دورهٔ آموزشی «علم داده»

Data Science Course

جلسهٔ هفتم:

چرا احتمالات؟

مدرس: محمد فزونی

عضو هیات علمی دانشگاه گنبدکاووس

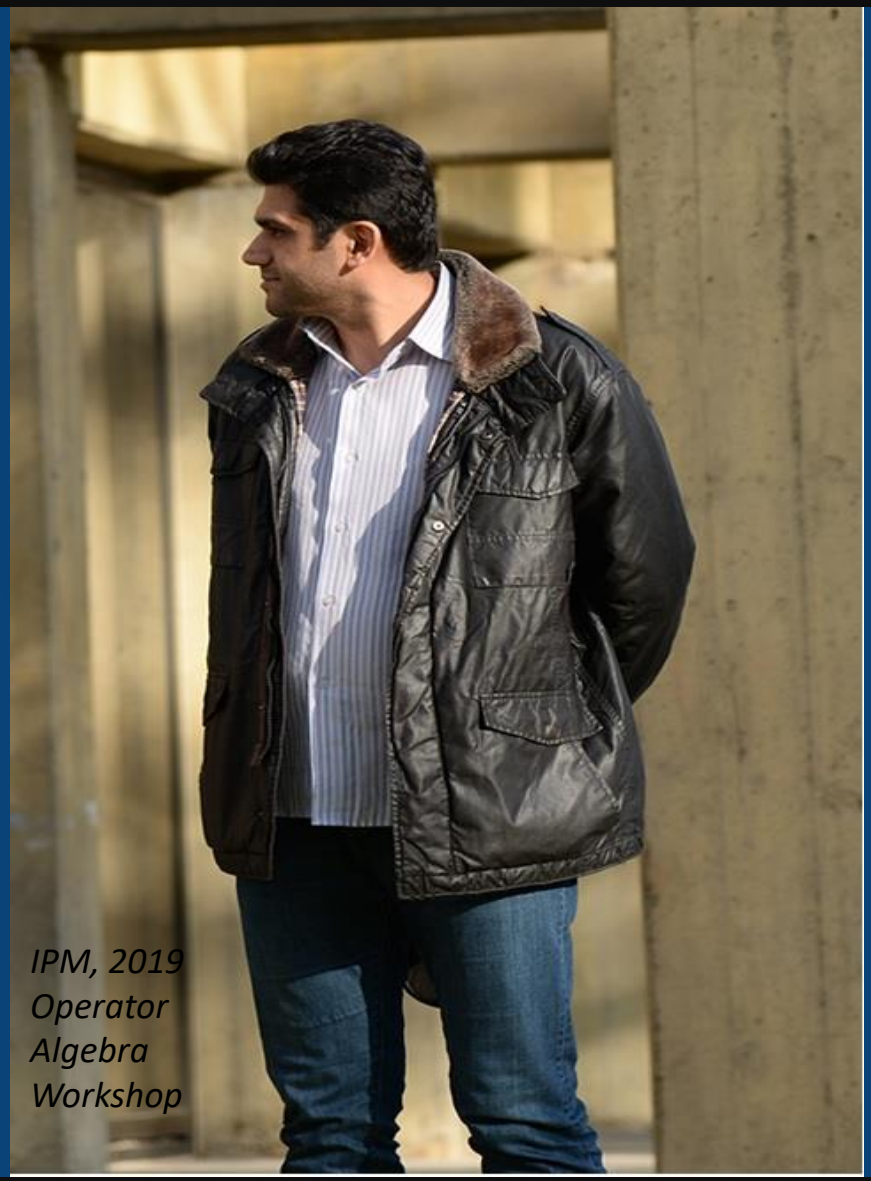
پائیز ۱۳۹۹

دربارهٔ من

Mohammad Fozouni (Ph.D.)
Dep. of Math & Statistics
Gonbad Kavous University

- fozouni@hotmail.com
- https://m-fozouni.ir
- http://profs.gonbad.ac.ir/fozouni/en
- https://www.aparat.com/elmedade

#data_science_fozouni



IPM, 2019
Operator Algebra Workshop

Probability and Finance

Tie-ins to Finance

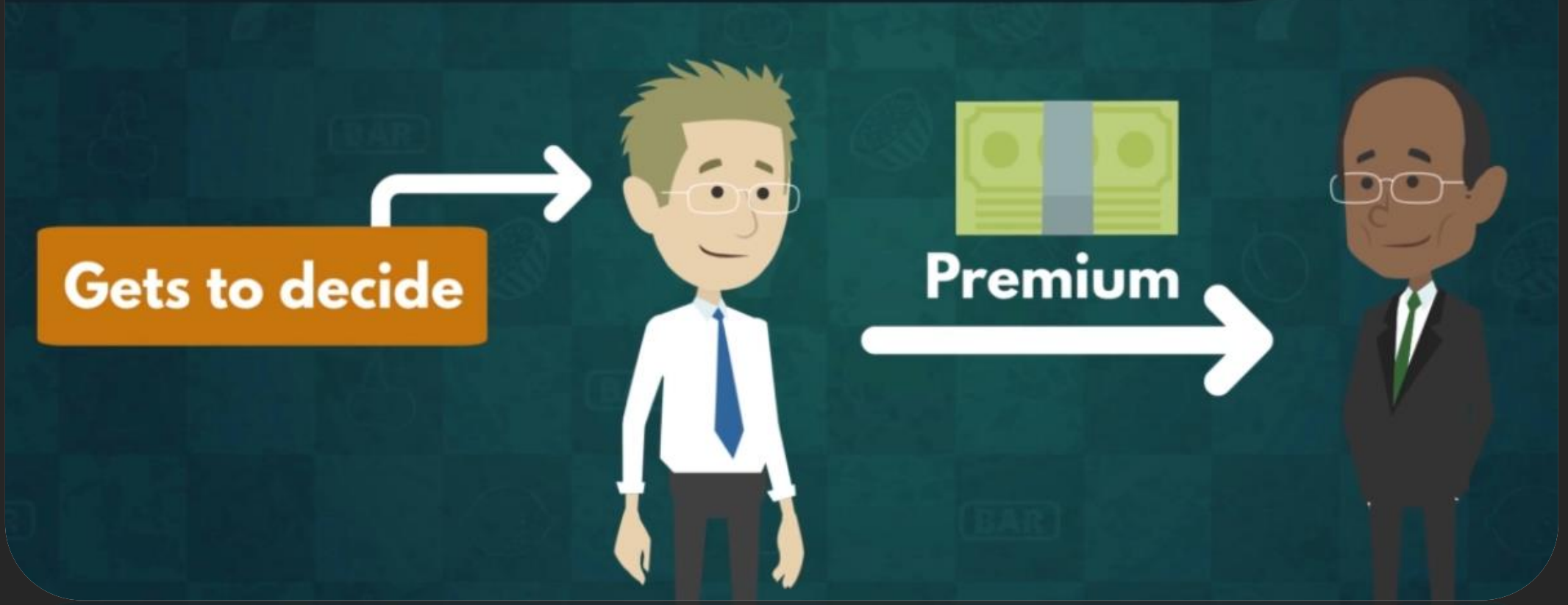
Option:

