Introduction to hypothesis tests

Overview

Last quick review of confidence intervals

More practice creating confidence intervals in R

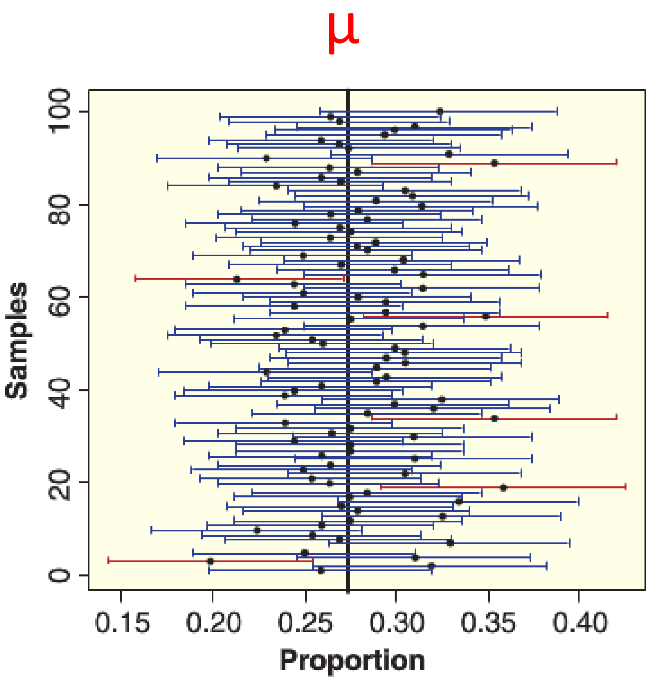
Introduction to hypothesis tests

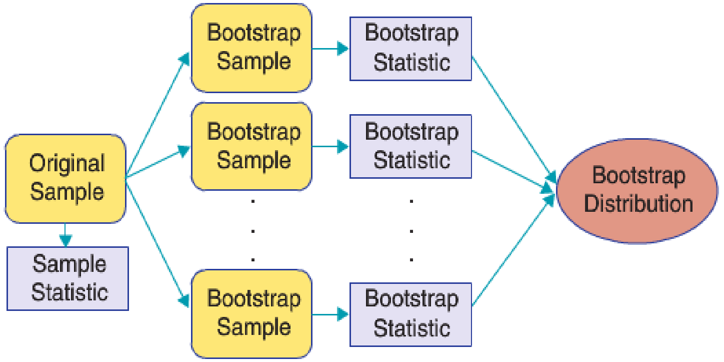
If there is time: Hypothesis tests in R

Very quick review of bootstrap confidence intervals

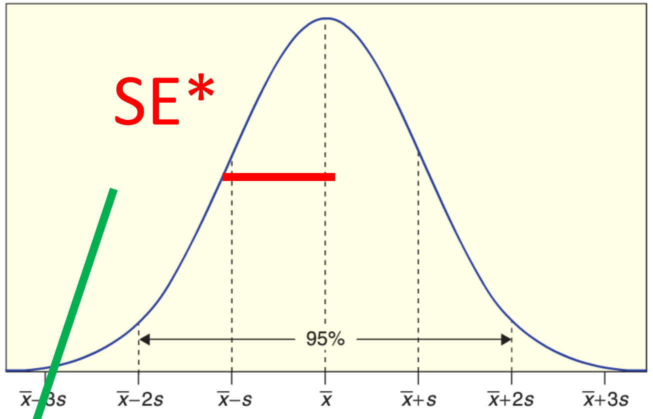
Very quick review: The big picture

