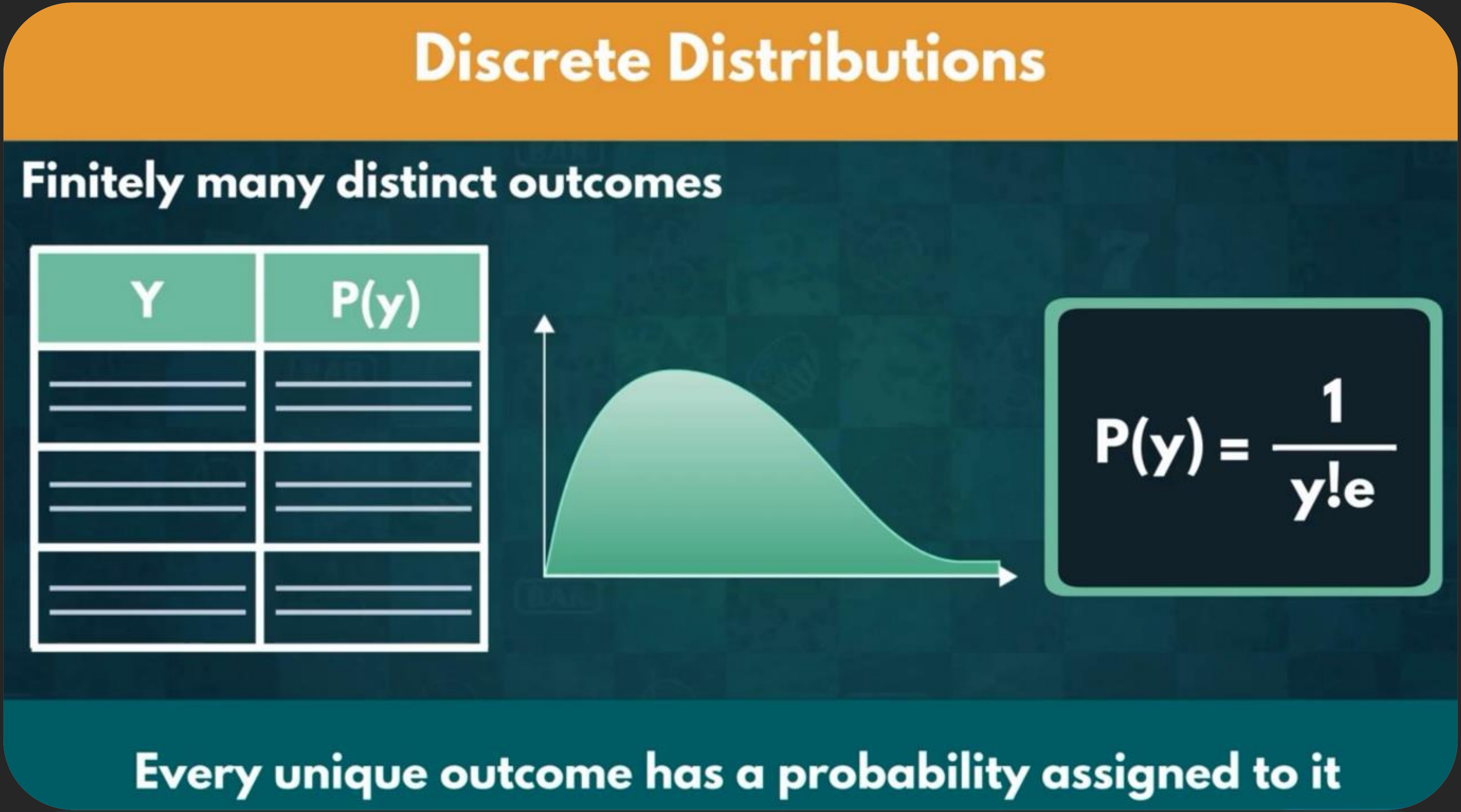
دورهٔ آموزشی «علم داده»

Data Science Course

جلسهٔ چهارم:

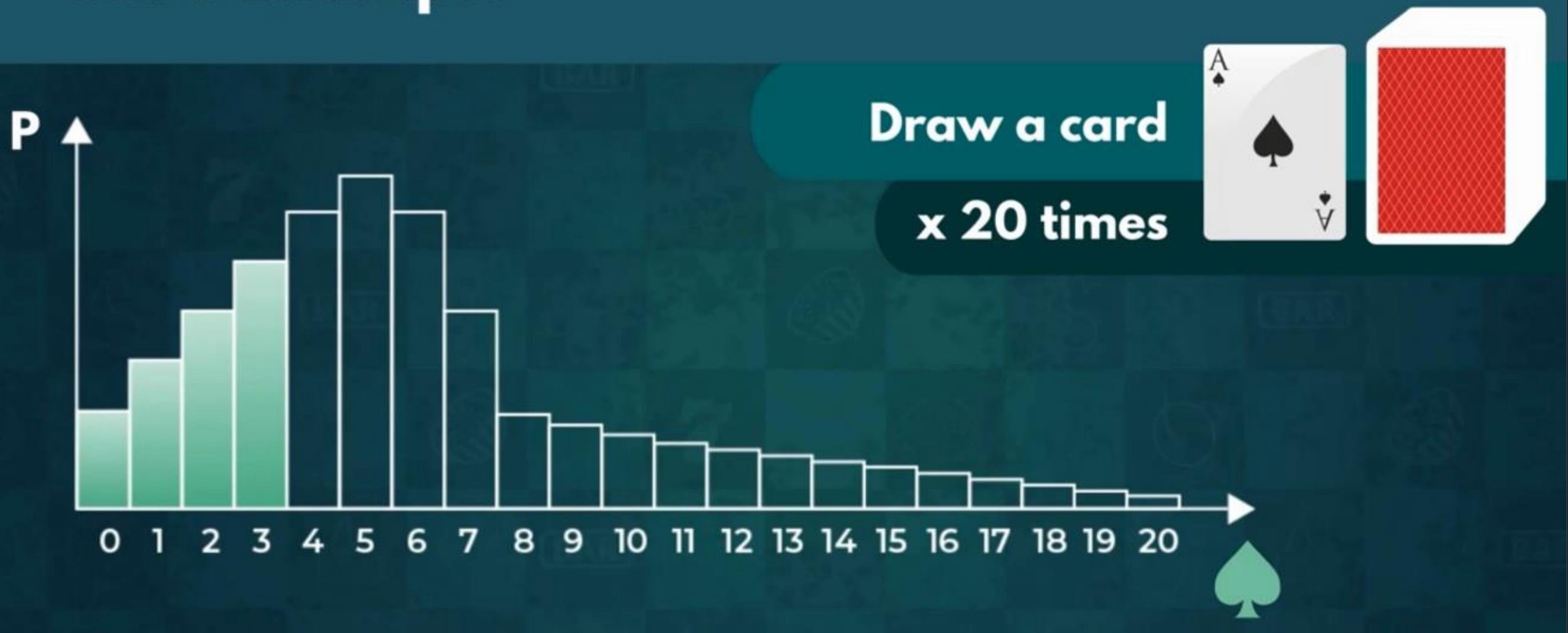
توزیع‌های گسسته

مدرس: محمد فزونی
عضو هیات علمی دانشگاه گنبدکاووس
پائیز ۱۳۹۹



www.m-fozouni.ir
Data Science Course

Card Example



$$P(0) + P(1) + P(2) + P(3) = P(y \leq 3)$$

Die Example

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6)$$

Main Takeaway

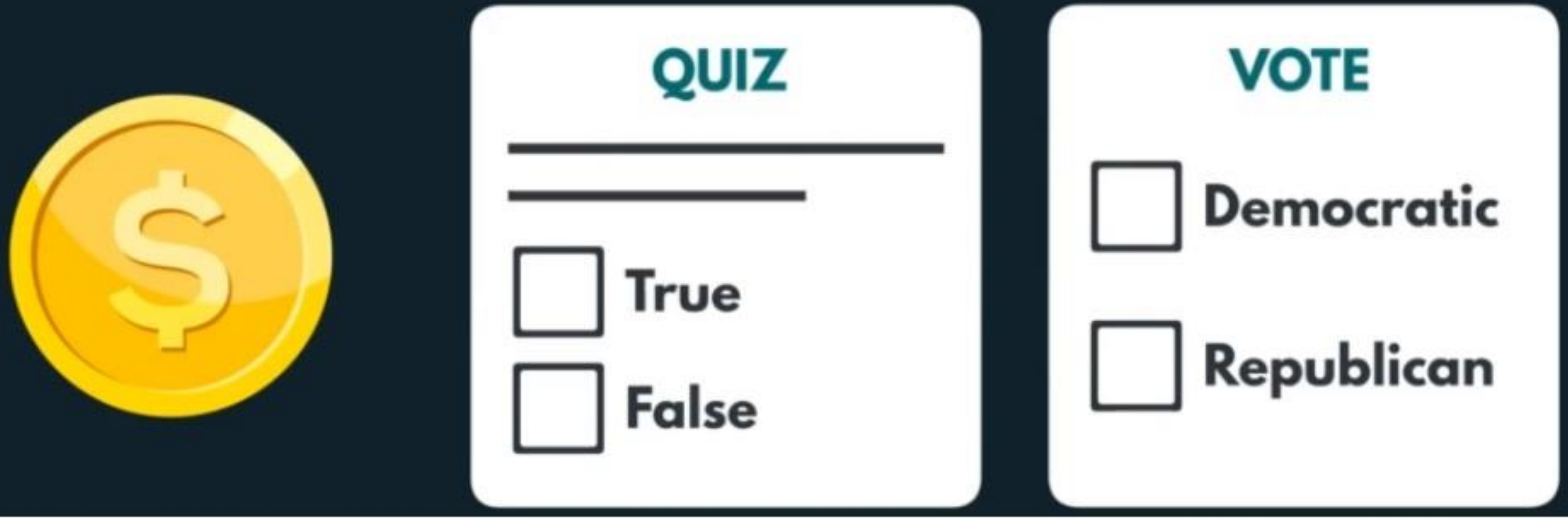
$$X \sim U(a, b)$$

- Each outcome is equally likely
- Both the mean and the variance are uninterpretable
- No predictive power

Bernoulli Distribution

Events with:

- 1 Trial
- 2 Possible outcomes



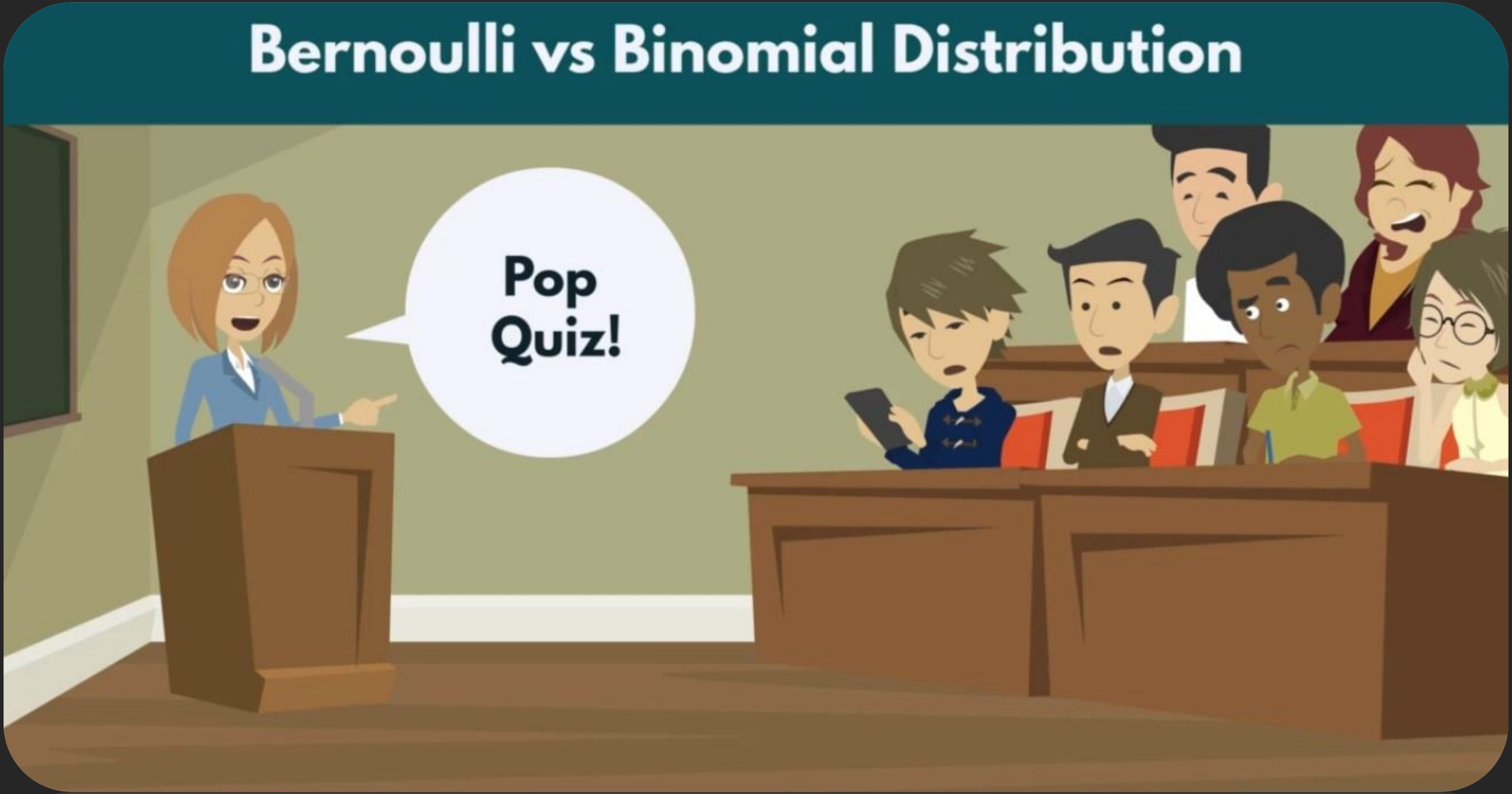
Binomial Distribution