An agreement between two parties for the price of a stock or item at a future point in time

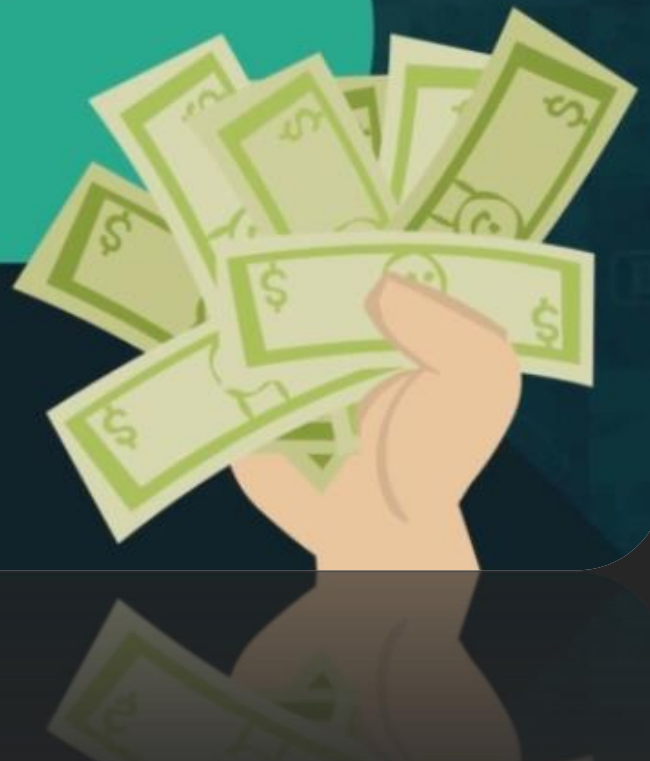
It allows one of the sides to decide

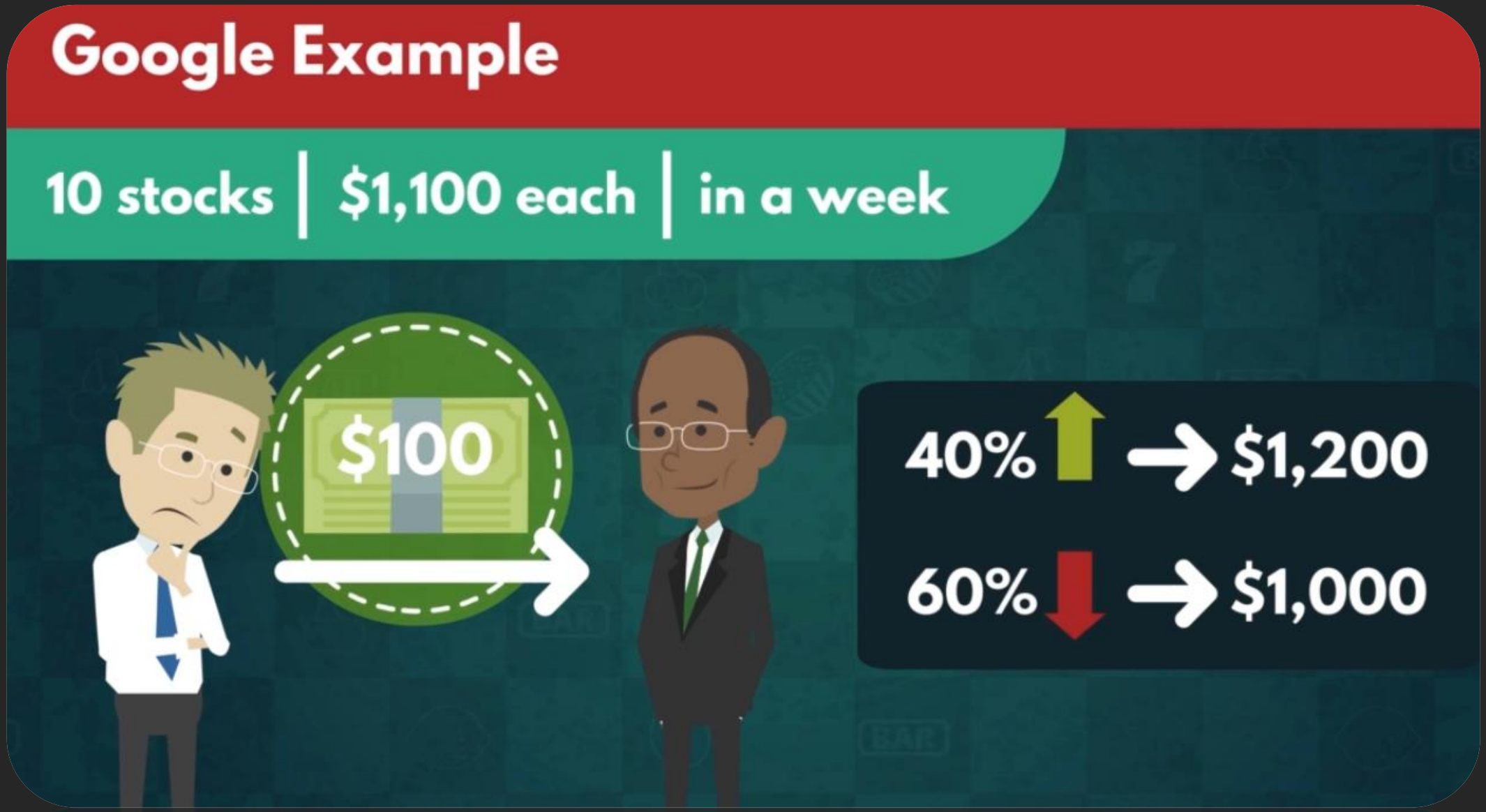
