1 the McNemar Test for Significance of Changes

Data The data consist of observation on \( n' \) independent bivariate random variables \((X_i, Y_i), i = 1, 2, \ldots, n'\). The measurement scale for the \( X_i \) and the \( Y_i \) is nominal with two categories, which we call "0" and "1"; that is, the possible values of \((X_i, Y_i)\) are \((0, 0), (0, 1), (1, 0), \) and \((1, 1)\). In the McNemar test the data are usually summarized in a \( 2 \times 2 \) contingency table, as follows.

<table>
<thead>
<tr>
<th>Classification of the ( Y_i )</th>
<th>( Y_i = 0 )</th>
<th>( Y_i = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification ( X_i = 0 )</td>
<td>( a ) (the number of pairs where ( X_i = 0 ) and ( Y_i = 0 ))</td>
<td>( b ) (the number of pairs where ( X_i = 0 ) and ( Y_i = 1 ))</td>
</tr>
<tr>
<td>( X_i = 1 )</td>
<td>( c ) (the number of pairs where ( X_i = 1 ) and ( Y_i = 0 ))</td>
<td>( d ) (the number of pairs where ( X_i = 1 ) and ( Y_i = 1 ))</td>
</tr>
</tbody>
</table>

Assumptions

1. The pairs \((X_i, Y_i)\) are mutually independent.

2. The measurement scale is nominal with two categories for all \( X_i \) and \( Y_i \).

3. The difference \( P(X_i = 0, Y_i = 1) - P(X_i = 1, Y_i = 0) \) is negative for all \( i \), or zero for all \( i \), or positive for all \( i \).

Test Statistic The test statistic for the McNemar test is usually written as

\[
T_1 = \frac{(b - c)^2}{b + c} \tag{1}
\]

However, for \( b + c \leq 20 \), the following test statistic is preferred.

\[
T_2 = b \tag{2}
\]

Null Distribution The null distribution of \( T_1 \) is approximately the chi-squared distribution with 1 degree of freedom when \( (b + c) \) is large. The exact distribution of \( T_2 \) is the binomial distribution with \( p = 1/2 \) and \( n = b + c \).

Hypotheses

\[
H_0 : P(X_i = 0, Y_i = 1) = P(X_i = 1, Y_i = 0) \quad \text{for all } i
\]

\[
H_0 : P(X_i = 0, Y_i = 1) \neq P(X_i = 1, Y_i = 0) \quad \text{for all } i
\]

Let \( n \) equal \( b + c \). If \( n \leq 20 \), use Table A3. If \( \alpha \) is the desired level of significance, enter Table A3 with \( n = b + c \) and \( p = 1/2 \) to find the table entry approximately equal to \( \alpha/2 \).
Call this entry $\alpha_1$, and the corresponding value of $y$ is called $t$. Reject $H_0$ if $T_2 \leq t$, or if $T_2 \geq n - t$, at a level of significance of $2\alpha_1$. Otherwise accept $H_0$. The $p$-value is twice the probability of $T_2$ being less than or equal to the observed value, or greater than or equal to the observed value, whichever is smaller. If $n$ exceeds 20, reject $H_0$ if $T_1$ exceeds the $(1 - \alpha)$ quantile of a chi-squared random variable with 1 degree of freedom. Otherwise accept $H_0$. The $p$-value is the probability of $T_1$ exceeding the observed value.

2 Cox and Stuart Test for Trend

**Data** The data consist of observations on a sequence of random variables $X_1, X_2, \ldots, X_{n'}$, arranged in a particular order, such as the order in which the random variables are observed. It is desired to see if a trend exists in the sequence. Group the random variables into pairs $(X_1, X_{1+c}), (X_2, X_{2+c}), \ldots, (X_{n'-c}, X_{n'})$, where $c = n'/2$ if $n'$ is even, and $c = (n' + 1)/2$ if $n'$ is odd. Replace each pair $(X_i, X_{i+c})$ with a "\(+\)" if $X_i < X_{i+c}$, or a "\(-\)" if $X_i > X_{i+c}$, eliminating ties. The number of untied pairs is called $n$.

This test may be used to detect any specified type of non-random pattern, such as a sine wave or other periodic pattern. The sequence of random variables is merely rearranged so that the smallest numbers, as predicted, will be near the beginning of the sequence and the larger numbers near the end. Then the presence of an upward trend in the rearranged sequence is evidence that the predicted pattern is present in the original arrangement of the sequence.

**Assumptions**

1. The random variables $X_1, X_2, \ldots, X_{n'}$ are mutually independent.

2. The measurement scale of the $X_i$s is at least ordinal.

3. Either the $X_i$s are identically distributed or there is a trend; that is, the later random variables are more likely to be greater than instead of less than the earlier random variables (or vice versa).

**Test Statistic**

\[ T = \text{total number of } +'s \]

**Null Distribution** The null distribution of the test statistic is the binomial distribution with $p = 1/2$ and $n = \text{the number of untied pairs}$, where $X_i$ is not equal to $X_{i+c}$.

**Hypotheses** The rest of test is identical to the sign test presented in the previous section. The null hypothesis is that no trend is present. An upper-tailed test is used to detect an upward trend. A lower-tailed test is used to detect a downward trend. The two-tailed test is used if the alternative hypothesis is that any type of trend (upward or downward) exists.
3 Exercises

3.4 #1, 2, 3, 5

1. One hundred thirty-five citizens were selected at random and were asked to state their opinion regarding U.S. foreign policy. Forty-three were opposed to the U.S. foreign policy. After several weeks, during which they received an informative newsletter, they were again asked their opinion; 37 were opposed, and 30 of the 37 were persons who originally were not opposed to the U.S. foreign policy. Is the change in number of people opposed to the U.S. foreign policy significant?

Solution. Let "1" represent opposing. Then $b = 30$ and $c = 36$. So $n = b + c = 66$.

$$T_1 = \frac{(b - c)^2}{b + c} = 0.54$$

The $p$-value is 0.4601.