Hypothesis Testing

Good Afternoon :)

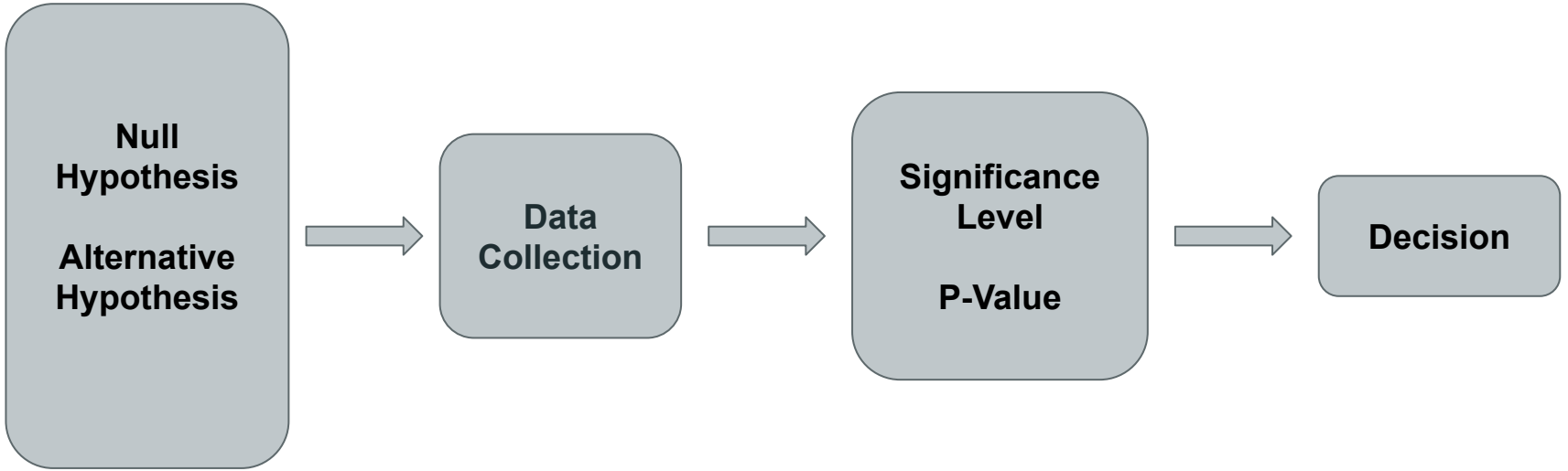
Statistical Inference

Confidence Intervals

Hypothesis Testing

What is Hypothesis Testing?

Hypothesis Testing is a mathematically rigorous way of making a yes-no decision



Step 1: A Real World Problem

I have a friend Xiao Ming, who always takes days to reply my Telegram messages.

I'm beginning to think that he **only** does this to me.

We need to make a **data-driven** decision

- Yes, get upset at XM
- No, don't get upset

Step 1: Null Hypothesis

Assume XM is innocent.

Being good data scientists, we want to gather data, then use this data to make a decision.

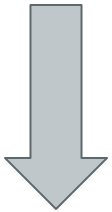
Random Error or Bias?

- Mismatch in timings? Maybe I always text him when he is busy with other things
- His phone just has problems receiving texts from me specifically?
- Maybe my messages are so thought-provoking that he needs more time to think?
- Maybe he secretly does not like me but has been struggling to say it?
- Maybe I am overthinking?

The best way to objectively test this is to see if XM does this to others.

Step 1: Null Hypothesis

Assume XM is innocent.



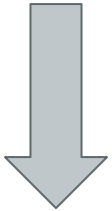
H_0 : Assume XM replies everyone **equally** slowly.

Or

H_0 : Average Time (Me) = Average Time (Others)

Step 1: Alternative Hypothesis

XM is not innocent.



H_1 : XM replies me slowly, with respect to everyone else.

