McNemar's Test for Paired Data

By: Kenneth Fasano & Elizabeth Hayden

OVERVIEW:
- The standard test considers paired binary response data displayed in a 2X2 contingency table
  - *For example*, if you have twins randomly placed in two treatment groups, control and test, you would then test the two treatments groups on a binary outcome, pass or fail
  - Thereby, there are four possible results for each pair: (1) both the control twin and the test twin fail, (2) the control twin fails while the test twin passes, (3) the control twin passes while the test twin fails, or (4) both the control twin and the test twin pass
- Also, this test is analogous with the paired t-test except for the fact that in this case, each variable is catergorical
- Catergorical variables, such as species or gender, are factors with two or more levels
- McNemar's test may be extended to a 3X3 or higher square tables by expanding the test statistic to include the sum of values obtained from all possible pairs of 2X2 tables

STANDARD MCNEMAR'S TEST:

Assumptions:
1. Paired exactly matched observations are made
2. Each pair is composed of dependent observations, X and Y

Contingency Table:

	Y=1	Y=2
X=1	CONCORDANT	DISCONCORDANT TYPE 1
X=2	DISCONCORDANT TYPE 2	CONCORDANT

To interpret this contingency table, you look at the diagonal cells of paired observations as either:
- CONCORDANT, which refers to an agreement in results between X and Y

  For example, referring back to the twins' case from above, the CONCORDANT results occur when both the control twin and the test twin fail, OR when both the control twin and the test twin pass
  <p align="center">OR</p>
- DISCONCORDANT, which refers to a lack of agreement in results between X and Y

  For example, with the twins' case, the DISCONCORDANT results occur when either the control twin fails while the test twin passes, OR when the control twin passes while the test twin fails

- Let your p-value represent the probability of your DISCONCORDANT TYPE 1 results

Hypotheses:

- *Null Hypothesis* – The p-value is equal to (1/2) which indicates that it is equally probable to obtain DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results

  Your null hypothesis in the twins' case would refer to an EQUAL chance of obtaining an outcome where either the control twin fails while the test twin passes, OR when the control twin passes while the test twin fails

- *Alternative Hypothesis* – The p-value is either greater than or less than (1/2) which indicates that it is NOT equally probable to obtain DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results

  Your alternative hypothesis in the twins' case would refer to an UNEQUAL chance of obtaining an outcome where either the control twin fails while the test twin passes, OR when the control twin passes while the test twin fails

Criterion for Normal Approximation:

- If the number of DISCONCORDANT pairs is greater than or equal to 20, THEN normal approximation may be used
- The number of DISCONCORDANT pairs is the sum of DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results
- If the number is less than 20, then the Exact Test is employed

Test Statistic:

$$\chi = \frac{(n_A - n_B)^2}{n_D}$$

In this equation:

n_A, is equal to the number of DISCONCORDANT TYPE 1 results
n_B, is equal to the number of DISCONCORDANT TYPE 2 results
n_D, is equal to the sum of DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results

- This equation is useful when calculating by hand, but R will calculate this value for you

Test Statistic Corrected for Continuity:

$$\chi_C = \frac{\left(\left| n_A - \frac{n_D}{2} \right| - \frac{1}{2} \right)^2}{\frac{n_D}{4}}$$

- This equation is useful when calculating by hand, but R will calculate this value for you if tell it to correct for continuity

Sampling Distribution of χ and χ_c:
- If assumptions hold and the *Null Hypothesis* is true, then χ and χ_c are distributed according to a chi-squared distribution with one degree of freedom

Critical Value of the Test:
- You must set your Type 1 error value:

  For example, $\alpha = 0.05$

- CV = qchisq(1-α, 1)

Decision Rule:

IF χ > CV THEN REJECT YOUR NULL HYPOTHESIS, OTHERWISE ACCEPT YOUR ALTERNATIVE HYPOTHESIS

IF χ_C >CV THEN REJECT YOUR NULL HYPOTHESIS, OTHERWISE ACCEPT YOUR ALTERNATIVE HYPOTHESIS

Probability Value:
- $P_\chi = (1\text{-pchisq}(\chi,1)$
- $P_C = (1\text{-pchisq}(\chi_C,1))$
- These equations are useful when calculating by hand, but R will calculate these values for you if you tell it to either NOT correct for continuity or to correct for continuity