Confidence intervals give range of plausible values for a parameter





Bootstrap distribution



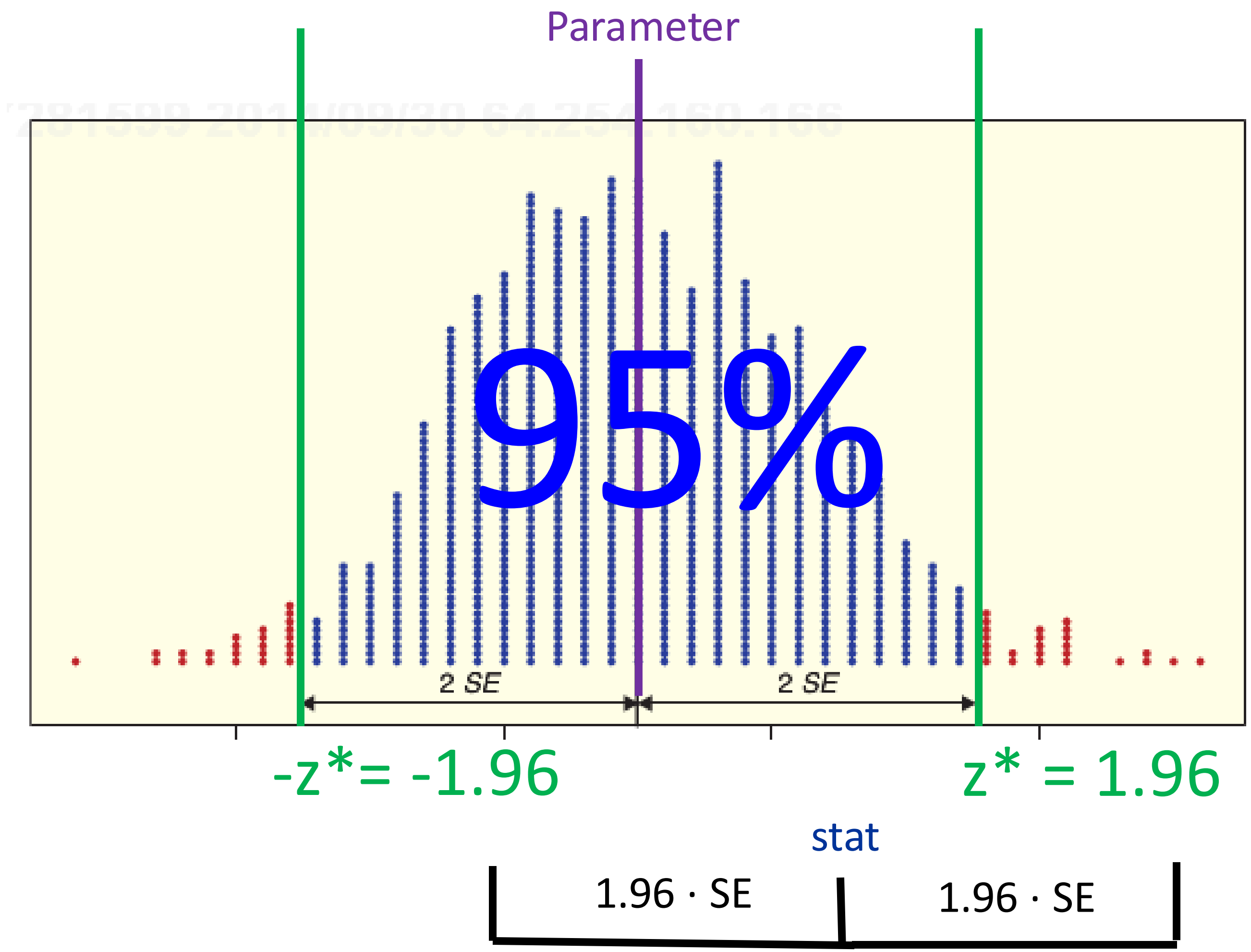
One sample of data

95% confidence interval formula: stat $\pm$ 2 $\cdot$ $\hat{SE}$

Critical value

For a 95% CI, we choose $z^* = 2$

- 95% of our statistics lie within
  -2 and 2 SE of the parameter
  - (1.96 to be more precise)




```
z_star <- mosaic::cnorm(.95)
```

