	0	48.87	No
Howasaki vs. Gemini	75	25	50	48.87	Yes

^{*} These tests should only be considered if the null hypothesis of equality was rejected by Cochran's Q Test.

[†] Individual Comparison Alpha = (Overall Alpha)/c = 0.05/3 = 0.01667.

Multiple Comparisons using the McNemar Test

H0: $\pi_i = \pi_j$ (The proportions of [Response = "1"] in the two groups are equal.) H1: $\pi_i \neq \pi_j$ (The proportions of [Response = "1"] in the two groups are not equal.) Number of Comparisons (c): 3

	Proportions			Chi-Square Statistic			Reject H0 with Overall α = 0.05?†
Comparison*	π_i (%) π_j (%)		Test	М	DF	P-Value	
Chenesco vs. Howasaki	25	75	Asymptotic Binomial Exact	6.0	1	0.01431 0.03125	Yes No
Chenesco vs. Gemini	25	25	Asymptotic Binomial Exact	0.0	1	1.00000 1.00000	No No
Howasaki vs. Gemini	75	25	Asymptotic Binomial Exact	4.5	1	0.03389 0.07031	No No

^{*} These tests should only be considered if the null hypothesis of equality was rejected by Cochran's Q Test.

The results from **NCSS** match Sheskin (2011), with slight differences due to rounding. Key matched items are highlighted in purple.

[†] Individual Comparison Alpha = (Overall Alpha)/c = 0.05/3 = 0.01667.