TYPES OF INFERENCES

~~1. IDENTIFYING THE POPULATION MEAN~~

~~2. IDENTIFYING THE POPULATION %~~

~~3. VERIFYING WHETHER THE POPULATION MEAN IS EQUAL TO A CERTAIN VALUE~~

~~4. VERIFYING WHETHER THE POPULATION % IS EQUAL TO A CERTAIN VALUE~~

5. VERIFYING WHETHER 2 POPULATION MEANS ARE DIFFERENT

6. VERIFYING WHETHER 2 POPULATION % ARE DIFFERENT

TYPES OF INFERENCES

~~1. IDENTIFYING THE POPULATION MEAN~~

~~2. IDENTIFYING THE POPULATION %~~

~~3. VERIFYING WHETHER THE POPULATION MEAN IS EQUAL TO A CERTAIN VALUE~~

~~4. VERIFYING WHETHER THE POPULATION % IS EQUAL TO A CERTAIN VALUE~~

~~5. VERIFYING WHETHER 2 POPULATION MEANS ARE DIFFERENT~~

6. VERIFYING WHETHER 2 POPULATION % ARE DIFFERENT

6. VERIFYING WHETHER 2 POPULATION % ARE DIFFERENT

CASE STUDY: CUSTOMER SURVEYS

CUSTOMERS ARE WEIRD ANIMALS

CUSTOMERS ARE WEIRD ANIMALS

CUSTOMER OBSESSION

IS A MANTRA

AT MOST OF THE SUCCESSFUL COMPANIES TODAY

CUSTOMER SATISFACTION SURVEYS

HELP COMPANIES KEEP TRACK OF

WHAT THEIR CUSTOMERS ARE THINKING

MR. N IS A CONSULTANT

MR. N IS A CONSULTANT

HE'S BEEN HIRED BY A COMPANY F
WHICH HAS BOTH A MOBILE APP AND A WEBSITE

HIS TASK IS TO DETERMINE

ARE CUSTOMERS WHO USE THE APP MORE SATISFIED THAN THOSE WHO USE THE WEBSITE?

STEP : 1

SET UP A CONTROLLED EXPERIMENT

MR. N PICKS 2 SAMPLES

A = 100 PEOPLE WHO USE THE APP

B = 100 PEOPLE WHO USE THE WEBSITE

HE ASKS ALL OF THEM

" WOULD YOU RECOMMEND OUR COMPANY TO YOUR FRIENDS AND FAMILY?"