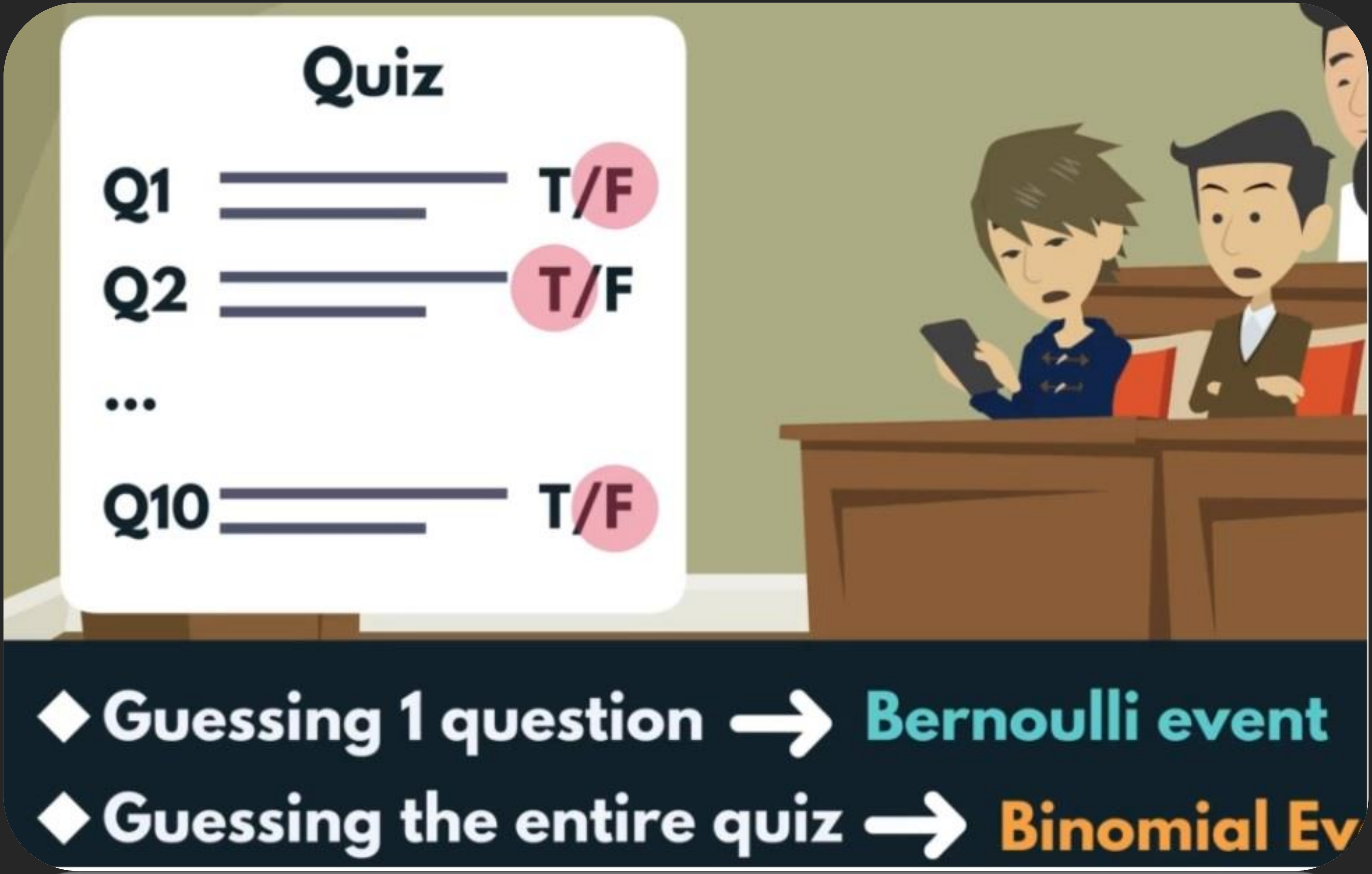
Notation

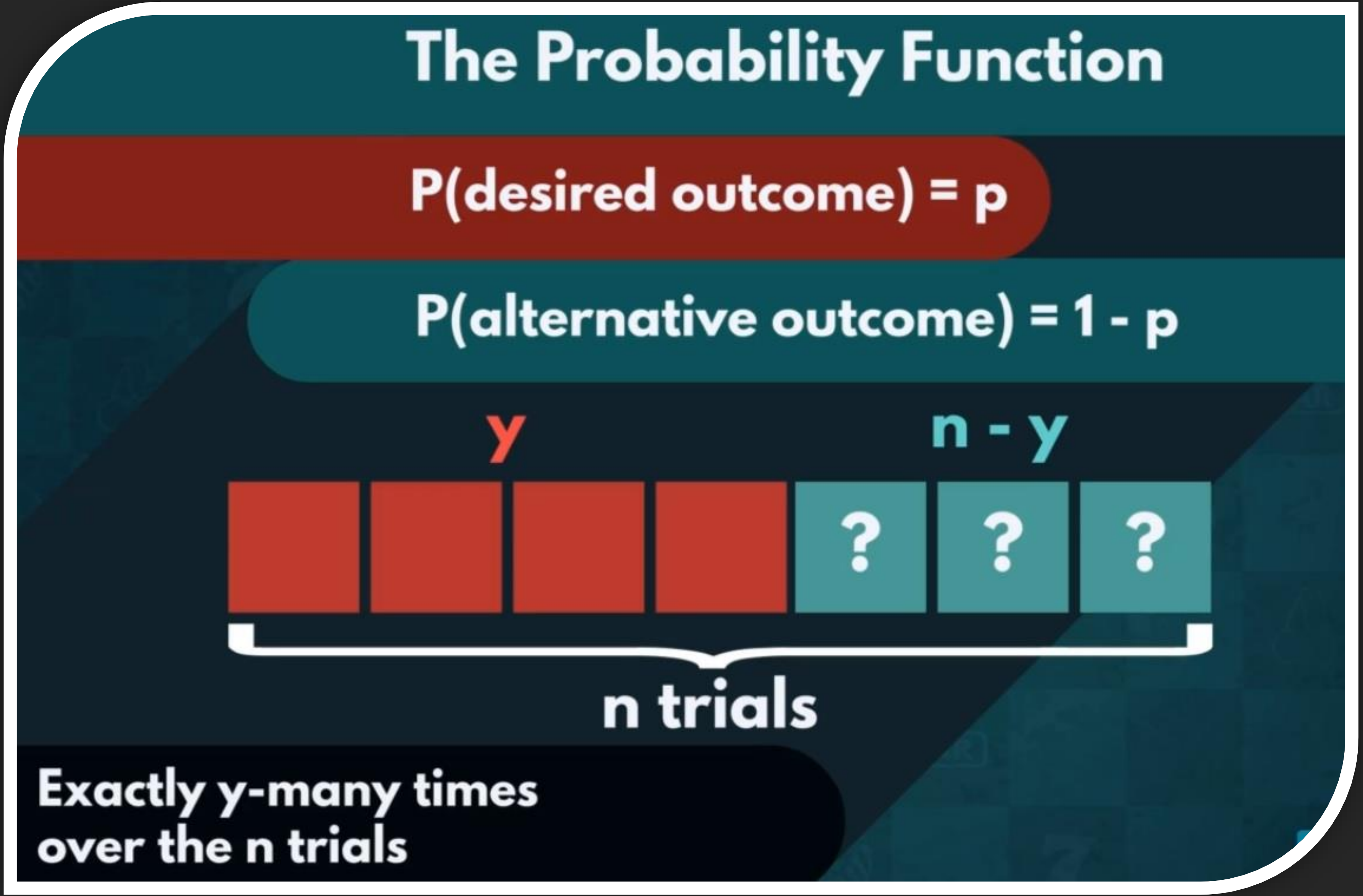