1.96

Critical value z^*

confidence interval: stat $\pm z^* \cdot \hat{SE}$

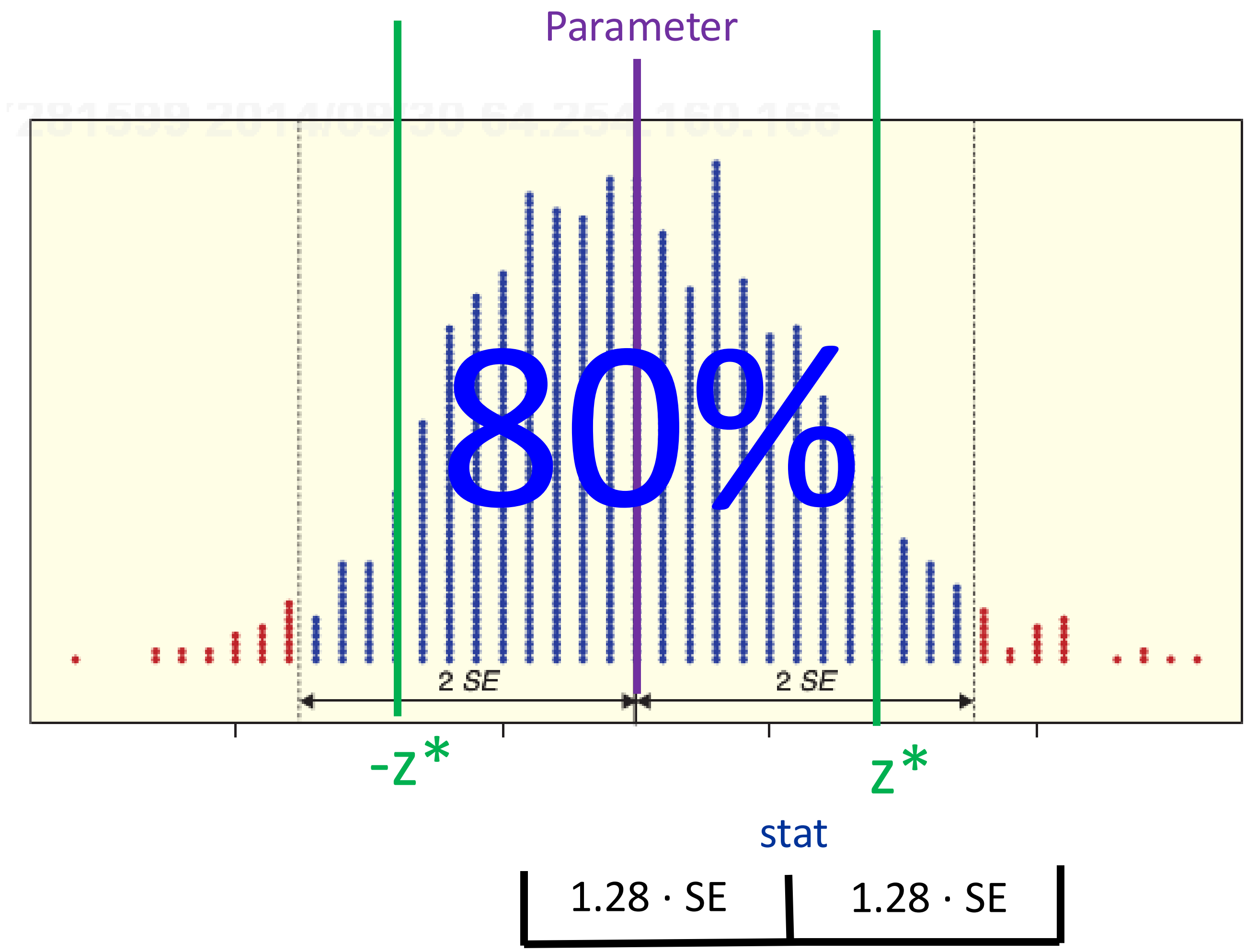
Critical value

We can choose critical values z* for other confidence levels

- E.g. for an 80% CI, we choose z* so that 80% of our statistics lie within -z* and z* SE of the parameter

```
z_star <- mosaic::cnorm(.80)
```