Option pricing

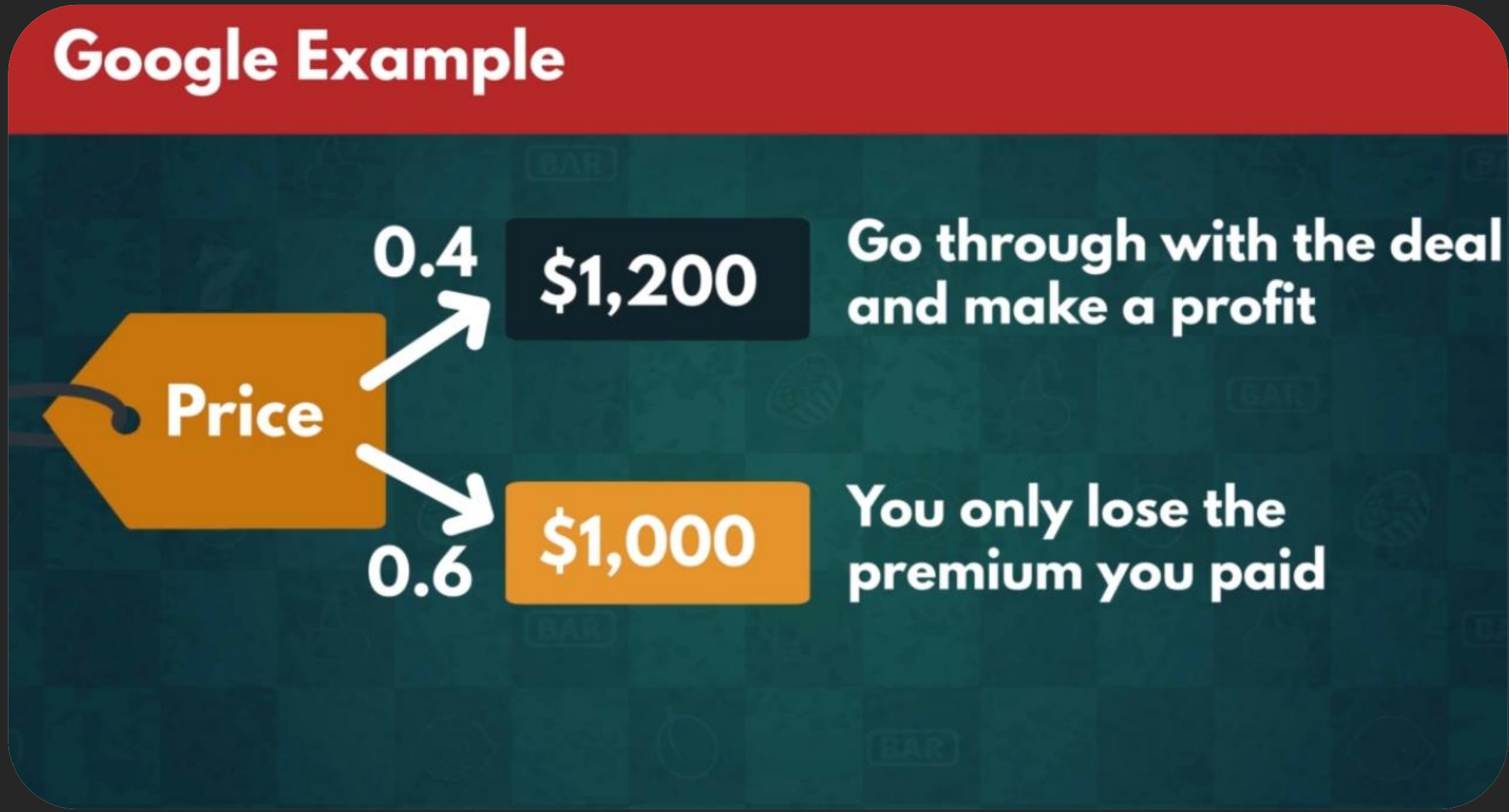
One of the parties at a clear disadvantage

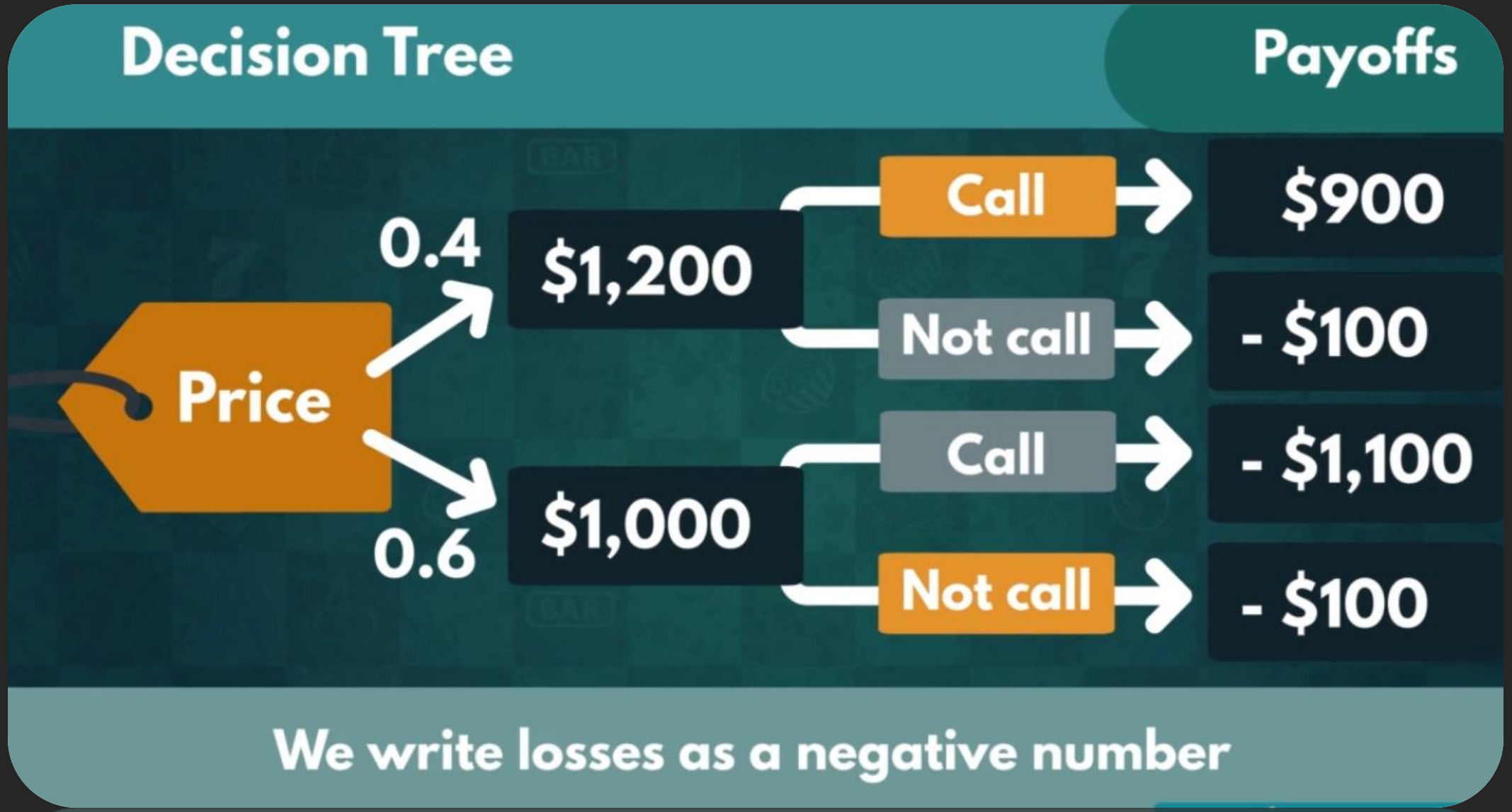


How much we are willing to pay to receive that pact?
(Highest premium)