Exact Test Probability Values:
- These probability values are used when the number of DISCONCORDANT pairs is less than 20
- $n_A < n_D/2$:

$$P_A = 2 * \sum_{k=0}^{n_A} combin(n_D, k) * \frac{1}{2}^{n_D}$$

- $n_A > n_D/2$:

$$P_B = 2 * \sum_{k=n_A}^{n_D} combin(n_D, k) * \frac{1}{2}^{n_D}$$

- $n_A = n_D/2$:

$$P_C = 1$$

- These values are not calculated in R so you must do them by hand when working with a case when the number of DISCONCORDANT pairs is less than 20

EXAMPLE OF A STANDARD MCNEMAR'S TEST:
- Using the example given by Bland (2000), the prevalence of symptoms of severe colds at age 12 and the prevalence of symptoms of severe colds at age 14, among a group of 1319 schoolchildren, were tested against each other

Assumptions:
1. Paired exactly matched observations are made
2. Each pair is composed of dependent observations, X and Y

Contingency Table:

SEVERE COLDS AT AGE 14

SEVERE COLDS AT AGE 12	YES	NO	TOTAL
YES	212	144	356
NO	256	707	963
TOTAL	468	851	1319

To interpret this contingency table, you look at the diagonal cells of paired observations as either:
* CONCORDANT, which refers to an agreement in results between X and Y

    The CONCORDANT results occur when schoolchildren HAVE severe colds at both age 12 and age 14, OR when schoolchildren do NOT have severe colds at both age 12 and age 14

    OR
* DISCONCORDANT, which refers to a lack of agreement in results between X and Y

    The DISCONCORDANT results occur when schoolchildren HAVE severe colds at age 12 but do NOT have severe colds at age 14, OR when schoolchildren do NOT have severe colds at age 12 but HAVE severe colds at age 12

* Let your p-value represent the probability of your DISCONCORDANT TYPE 1 results which in this example represents the schoolchildren who HAVE severe colds at age 12 but do NOT have severe colds at age 14

Hypotheses:
* *Null Hypothesis* – The p-value is equal to (1/2) which indicates that it is equally probable to obtain DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results

    Your null hypothesis is that there is an EQUAL chance of obtaining an outcome where schoolchildren HAVE severe colds at age 12 but do NOT have severe colds at age 14, and when schoolchildren do NOT have severe colds at age 12 but HAVE severe colds at age 12

* *Alternative Hypothesis* – The p-value is either greater than or less than (1/2) which indicates that it is NOT equally probable to obtain DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results

    Your alternative hypothesis is that there is an UNEQUAL chance of obtaining an outcome where schoolchildren HAVE severe colds at age 12 but do NOT have severe colds at age 14, and when schoolchildren do NOT have severe colds at age 12 but HAVE severe colds at age 12

Normal Approximation:
* If the number of DISCONCORDANT pairs is greater than or equal to 20, THEN normal approximation may be used
* The number of DISCONCORDANT pairs is the sum of DISCONCORDANT TYPE 1 results and DISCONCORDANT TYPE 2 results

- The number of DISCONCORDANT TYPE 1 results is equal to 144, while the number of DISCONCORDANT TYPE 2 results is equal to 256
- The sum of 144 and 256 is equal to 400 which is definitely greater than 20
- Therefore, the normal approximation is used, NOT the Exact Test

Test Statistic, Test Statistic Corrected for Continuity, Critical Value of the Test, and Probability Value:
- All of these values can be calculated directly by R

Sampling Distribution and Decision Rule
- Once R calculates the values indicated above, the Sampling Distribution and Decision Rule can be determined

Prototype in R:

 # McNemar's Test without continuity correction
X=matrix(c(212,144,256,707),nrow=2,byrow=T) here you are setting X equal to a matrix using the concatenate function, "c(212,144,256,707)", to order the values by row, "byrow=T", which there are two of, "nrow=2"
X a table will appear that correctly matches the contingency table constructed above:

```
     [,1]        [,2]
[1,] 212        144
[2,] 256        707
```

mcnemar.test(X,correct=F) this is employing the McNemar Test without correcting for continuity

A table will appear:

McNemar's Chi-squared test

data: X
McNemar's chi-squared = 31.36, df = 1, p-value = 2.144e-08

 # To determine the critical value of the test
alpha=0.05 this will determine the stringency of the test
CV=qchisq(1-alpha,1)
CV

 A value will appear:

3.841459

Conclusion:
- Since the results give you a test statistic, χ or McNemar's chi-squared value, equal to 31.36, which is greater than the critical value, CV, equal to 3.841459, then using the Decision Rule it is fair to reject your *Null Hypothesis* and accept your *Alternative Hypothesis*
- Also, with a p-value equal to 2.144e-08, which is significantly less than alpha which is 0.05, it is doubly fair to reject your *Null Hypothesis* and accept your *Alternative Hypothesis*

- Therefore, in accepting your *Alternative* Hypothesis, you are saying that there is an UNEQUAL chance of obtaining an outcome where schoolchildren HAVE severe colds at age 12 but do NOT have severe colds at age 14, and when schoolchildren do NOT have severe colds at age 12 but HAVE severe colds at age 12
- In short, there is a significant difference between DISCONCORDANT TYPE 1 values and DISCONCORDANT TYPE 2 values
- Since you reject your *Null Hypothesis*, according to your Sampling Distribution characteristics, it appears that your Test Statistic is not from a chi-squared distribution with one degree of freedom

 # McNemar's Test with continuity correction
X=matrix(c(212,144,256,707),nrow=2,byrow=T) here you are setting X equal to a matrix using the concatenate function, "c(212,144,256,707)", to order the values by row, "byrow=T", which there are two of, "nrow=2"
X a table will appear that correctly matches the contingency table constructed above:

```
[,1]         [,2]
[1,] 212     144
[2,] 256     707
```

mcnemar.test(X,correct=T) this is employing the McNemar Test with correcting for continuity

A table will appear:

McNemar's Chi-squared test with continuity correction

data: X
McNemar's chi-squared = 30.8025, df = 1, p-value = 2.857e-08

 # To determine the critical value of the test
alpha=0.05 this will determine the stringency of the test
CV=qchisq(1-alpha,1)
CV

A value will appear:

3.841459

Conclusion:
- Since the results give you a test statistic, χ_C or McNemar's chi-squared value, equal to 30.8025, which is greater than the critical value, CV, equal to 3.841459, then using the Decision Rule it is fair to reject your *Null Hypothesis* and accept your *Alternative Hypothesis*
- Also, with a p-value equal to 2.857e-08, which is significantly less than alpha which is 0.05, it is doubly fair to reject your *Null Hypothesis* and accept your *Alternative Hypothesis*
- Therefore, in accepting your *Alternative* Hypothesis, you are saying that there is an UNEQUAL chance of obtaining an outcome where schoolchildren HAVE severe colds at age 12 but do NOT have severe colds at age 14, and when schoolchildren do NOT have severe colds at age 12 but HAVE severe colds at age 12

- In short, there is a significant difference between DISCONCORDANT TYPE 1 values and DISCONCORDANT TYPE 2 values
- Since you reject your *Null Hypothesis*, according to your Sampling Distribution characteristics, it appears that your Test Statistic is not from a chi-squared distribution with one degree of freedom

Reference:
- Bland M (2000) An Introduction to Medical Statistics, 3rd ed. Oxford: Oxford University Press.