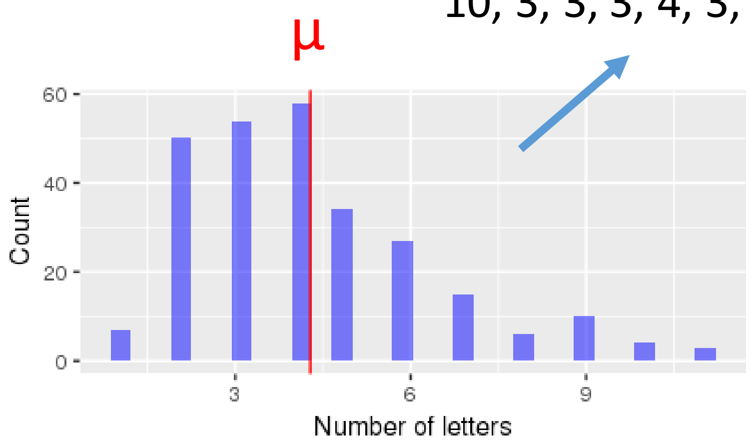
1.28



Critical value z*

confidence interval: stat ± z* $\cdot \hat{SE}$

Review: Using the Bootstrap to get SE*



<u>The sample (n = 10)</u>
10, 3, 3, 3, 4, 3, 2, 6, 4, 5

3, 3, 3, 5, 3,
4, 5, 2, 2, 10

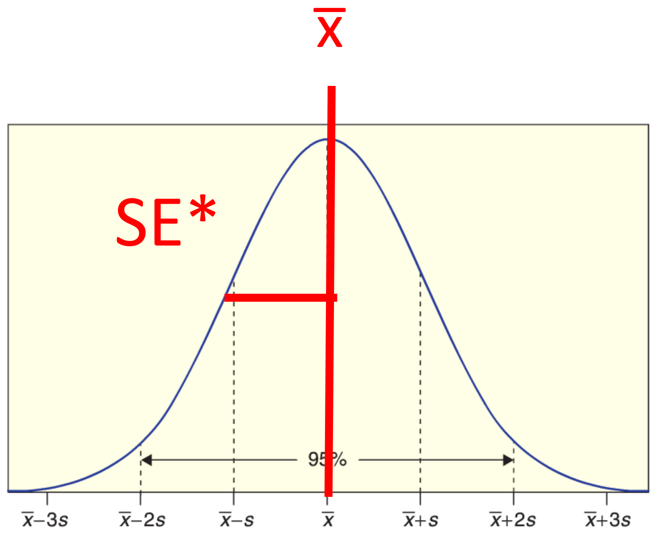
$\overline{x}^* = 4$

3, 3, 2, 3, 6,
4, 6, 5, 3, 6

$\overline{x}^* = 4.1$

5, 3, 2, 3, 3,
3, 10, 3, 4, 3

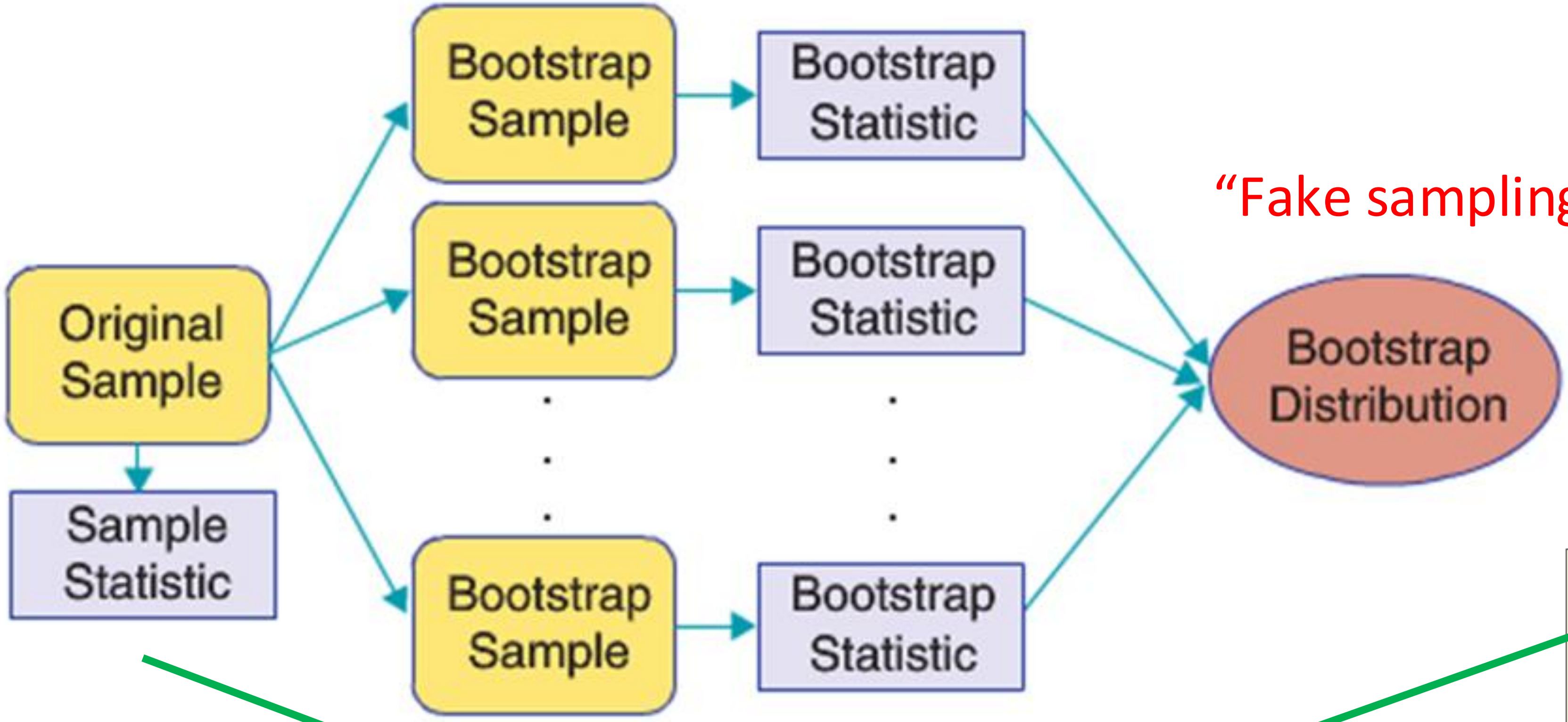
$\overline{x}^* = 3.9$



Bootstrap distribution!

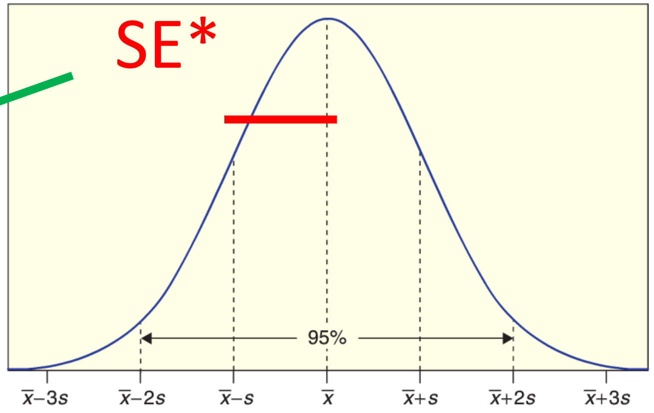
Confidence interval: stat $\pm$ $z^* \cdot$ SE*

Using the bootstrap to estimate the standard error (SE*)



"Fake sampling distribution"

Bootstrap distribution



Confidence interval: stat ± $z^* \cdot SE^*$

Bootstrap distribution in R for μ

```
my_sample <- c(21, 29, 25, 19, 24, 22, 25, 26, 25, 29)      # n = 10 points here
my_stat <- mean(my_sample)     #  x-bar

bootstrap_dist <-  do_it(10000) * {

        curr_boot <- sample(my_sample , 10, replace = TRUE)
        mean(curr_boot)

}

SE_boot   <-  sd(bootstrap_dist)
z_star <- mosaic::cnorm(.95)
CI     <-    c(my_stat    -   z_star  * SE_boot,        my_stat    +   z_star  * SE_boot)
```