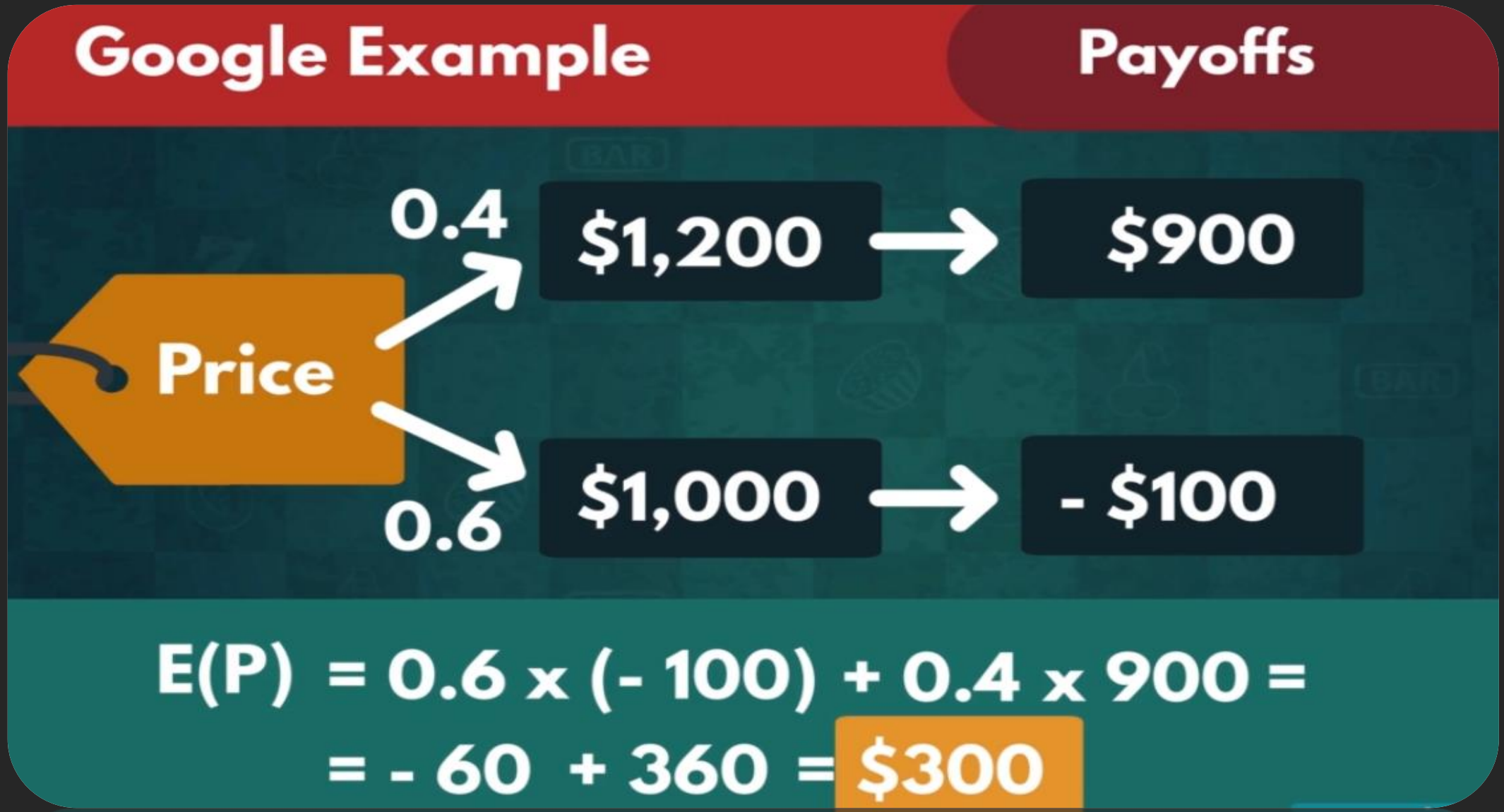

Google Example
10 stocks | $1,100 each | in a week
$100
40% → $1,200
60% → $1,000

