Tutorial: Inference for Paired Data using McNemar's Test
Part 1

Consider the following study from Dekkers et al. (2011) that compared two different screening tests for determining adrenal insufficiency. Adrenal insufficiency is a condition in which the adrenal glands do not produce adequate amounts of certain hormones. The screening test involves measuring a patient's cortisol response after administration of an intravenous bolus of adrenocorticotropic hormone (ACTH).

Currently, two doses of ACTH are used for diagnostic purposes in patients with suspected adrenal insufficiency: 1 µg and 250 µg (Dekkers et al. 2011). There is an ongoing debate about which dose should be used for the initial assessment of adrenal function (Dekkers et al. 2011).

The goal of this study was to compare the cortisol response of the 1 µg and 250 µg ACTH test among patients with suspected adrenal insufficiency. Patients with cortisol concentrations of ≥550 nmol/l after ACTH stimulation (considered normal cortisol response) were classified as not having adrenal insufficiency. This was a retrospective cohort study whereby patients who received both the 1 µg and 250 µg ACTH test between January 2004 and December 2007 were included for analysis. The data can be found in the `AI.dta` dataset.

Source: Dekkers OM, Timmermans JM, Smit JW, Romijn JA, Pereira AM. Comparison of the cortisol responses to testing with two doses of ACTH in patients with suspected adrenal insufficiency.*Eur J Endocrinol* 2011 Jan;164(1):83-7

1. Since this is paired data, we decide to use McNemar's test. State the null and alternative hypothesis for McNemar's test.

   Null: The proportion of patients classified as having adrenal insufficiency using the 1 µg test is the same as the proportion of patients classified as having adrenal insufficiency using the 250 µg test.
   Alternative: Those proportions are not equal.

   Is this the same as testing that the proportion of patients classified as *not* having adrenal insufficiency using the 1 µg test is the same as the proportion of patients classified as *not* having adrenal insufficiency using the 250 µg test?

2. Use the `table` command to summarize the data.

```
. tabulate one two
```

```
           |          two
       one |  Abnormal     Normal |     Total
-----------+----------------------+----------
  Abnormal |        42         19 |        61
    Normal |        14        132 |       146
-----------+----------------------+----------
     Total |        56        151 |       207
```

   a. How many discordant pairs are there?

19 + 14 = 33

3. Carry out McNemar's test in Stata at the α = 0.05 significance level.

```
. mcc one two


                 | Controls                |
Cases            |   Exposed   Unexposed   |      Total
-----------------+-------------------------+-----------
         Exposed |       132          14   |        146
       Unexposed |        19          42   |         61
-----------------+-------------------------+-----------
           Total |       151          56   |        207

McNemar's chi2(1) =      0.76    Prob > chi2 = 0.3841
Exact McNemar significance probability       = 0.4869

Proportion with factor
        Cases         .705314
        Controls     .7294686     [95% Conf. Interval]
                    ---------     --------------------
        difference  -.0241546     -.0832778   .0349687
        ratio        .9668874      .8962794   1.043058
        rel. diff.  -.0892857     -.2991256   .1205541

        odds ratio   .7368421      .3418529   1.550025   (exact)
```

a. What is the test statistic? Null distribution? P-value?

The test statistic is 0.76. The null distribution of the test statistic is chi-squared with 1 degree of freedom. The p-value is 0.3841. Note: there is an exact test version of McNemar's test based on the binomial distribution leading to a p-value of 0.4869, which was the p-value reported in the paper.

b. What is your conclusion?

Since our p-value is greater than 0.05 we fail to reject the null hypothesis. Thus, we have no evidence that the proportion of patients classified as having adrenal insufficiency using the 1 µg test is different from the proportion of patients classified as having adrenal insufficiency using the 250 µg test.