Bootstrap distribution in R for π

```
my_sample <- c("heads", "tails", "tails", "heads", "heads")      # n = 5
my_stat <- get_proportion(my_sample, "heads")     #  p-hat

bootstrap_dist <-  do_it(10000) * {

        curr_boot <- sample(my_sample , 5, replace = TRUE)
         get_proportion(my_sample, "heads")

}

SE_boot   <-  sd(bootstrap_dist)
z_star <- mosaic::cnorm(.95)
CI     <-    c(my_stat    - z_star  * SE_boot,        my_stat    +  z_star  * SE_boot)
```

Introduction to hypothesis tests

A quick note on probability

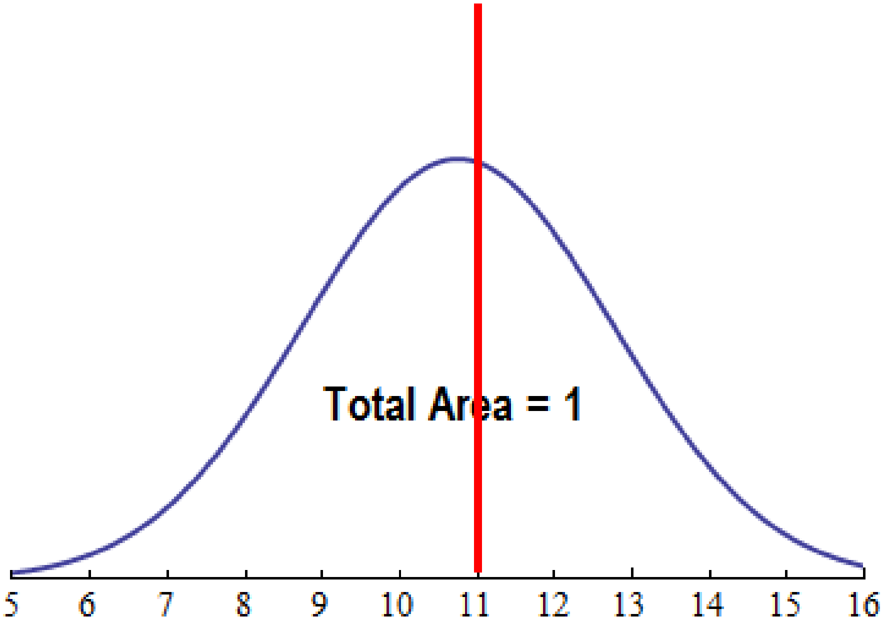
Probability is a way of measuring the likelihood that an event will occur

Probability models assigns a number between 0 and 1 to the outcome of an event (outcome) occurring

We can use a probability model to calculate the probability of an event

For example:

- $P(X < 11) = 0.55$
- $P(X > 20) = 0$



Statistical tests   (hypothesis test)

A **statistical test** uses data from a sample to assess a claim about a population

<u>Example 1</u>: we might make the claim that Trump's approval rating for all US citizens is 42%

How can we write this using symbols?

Statistical tests (hypothesis test)

A **statistical test** uses data from a sample to assess a claim about a population.

<u>Example 2</u>: we might make the claim that the average height of a baseball player is 72 inches

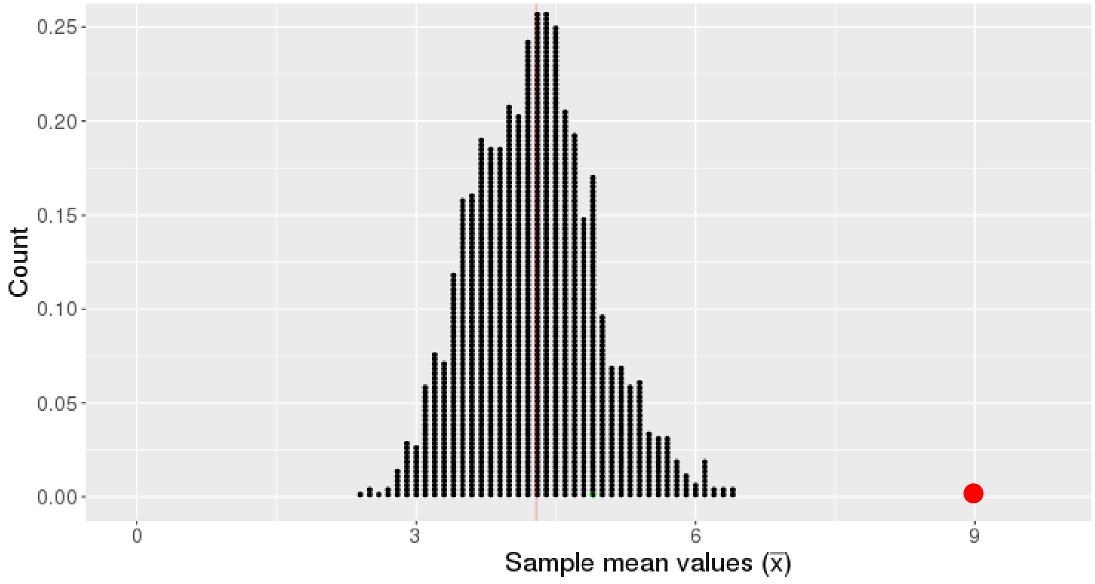
How can we write this using symbols?