EXTENDED MCNEMAR'S TEST FOR HIGHER ORDER TABLES:

Overview:
- This extension should be used for higher order tables (3x3)
- The test statistic includes the sum of the values of the possibilities of all pairs of (2x2) tables within the larger table

Assumptions:
1. Paired data and matched data is taken
2. The X variable and Y variable refer to paired dependent values

Contingency Table:

| | Father's Religion | | | |
Son's Religion	protestent	catholic	jewish	total
protestant	173	20	7	200
catholic	15	51	2	68
jewish	5	3	24	268
total	193	74	33	300

Hypothesis:
- *Null Hypothesis* - is that the discordant results are equally probable, p=(1/2), and there is no difference between treatments
- *Alternative Hypothesis* - is that there is a difference between treatments and the probability, $p \neq (1/2)$

For Normal Approximation:

- May be used when the number of discordant pairs is greater than or equal to twenty

Test Statistic:

$$\sum_i \frac{(n_A - n_B)^{\wedge}2}{n_D}$$

Critical Value:

- You must set your Type 1 error value:

  For example, $\alpha = 0.05$

- df = r (r-1) /2
- CV = qchisq(1-α, df)

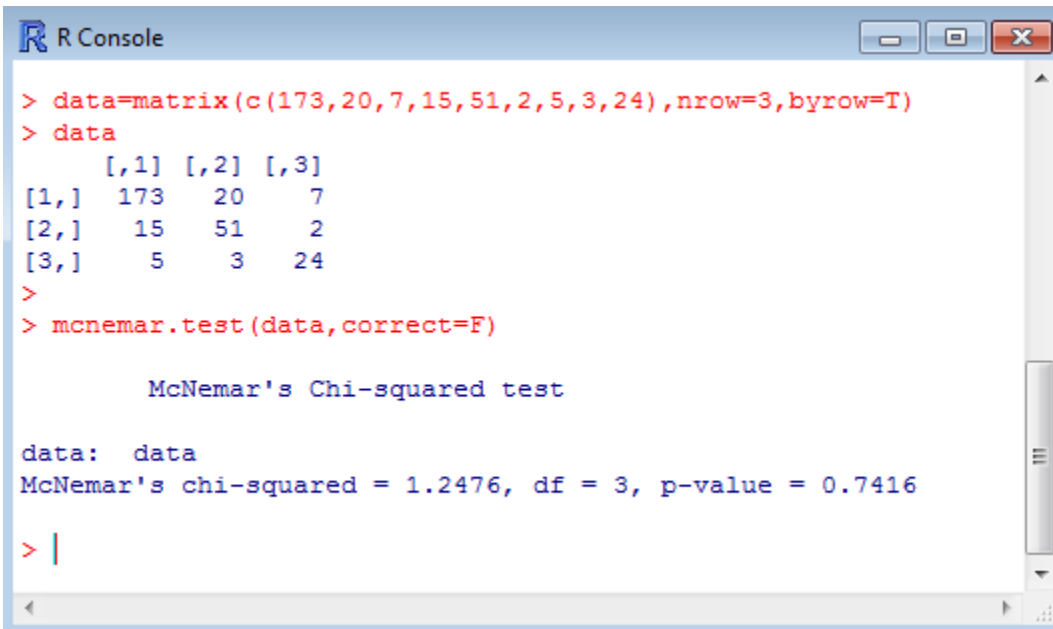*Decision Rule*:

IF χ > CV THEN REJECT YOUR NULL HYPOTHESIS, OTHERWISE ACCEPT YOUR ALTERNATIVE HYPOTHESIS

Probability Value:

- $P_\chi = (1\text{-pchisq}(\chi,\text{df})$
- This equation is useful when calculating by hand, but R will calculate this value for you

For example using the data from the contingency table above:

```
R R Console                                              [_][□][✕]

> data=matrix(c(173,20,7,15,51,2,5,3,24),nrow=3,byrow=T)
> data
      [,1] [,2] [,3]
[1,]  173   20    7
[2,]   15   51    2
[3,]    5    3   24
>
> mcnemar.test(data,correct=F)

        McNemar's Chi-squared test

data:   data
McNemar's chi-squared = 1.2476, df = 3, p-value = 0.7416

> |
```

Conclusion:

- Since p (.7416) > α (.05), you fail to reject the null hypothesis and conclude that there is no difference between the father's religion and the son's religion
- Since the discordant values are equally probable, it shows that it will be equally likely that the son chooses any of the religions that is different from his father's religion