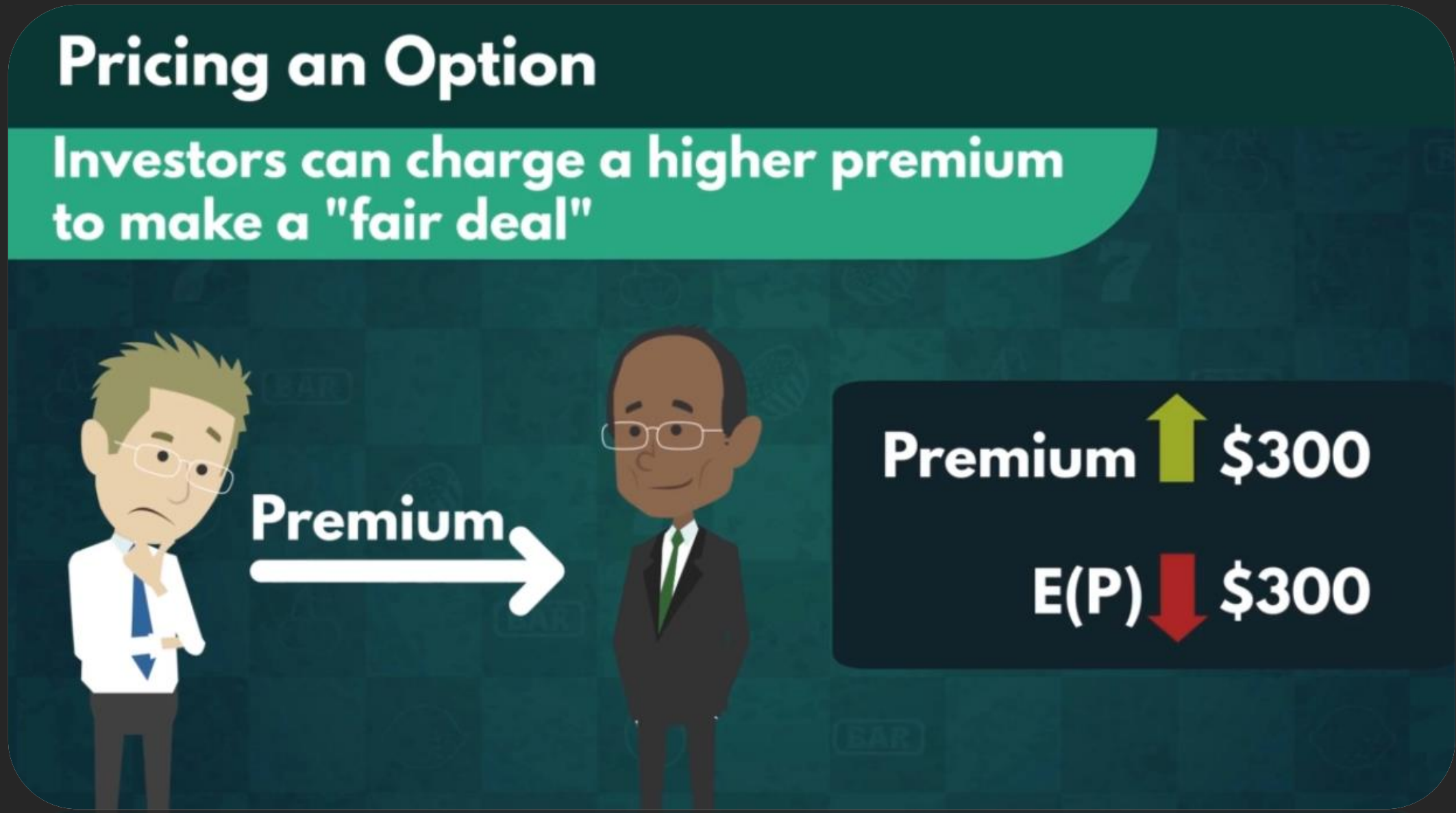

Google Example
0.4
$1,200
Go through with the deal and make a profit
Price
0.6
$1,000
You only lose the premium you paid


Decision Tree
Payoffs
Price
0.4
$1,200
Call
$900
Not call
- $100
0.6
$1,000
Call
- $1,100
Not call
- $100
We write losses as a negative number

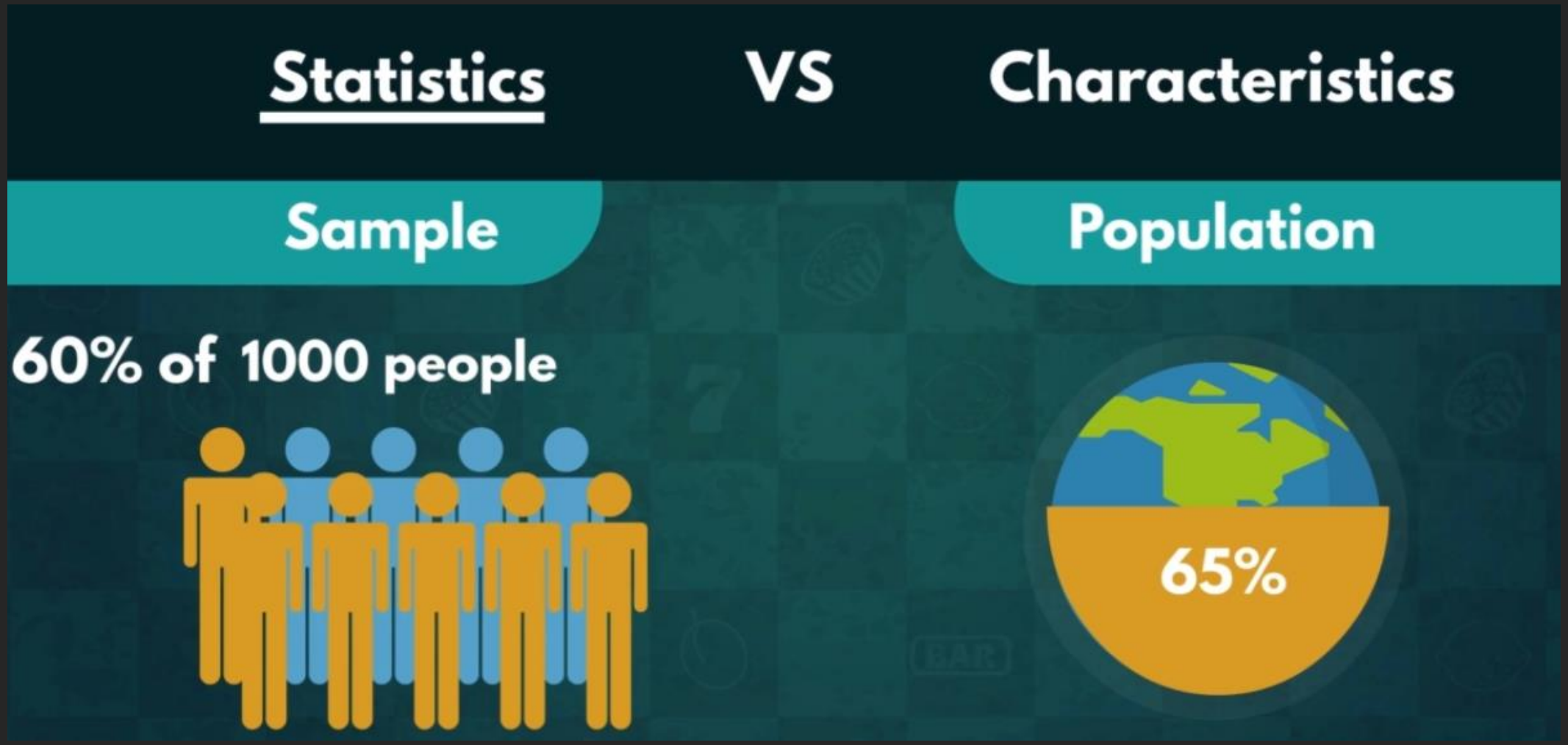
Expected Payoffs

$E(P) < 0$	Disadvantageous (Avoid buying this option)
$E(P) = 0$	"Fair deal" (You expect to make as much as you paid)
$E(P) > 0$	Favourable (Go through with the deal)


Google Example
Payoffs
0.4
$1,200
$900
Price
0.6
$1,000
- $100
E(P) = 0.6 x (- 100) + 0.4 x 900 =
= - 60 + 360 = $300


Pricing an Option
Investors can charge a higher premium to make a "fair deal"
Premium
Premium $300
E(P) $300