Basic hypothesis test logic

We start with a claim about a population parameter

- E.g., $\mu = 4$

This claim implies we should get a certain distribution of statistics



If our observed statistic is highly unlikely, we reject the claim

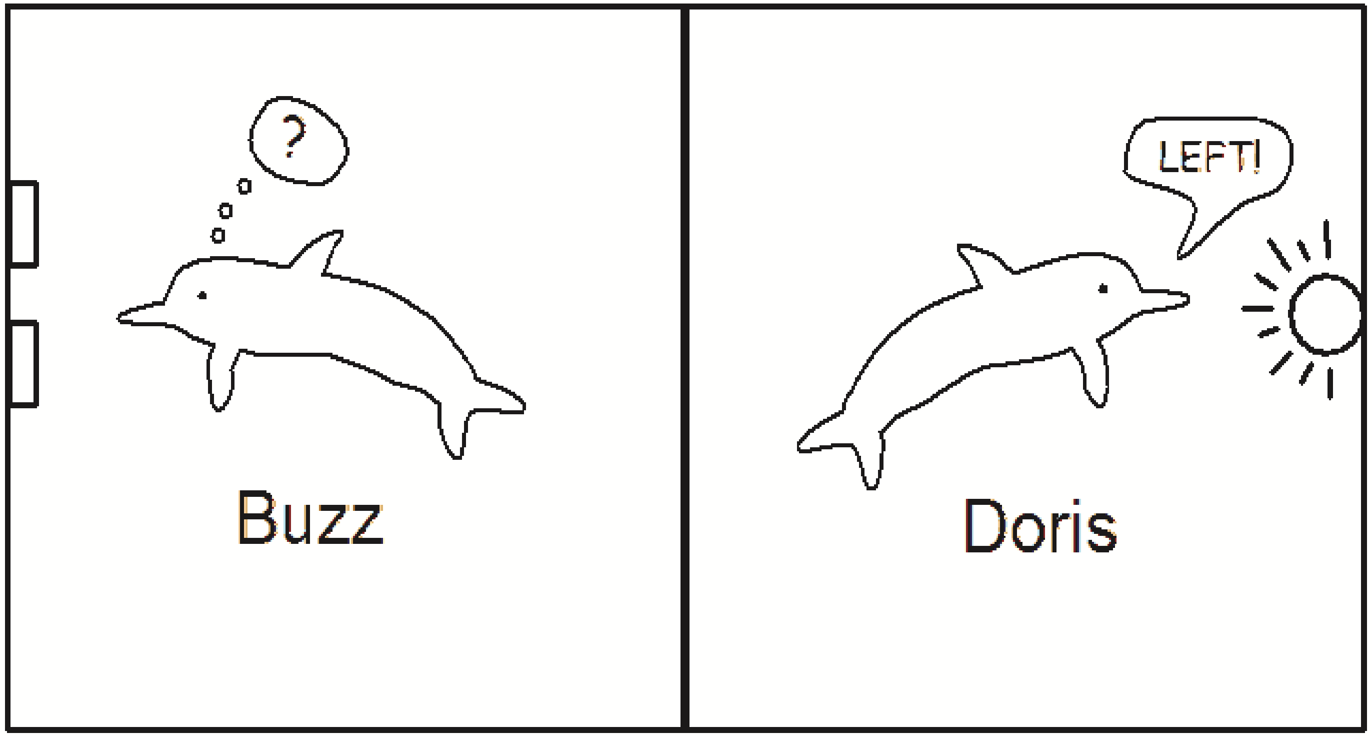
Are dolphins capable of abstract communication?