STEP : 2

COMPUTE SAMPLE STATISTICS FOR BOTH GROUPS A, B

GROUP

A

% WHO SAY YES = 67%

$$SD = SE = \sqrt{\frac{67\% * 33\%}{100}}$$

= 0.5 % PTS

GROUP

B

% WHO SAY YES = 63%

$$SD = SE = \sqrt{\frac{63\% * 37\%}{100}}$$

= 0.5 % PTS

STEP : 3

A TEST OF SIGNIFICANCE COMPARING 2 POPULATION %

NULL HYPOTHESIS

ALL VARIATIONS OBSERVED ARE DUE TO CHANCE I.E. A FLUKE

ALTERNATIVE HYPOTHESIS

THE VARIATIONS OBSERVED CANNOT JUST BE EXPLAINED BY CHANCE

NULL HYPOTHESIS VS ALTERNATIVE HYPOTHESIS

A TEST OF SIGNIFICANCE WILL TELL YOU WHICH OF THESE IS BETTER

A TEST OF SIGNIFICANCE

NULL HYPOTHESIS VS ALTERNATIVE HYPOTHESIS

THIS INVOLVES

1) COMPUTING A TEST STATISTIC

SOME VARIABLE WHOSE PROBABILITY DISTRIBUTION IS KNOWN

A TEST OF SIGNIFICANCE

NULL HYPOTHESIS VS ALTERNATIVE HYPOTHESIS

THIS INVOLVES

1) COMPUTING A TEST STATISTIC

2) COMPUTE THE PROBABILITY IF THE NULL HYPOTHESIS IS TRUE

A TEST OF SIGNIFICANCE

NULL HYPOTHESIS VS ALTERNATIVE HYPOTHESIS

THIS INVOLVES

1) COMPUTING A TEST STATISTIC

2) COMPUTE THE PROBABILITY IF THE NULL HYPOTHESIS IS TRUE

3) IF THE PROBABILITY IS TOO LOW, REJECT THE NULL HYPOTHESIS, ELSE ACCEPT IT

A 67%

B 63%

SAMPLE % ARE DIFFERENT

NULL HYPOTHESIS

THE DIFFERENCE BETWEEN A AND B IS DUE TO CHANCE

ALTERNATIVE HYPOTHESIS

THE DIFFERENCE BETWEEN A AND B IS REAL

STEP : 3 PERFORM A TEST OF SIGNIFICANCE

STEP : 3A

COMPUTE A TEST STATISTIC

WE'LL USE THE Z-STATISTIC

Z-STATISTIC FOR THE NULL HYPOTHESIS

SAMPLE % A

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

WE'LL USE THE Z-STATISTIC

Z-STATISTIC FOR THE NULL HYPOTHESIS

SAMPLE % B

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

WE'LL USE THE Z-STATISTIC

Z-STATISTIC FOR THE NULL HYPOTHESIS

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

SQUARE OF STANDARD ERROR A

WE'LL USE THE Z-STATISTIC

Z-STATISTIC FOR THE NULL HYPOTHESIS

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

SQUARE OF STANDARD ERROR B

WE'LL USE THE Z-STATISTIC

Z-STATISTIC FOR THE NULL HYPOTHESIS

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

A 67% STANDARD ERROR : 0.5

B 63% STANDARD ERROR : 0.5

67 63

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

0.5^2 0.5^2

= 5.7

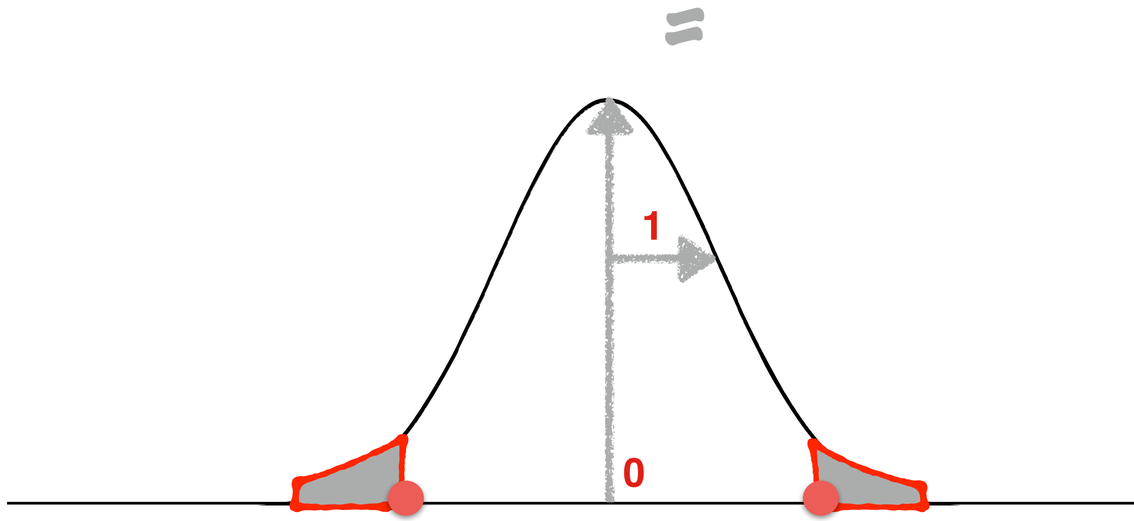
COMPUTE THE VALUE FOR THE NULL HYPOTHESIS