$$B\,(n,\ p)$$

$$X \sim B\,(10,\ 0.6)$$

$$\Rightarrow \mathrm{Bern}\,(p) = B\,(1,\ p)$$

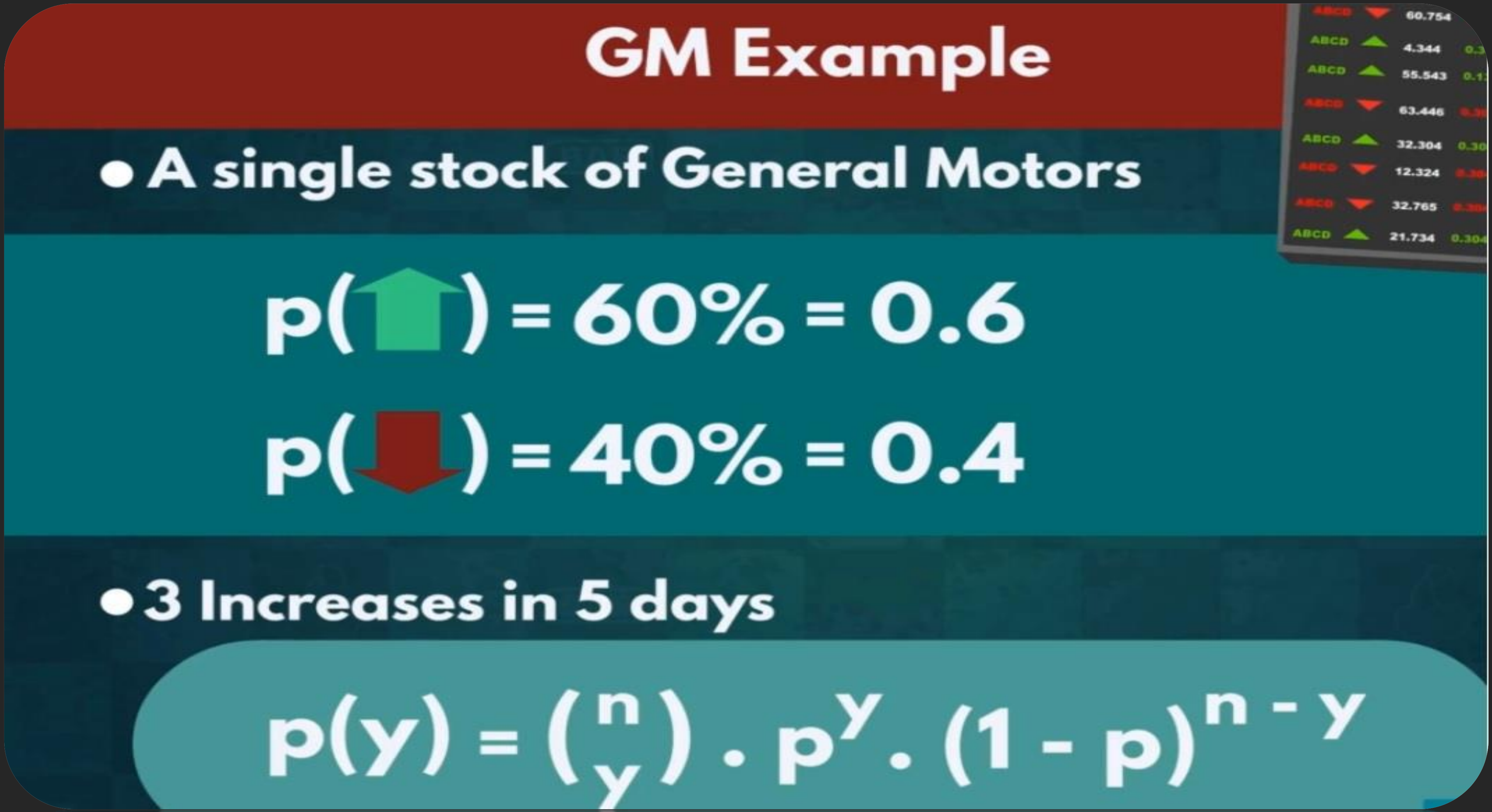

Bernoulli vs Binomial Distribution
Pop Quiz!