Probability and Statistics


Statistics
VS
Characteristics
Sample
Population
60% of 1000 people
65%

Statistics

Statistics focuses predominantly on samples and incomplete data

Uncertainty

Expected value

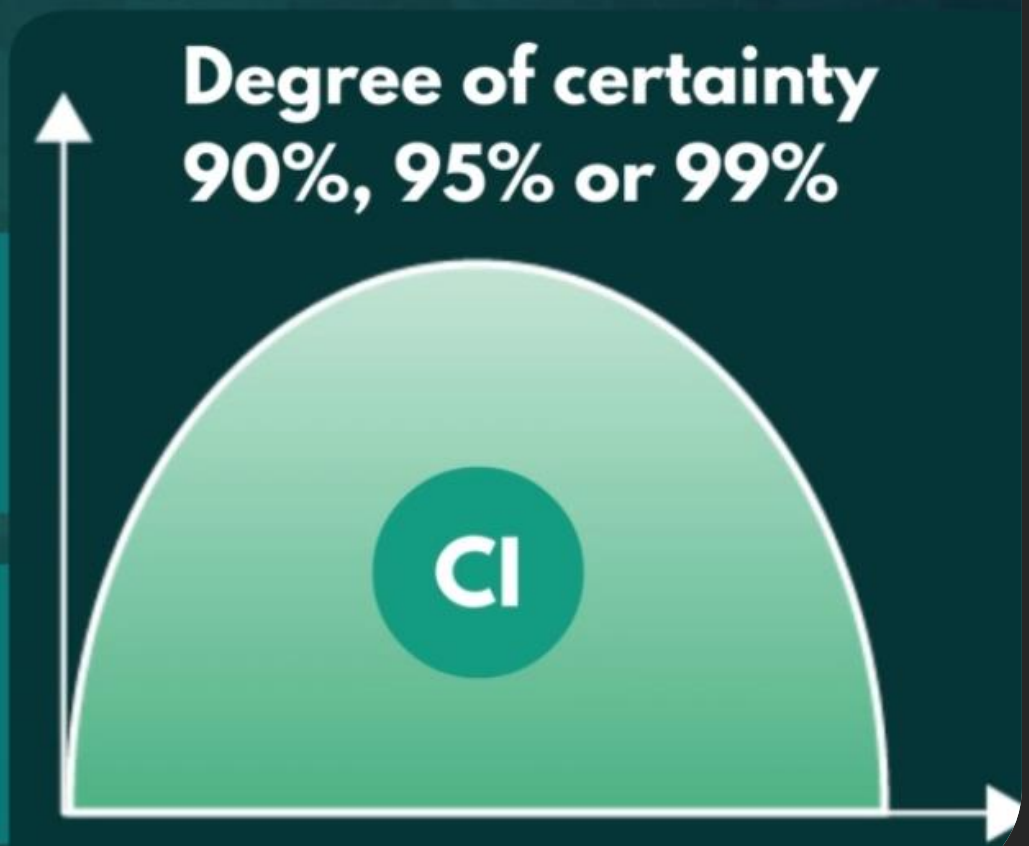
Prediction intervals


Data Science
Statistics
Probability
Concrete
General

Statistics

Experimental data

- Many useful concepts based on probability theory
- Express the likelihood of the population mean being within that interval



Confidence Intervals

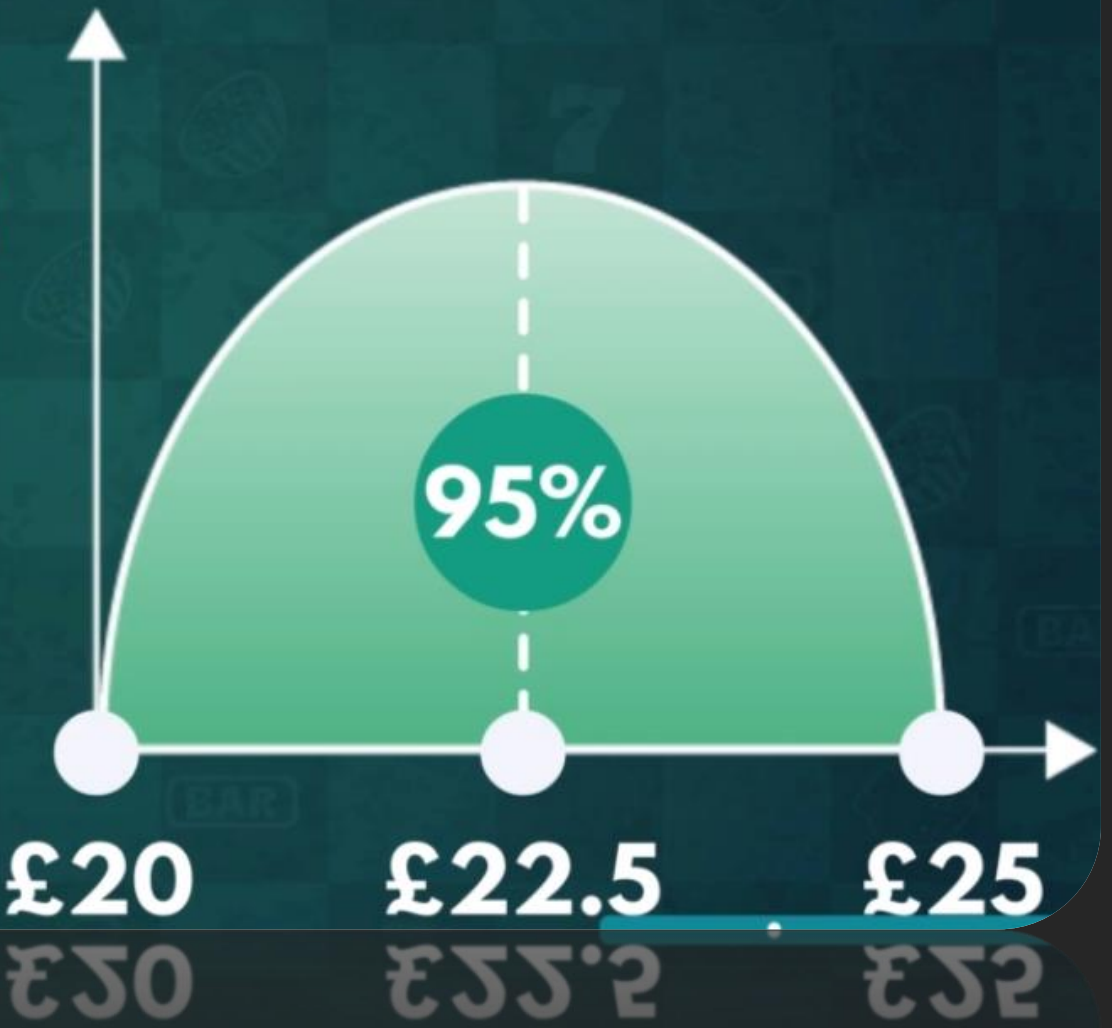
To calculate these CIs we must know :