STEP : 3 PERFORM A TEST OF SIGNIFICANCE

STEP : 3B

COMPUTE THE PROBABILITY IF
THE NULL HYPOTHESIS IS TRUE

$P(Z > 5.7)$



1 SIDED TEST (SINCE WE ARE CHECKING A>B NOT JUST INEQUALITY)

NULL HYPOTHESIS

ALL VARIATIONS OBSERVED ARE DUE TO CHANCE I.E. A FLUKE

ALTERNATIVE HYPOTHESIS

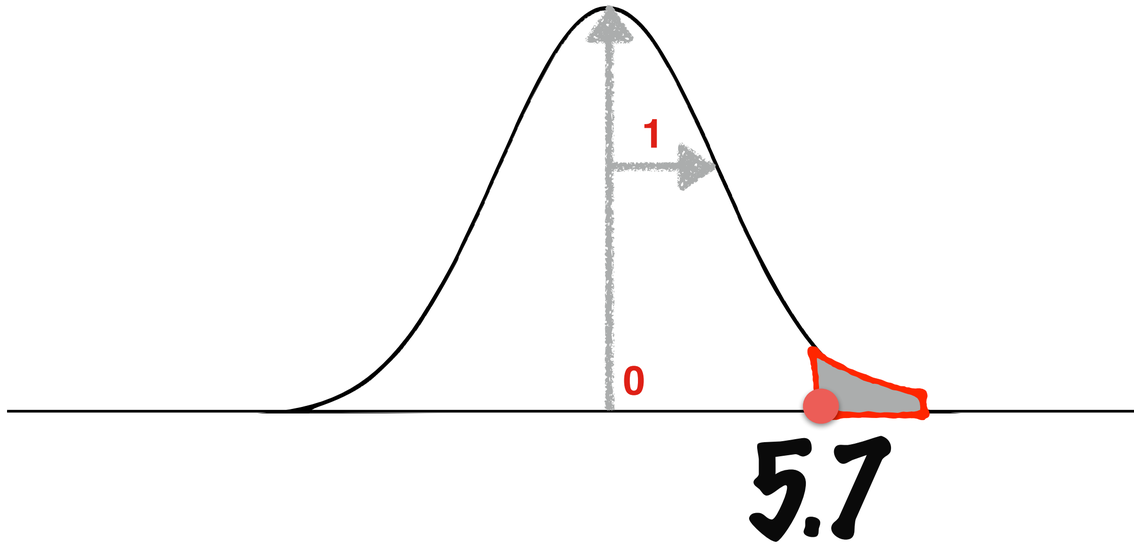
THE VARIATIONS OBSERVED CANNOT JUST BE EXPLAINED BY CHANCE

P-VALUE

P(Z > 5.7)

= AREA UNDER THE CURVE ABOVE 5.7

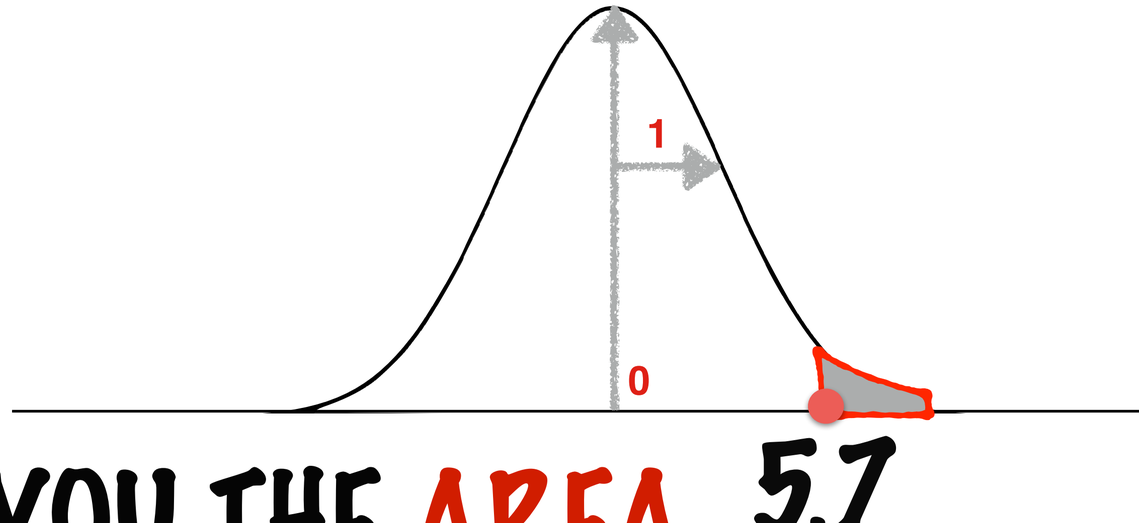
= ?