Quiz
Q1
T/F
Q2
T/F
...
Q10
T/F
◆ Guessing 1 question → Bernoulli event
◆ Guessing the entire quiz → Binomial Ev


The Probability Function
P(desired outcome) = p
P(alternative outcome) = 1 - p
y
n - y
?
?
?
n trials
Exactly y-many times
over the n trials

The Probability Function

$$p(y) = \binom{n}{y} \cdot p^{y} \cdot (1 - p)^{n - y}$$


GM Example
A single stock of General Motors
p( ) = 60% = 0.6
p( ) = 40% = 0.4
3 Increases in 5 days
$p(y) = \binom{n}{y} \cdot p^y \cdot (1-p)^{n-y}$

GM Example

$y \Rightarrow 3$
$n \Rightarrow 5$
$p \Rightarrow 0.6$

$$p(y) = \binom{n}{y} \cdot p^{y} \cdot (1-p)^{n-y} =$$

$$= C_3^5 \times 0.6^3 \times 0.4^2 =$$

= 34.56%

of getting exactly 3 increases

Expected Values

$$E(X) = x_0 \cdot p(x_0) + x_1 \cdot p(x_1) + \ldots x_n \cdot p(x_n)$$

$$Y \sim B(n, p)$$

$$E(Y) = p \cdot n$$

Variance and Standard Deviation

$$\sigma^2 = E(y^2) - E(y)^2 =$$

$$= n \cdot p \cdot (1 - p)$$

Formula

Poisson Distribution is wildly different

$$P(Y) = \frac{\lambda^{y} e^{-\lambda}}{y!}$$

Its probability function is much different

Expected Value

$p(y) \Rightarrow E(y)$

$$E(y) = y_0 \frac{\lambda^{y_0} e^{-\lambda}}{y_0!} + y_1 \frac{\lambda^{y_1} e^{-\lambda}}{y_1!} + \dots = \lambda$$

Variance

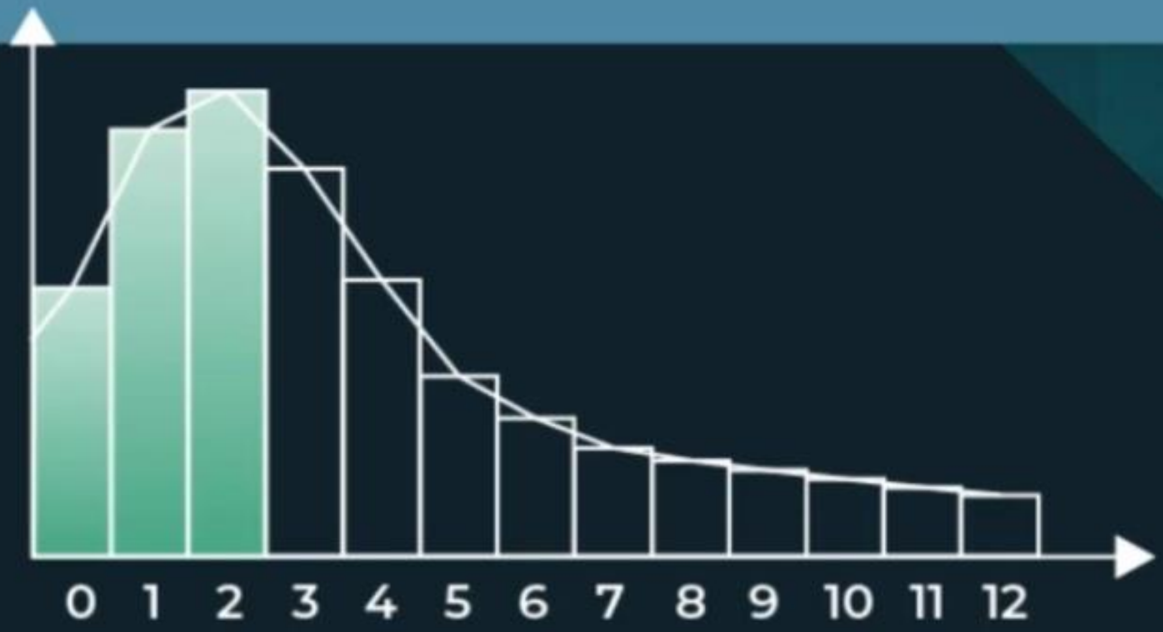
$$\sigma^2 = (y_0 - \mu)^2 + (y_1 - \mu)^2 + \ldots =$$

