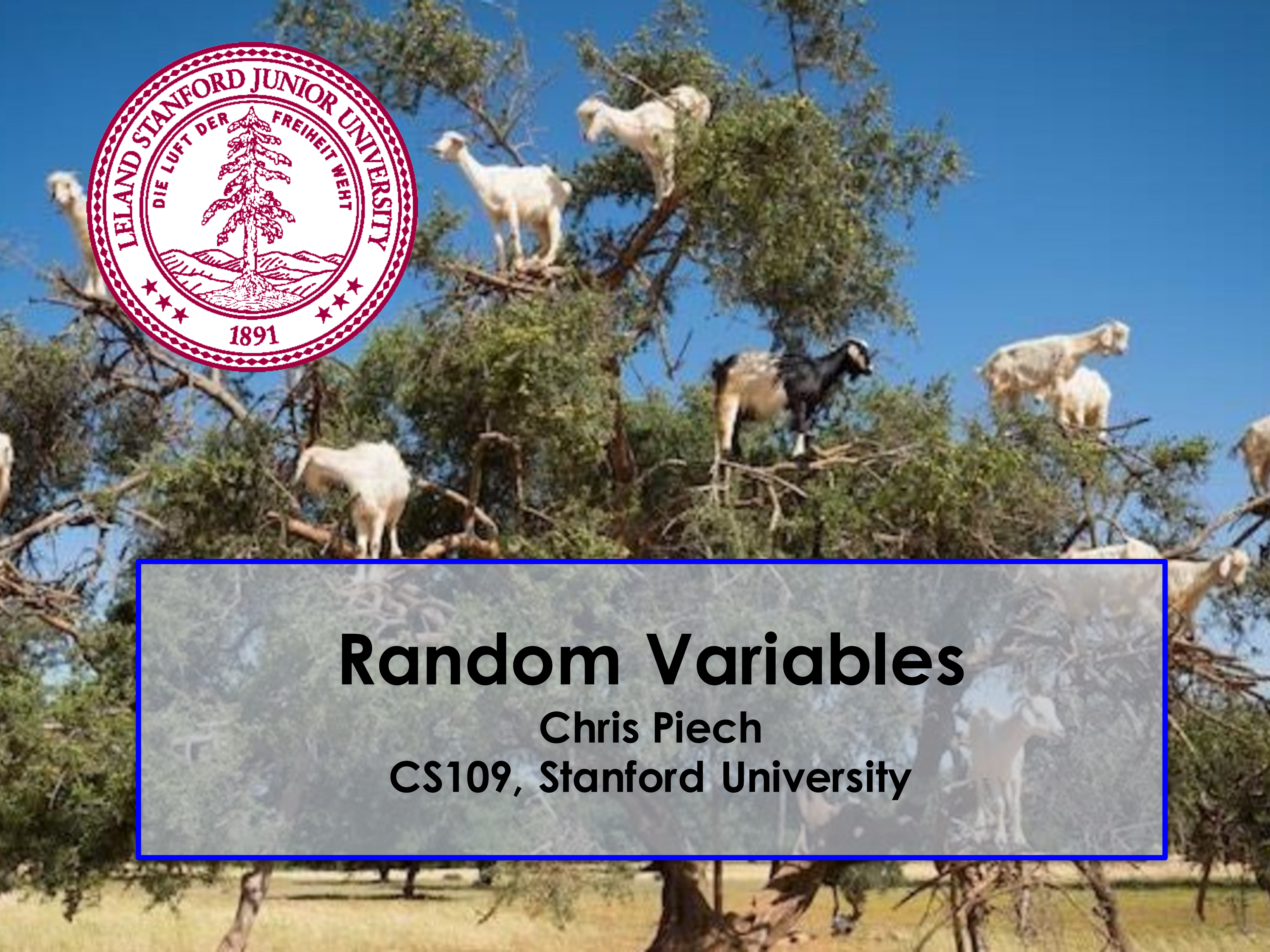
Fruit of the random tree.