Or

H_1 : Average Time (Me) > Average Time (Others)

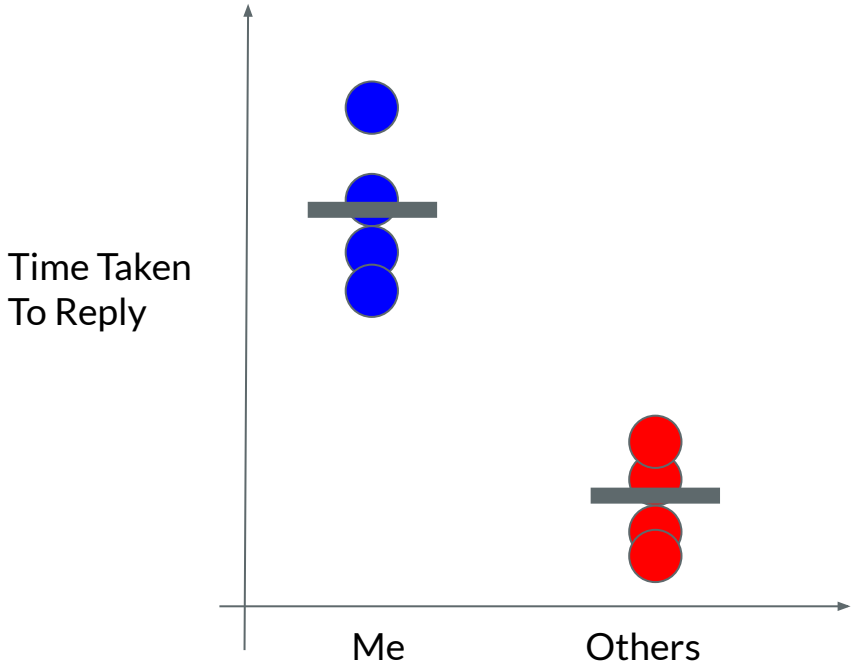
Step 1: Hypothesis Testing Statement

H_0 : Average Time (Me) = Average Time (Others)

H_1 : Average Time (Me) > Average Time (Others)

Note: Innocent until proven guilty

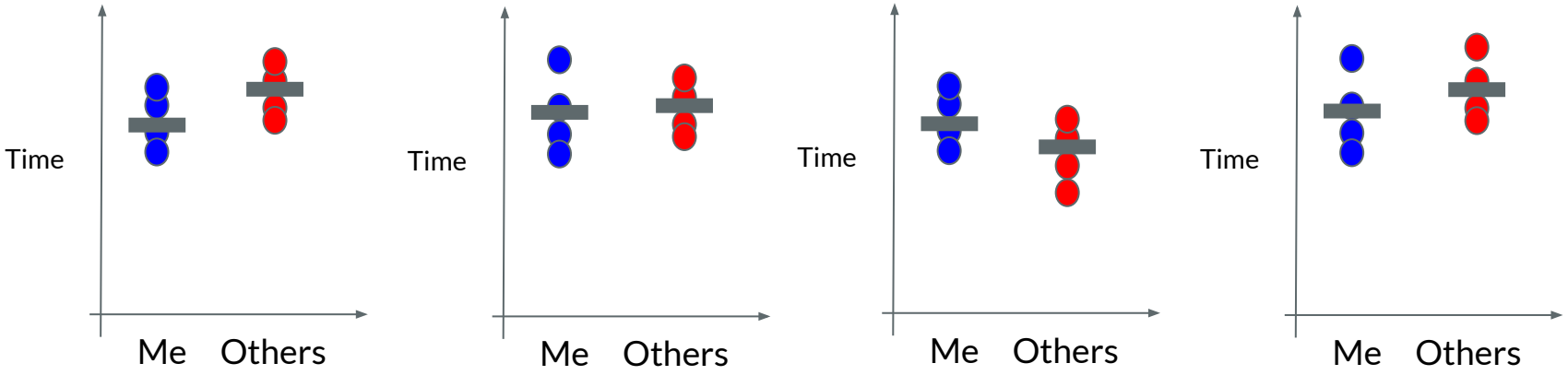
Step 2: Data Collection [Simplistic]



I have proven that he replies me slowly!

Not Yet

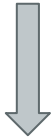
Step 2: Data Collection [Rigorous]



X 100

Step 3: Significance Level and P-Value

The p-value is the probability of obtaining a test result at least as extreme as the result we observed, assuming the null hypothesis is true.



The p-value is the probability that the slow replies to me happened on their own, **assuming** that XM is innocent.