$$= \lambda$$

So we have:

$$\mu = \sigma^2 = \lambda$$

Interval



- Same steps as discrete distributions
- The joint probability of all individual elements

SUBJECT OF THE NEXT VIDEO

Continuous Distributions

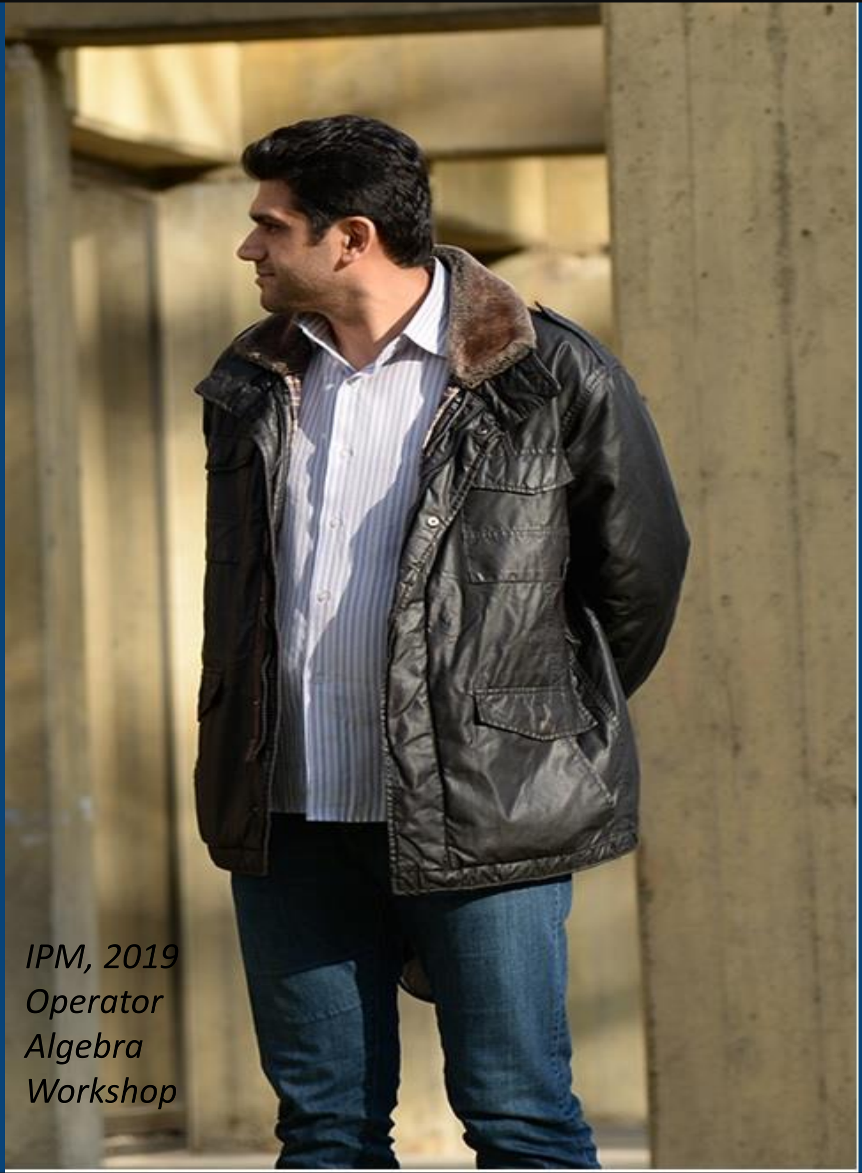
In details

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