1 SIDED TEST (SINCE WE ARE CHECKING A>B NOT JUST INEQUALITY)

P-VALUE

$P(Z > 5.7)$



FUNCTION IN R

PNORM() IN R WILL TELL YOU THE AREA UNDER THE CURVE FROM -INF TO Z (CUMULATIVE DISTRIBUTION FUNCTION)

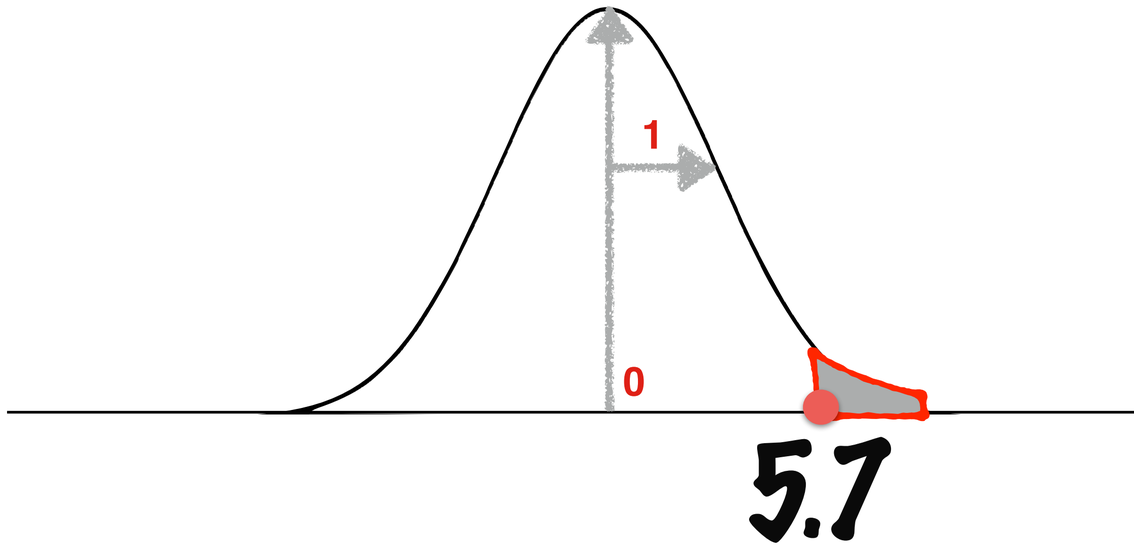
```
> 1-pnorm(5.7)

[1] 5.990371e-09
```

P-VALUE

$P(Z > 5.7)$

$= 5.9\ E\text{-}9$



STEP : 3 PERFORM A TEST OF SIGNIFICANCE

STEP : 3C

IF THE PROBABILITY IS TOO LOW, REJECT THE NULL HYPOTHESIS, ELSE ACCEPT IT

P-VALUE = 5.99 E-9

SINCE THE P-VALUE IS TOO LOW

THE NULL HYPOTHESIS IS REJECTED

TYPES OF INFERENCES

~~1. IDENTIFYING THE POPULATION MEAN~~

~~2. IDENTIFYING THE POPULATION %~~

~~3. VERIFYING WHETHER THE POPULATION MEAN IS EQUAL TO A CERTAIN VALUE~~

~~4. VERIFYING WHETHER THE POPULATION % IS EQUAL TO A CERTAIN VALUE~~

~~5. VERIFYING WHETHER 2 POPULATION MEANS ARE DIFFERENT~~

6. VERIFYING WHETHER 2 POPULATION % ARE DIFFERENT