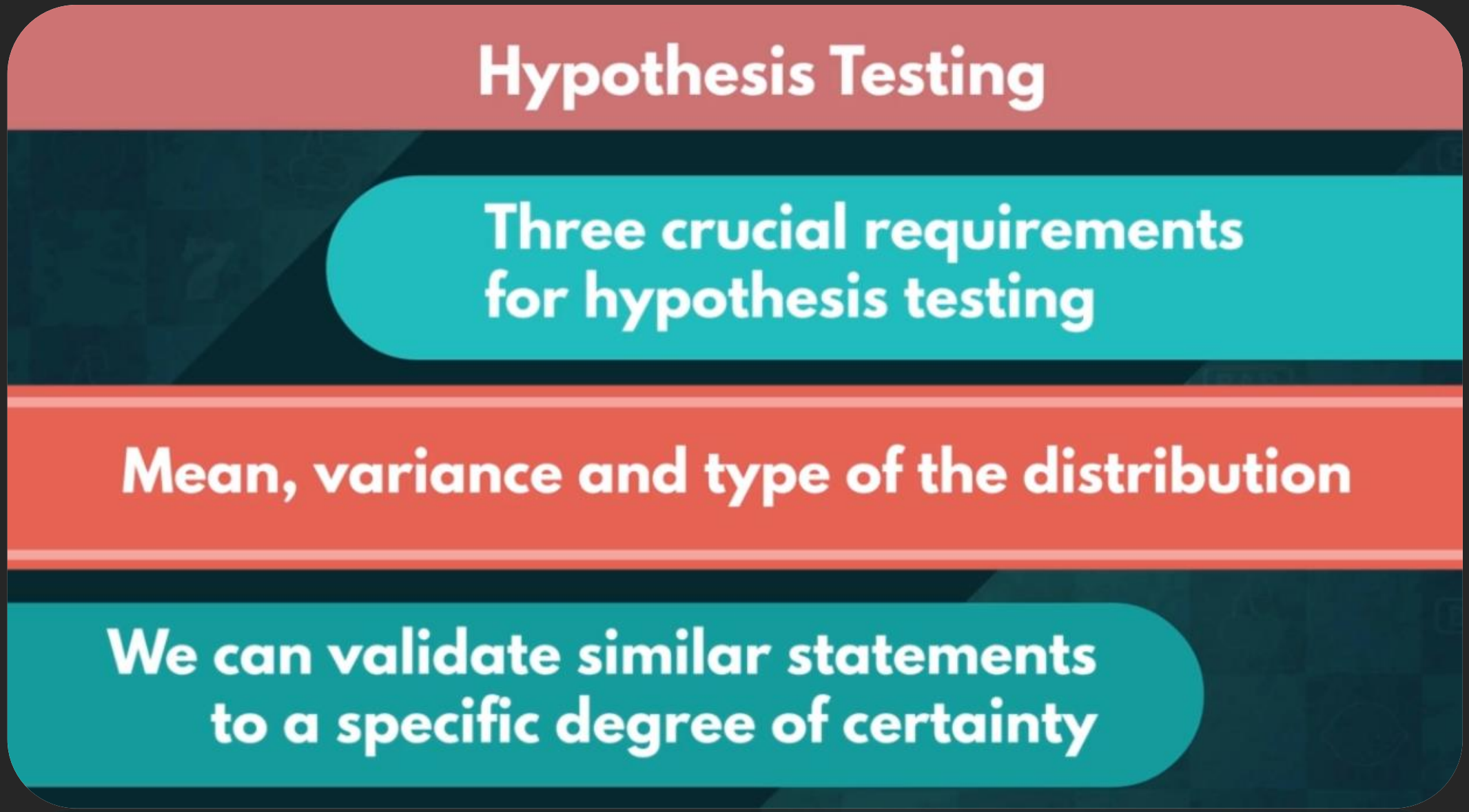
Mean, variance and standard deviation

A good understanding of probability is crucial

London Example

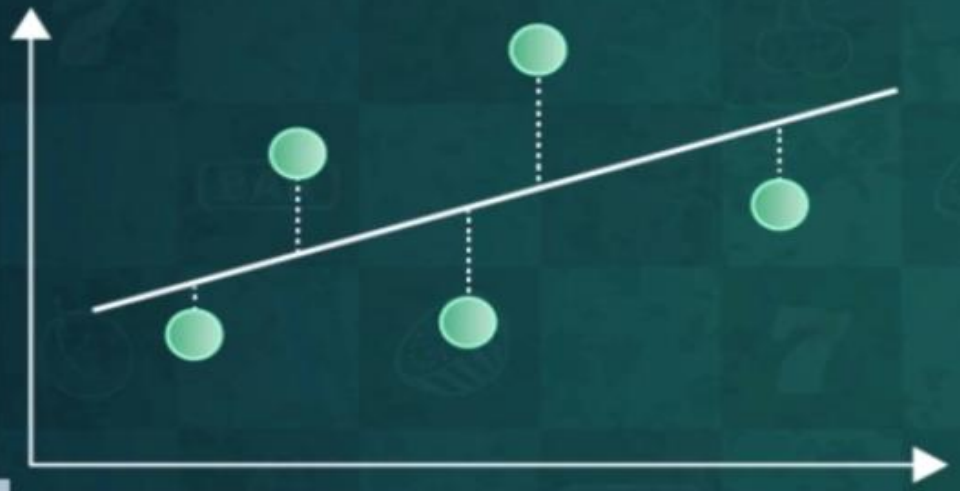
CI = [20; 25]




Hypothesis Testing
Three crucial requirements for hypothesis testing
Mean, variance and type of the distribution
We can validate similar statements to a specific degree of certainty