Dr. Jarvis Bastian is the 1960's wanted to know whether dolphins are capable of abstract communication

He used an old headlight to communicate with two dolphins (Doris and Buzz)

- Stead light = push button on right to get food
- Flashing = push button on the left to get food

A canvas was then put in the middle of the pool with Doris on one side and Buzz on the other



Left

Right

Buzz got 15 out of 16 trials correct

On Canvas, answer the following questions

1. What are the cases here?

2. What is the variable of interest and is it categorical or quantitative?

3. What is the observed statistic - and what symbols should we use to denote it?

4. What is the population parameter we are trying to estimate - and what symbol should we use to denote it?

5. Do you think the results are due to chance?
   - i.e., how many correct answers do you think Buzz would have gotten if he was guessing?

6. Are dolphins capable of abstract communication?

The dolphin communication study

7. If Buzz was just guessing, what would we expect the value of the parameter to be?

8. If Buzz was not guessing, what would we expect the value of the parameter to be?

Chance models

How can we assess whether 15 out of 16 correct trials ($\hat{p}$ = .975) is beyond what we would expect to see by chance?

- i.e., beyond what we would expect to see if $\pi = 0.5$?

If buzz was guessing we can model his guesses as a coin flip:

Heads = correct guess

Tails = incorrect guess

Let's flip a coin 16 times and see how many times we get 15 heads

Chance models

To really be sure, how many repetitions of flipping a coin 16 times should we do?

Any ideas how to do this?

Flipping coins using SDS1000 functions

rflip_count() returns the number of heads out of num_flip coin flips:

```
rflip_count(num_flips, prob = .5)
```

num_flips: the number of times to flip a coin

- 16 for Doris/Buzz

prob: the probability of success on each trial

- .5 if Buzz was guessing

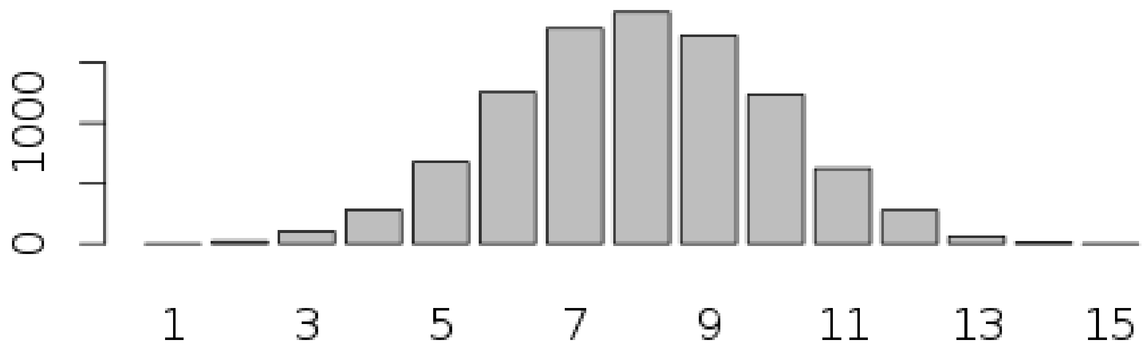
We can repeat flips many times using the do_it() function:

```
library(SDS1000)

flip_simulations <- do_it(10000) * {
    rflip_count(16, prob = .5)
}
```

Simulating Flipping 16 coins 10,000

0	0
1	1
2	22
3	105
4	283
5	679
6	1257
7	1786
8	1920
9	1726
10	1238
11	623
12	279
13	63
14	15
15	3
16	0



Q10: Is it likely that Buzz was guessing?

Q11: Are dolphins capable of abstract communication?

Let’s try it in R!