

Aaron Goyzueta
LING 82100
HW 2

Part 1

2,360 occurrences of the dative form were found in a large corpus of American English phone conversation, of which 501 were prepositional dative and 1,859 were double object constructions. This was significant at $\alpha = .05$, shown by performing a binomial test. $p < .001$, 95% confidence interval = .196, .229.

R-commands:

```
x <- 501
n <- 501 + 1859
binom.test(x, n, .5)
```

Part 2

Reading in the data:

```
path <- "/Users/aarongoyzueta/Desktop/CUNY semester 2/Stats w: Kyle/HWS/HW
2/hw02-AaronGoyzueta/PTB.tsv"
df <- read.table(path, sep="\t", header=TRUE)
```

```
Stan.correct <- df$gold.tag == df$Stanford.tag
Nlp.correct <- df$gold.tag == df$NLP4J.tag
```

Stanford wins:

```
x1 <- sum(Stan.correct & !Nlp.correct)
Answer: The Stanford tagger had 943 wins over the NLP4J tagger.
```

NLP4J wins:

```
x2 <- sum(Nlp.correct & !Stan.correct)
Answer: The NLP4J tagger had 1016 wins over the Stanford tagger.
```

McNemar's test:

```
x <- min(x1, x2)
n <- x1 + x2
binom.test(x, n, .5)
```

The test was insignificant at $\alpha = .05$, $p = .104$, 95% confidence interval = .459, .504. Therefore, there is no reason to conclude that one tagger outperformed the other.