Step 3: Significance Level and P-Value

P-value = P(XM replies me slowly | XM innocent)

- What is the probability that we observed this data, **assuming** our null hypothesis is true?
- We want this probability to be as small as possible, therefore we set a threshold
- **Threshold = Level of Significance (α)**

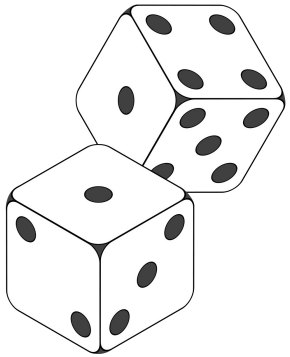
Step 3: Significance Level and P-Value

P-Value: 0.03< Level of Significance (α): 0.05

Step 4: Decision

- We can only decide if XM is guilty or not guilty, but never can we decide that he is innocent.
- The P-Value is less than the threshold
- XM did not intentionally reply me slowly
- There is not enough evidence to reject the null hypothesis that he is innocent.
- There is not enough evidence to prove that XM is guilty.

A (Tougher) Real World Problem



Step 1: Problem

I believe that a player brought a particular die that is “loaded” and thus biased towards one of the six sides, for example, the side showing the digit 6.

Any toss of the die would more likely show 6 rather than the other five sides.

How do I prove it?

Step 1: Hypothesis Testing Statement

H_0 : “The die is fair”

$P(X)=\frac{1}{6}$ for $X=\{1,2,3,4,5,6\}$

H_1 : “The die is biased towards 6”

$P(6) > \frac{1}{6}$

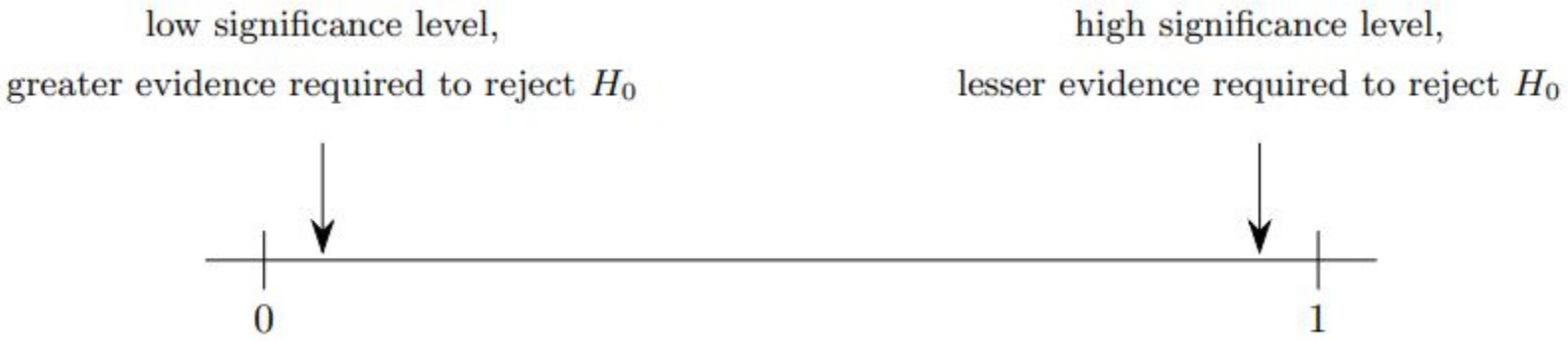
Step 2: Data Collection

Jupyter Notebook

www.github.com/joel-quek

Step 3: Significance Level and P-Value

- We choose the significance level of our test.
- How “convincing” we need our evidence to be before we reject the null hypothesis in favour of the alternative hypothesis.



Step 3: Significance Level and P-Value

- We calculate the p-value
- p-value is the probability of observing a test result that favours the alternative hypothesis while assuming that the null hypothesis is true.

small p-value, **unlikely** to observe a test result that is at least as extreme as what was observed in the sample if H_0 was true.

large p-value, **more likely** to observe a test result that is at least as extreme as what was observed in the sample if H_0 was true.

Step 3: Significance Level and P-Value

Suppose in eight independent tosses of the die, the side with 6 was observed seven times. What is the p-value in this case?

Step 3: Significance Level and P-Value

$$X \sim \text{Binomial}(8, \tfrac{1}{6})$$

$$P(7 \text{ out of } 8 \text{ tosses result in } 6 \text{ facing up} \mid H_0 \text{ is true})$$

$$+$$

$$P(8 \text{ out of } 8 \text{ tosses result in } 6 \text{ facing up} \mid H_0 \text{ is true})$$

Step 4: Decision

Jupyter Notebook

www.github.com/joel-quek

Thank You :)