Porru Ilaria 3/4/2020
Statistics HW 2

Part 1. Reporting a binomial test

According to the null hypothesis prepositional datives and double object datives have the same probability to appear. In our data we can observe that the prepositional datives occurred 501 times on a total of 2360 trials.

To test for this, we ran a binomial test:

```
Source on Save
1  x <- 501
2  n <- (501 + 1859)
3  binom.test(x, n, p <- 0.5)
4
5
6
```

which resulted in

```
        Exact binomial test

data:  x and n
number of successes = 501, number of trials = 2360, p-value < 2.2e-16
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.1959431 0.2293504
sample estimates:
probability of success
             0.2122881
```

We observe that the p-value is < 2.2e-16. Since this value is less than $\alpha = .05$, we reject the null hypothesis. The given datives do not have the same probability to occur.

Part 2. McNemar's test.

```
16
17 PTB <- read.table(file='PTB.tsv', header = TRUE, comment.char = "#")
18
19 NLP4J.correct <- PTB$gold.tag == PTB$NLP4J.tag
20 Stanford.correct <- PTB$gold.tag == PTB$Stanford.tag
21 #WINS OF STANFORD OVER NLP4J
22 x1 <- sum(Stanford.correct & !NLP4J.correct)
23 x1
24
25 #WINS OF NLP4J OVER STANFORD
26 x2 <- sum(NLP4J.correct & !Stanford.correct)
27 x2
28
29 x <- min(x1, x2)
30 n <- x1+x2
31
32 #Binomial
33 binom.test(x,n, p=0.5)
34
35
```

- The number of wins of the Stanford column over the NLP4J column is 943.
- The number of wins of the NLP4J column over the Stanford column is 1016.

- The p-value we obtained is not significant with α = .05, so we can't say that one tagger is significantly better than the other.

```
        Exact binomial test

data:  x and n
number of successes = 943, number of trials = 1959, p-value = 0.1038
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.459029 0.503763
sample estimates:
probability of success
              0.481368

>
```