

Notes: MS 204 Chapter 2 part II

Overview

- Hypothesis testing framework – single proportion
- Simulating an experiment – single proportion

Motivating example

Lucky or good? Patriots winning coin tosses at nearly impossible rate

The New England Patriots' continued struggles against tired "gate" suffixes reached a new nadir earlier this week when a statistical quirk regarding the team's pregame coin toss win rate became an actual topic of conversation. The Patriots have won 19 of their last 25 pregame coin flips, which is pretty remarkable considering the whole 50:50 odds thing.

```
library(tidyverse)
library(mosaic)
tosses <- c(rep(0, 6), rep(1, 19))
tally( ~ tosses)
```

```
## tosses
##    0    1
##    6   19
```

General question: How likely is it for data as extreme as what is in the sample to occur under the null?

Ex: How likely is it for the Patriots to win 19 or more coin tosses in 25 attempts if the coin toss was really a 50-50 proposition?

Setting the stage

- Study type, variable types, sample/population

Possible claims

Simulation

Additional hypothesis test notes

```
set.seed(0)
tosses <- c(rep(0, 6), rep(1, 19))
tally( ~ tosses)
```

```
## tosses
##  0  1
##  6 19
```

```
mean( ~ tosses)
```

```
## [1] 0.76
```

```
obs <- mean( ~ tosses)
obs
```

```
## [1] 0.76
```

```
rflip(5, prob = 0.5)
```

```
##
## Flipping 5 coins [ Prob(Heads) = 0.5 ] ...
##
## H T T H H
##
## Number of Heads: 3 [Proportion Heads: 0.6]
```

```
do(1)*rflip(25, prob = 0.5)
```

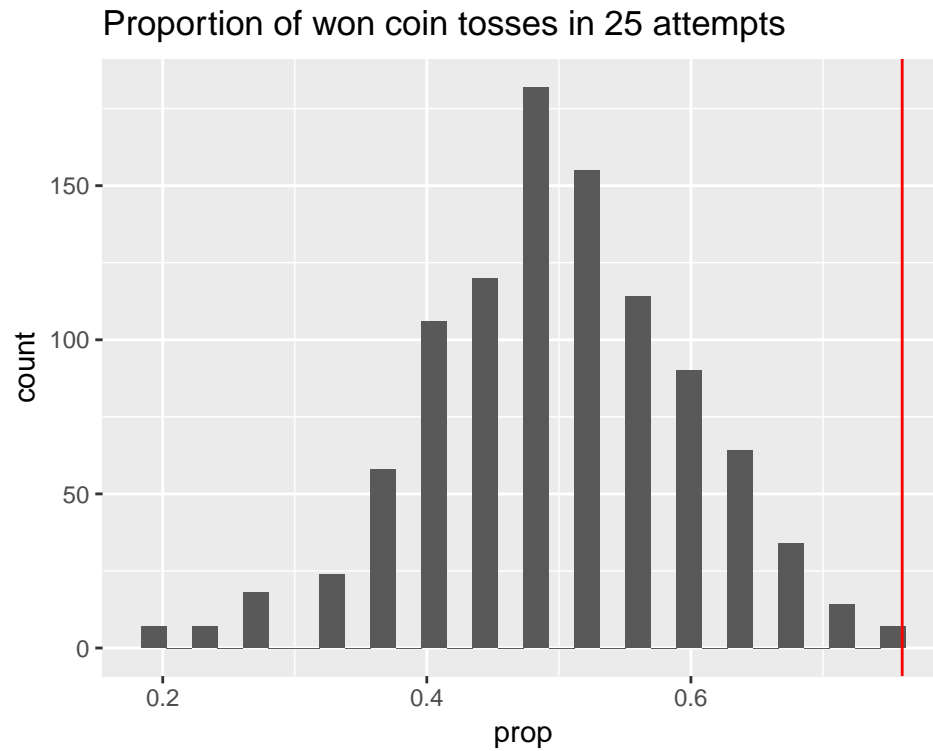
```
##      n heads tails prop
## 1 25      14      11 0.56
```

```
NFL.null <- do(1000)*rflip(25, prob = 0.5)
```

```
NFL.null %>% head()
```

```
##      n heads tails prop
## 1 25      12      13 0.48
## 2 25      10      15 0.40
## 3 25      16       9 0.64
## 4 25      13      12 0.52
## 5 25      14      11 0.56
## 6 25      13      12 0.52
```

```
qplot(x = prop, data = NFL.null) +
  geom_vline(aes(xintercept = obs), colour = "red") +
  ggtitle("Proportion of won coin tosses in 25 attempts")
```



```
NFL.null %>% summarise(p.value = mean(prop >= obs))
```

```
##   p.value
## 1  0.007
```

```
NFL.null %>% summarise(p.value = mean(prop >= obs)*2)
```

```
##   p.value
## 1  0.014
```

Summary