Knowing the Type of a Distribution

Create different models



Computationally expensive

Computer software

MATHEMATICAL MODELING

Mathematical Modeling

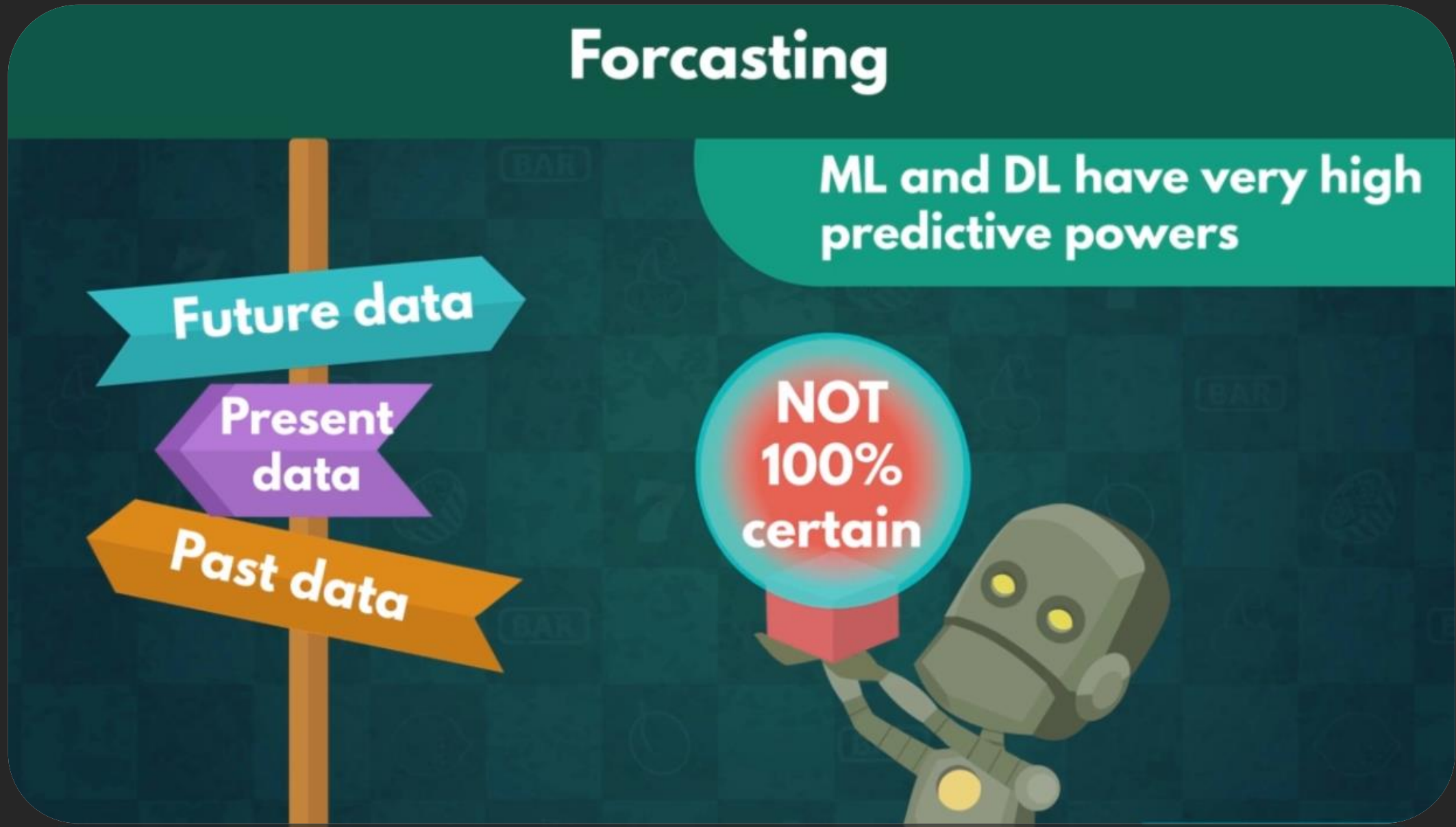
An extension of statistics that data scientists deal with

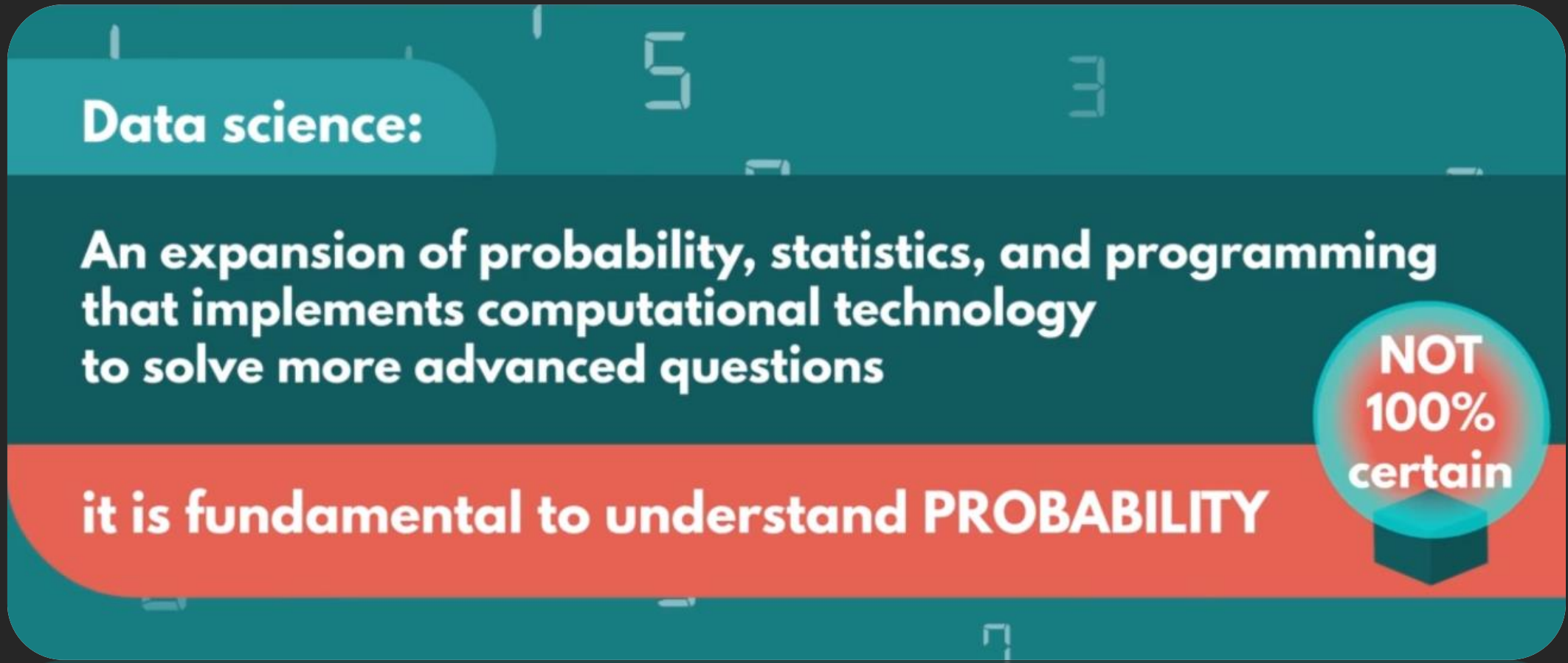
Probability and Data Science

Expected Values

ML → An extremely fast paced trial-and-error process

The more predictions it makes, the more precise they become


Forcasting
ML and DL have very high predictive powers
Future data
Present data
Past data
NOT 100% certain


Data science:
An expansion of probability, statistics, and programming
that implements computational technology
to solve more advanced questions
it is fundamental to understand PROBABILITY
NOT
100%
certain