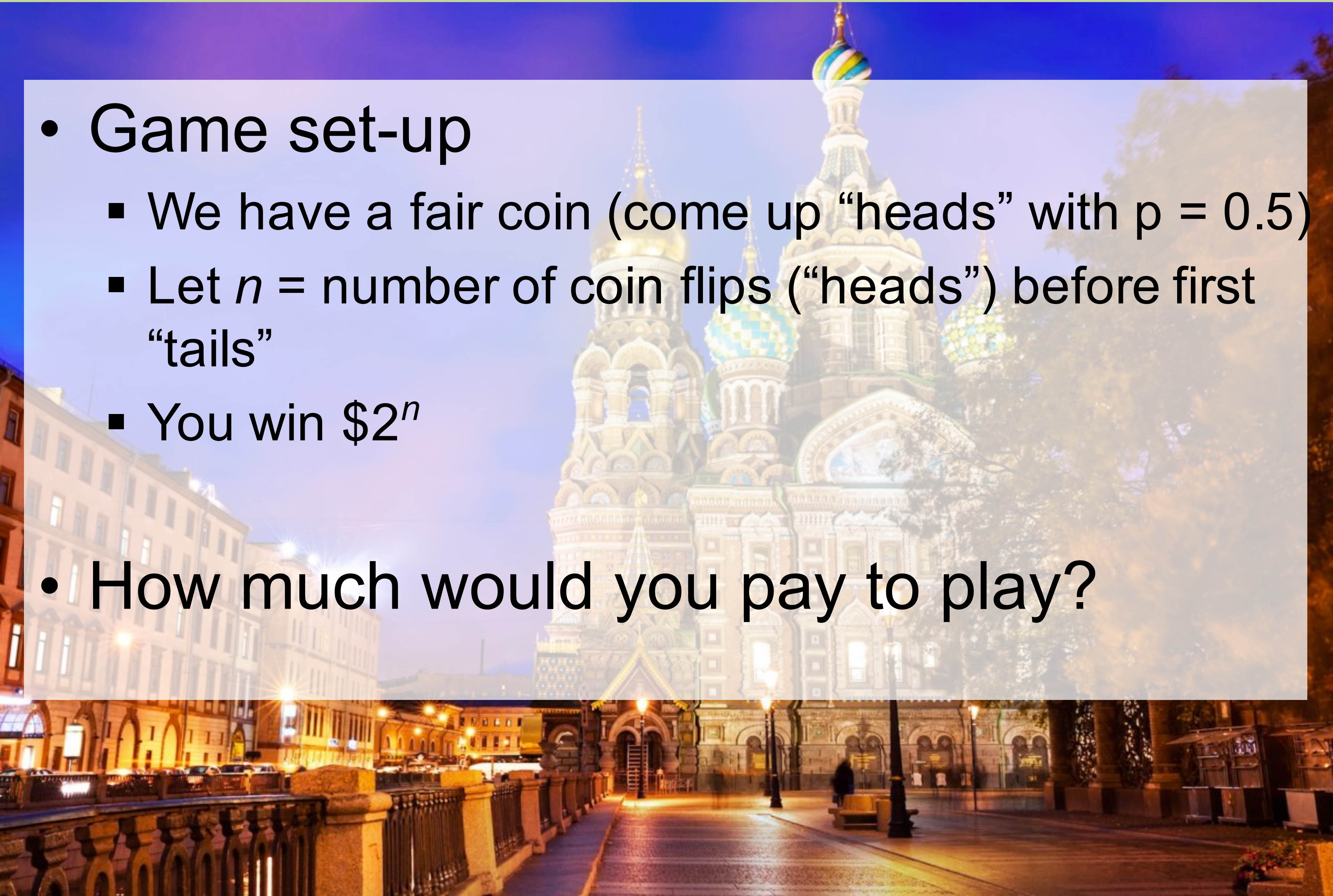
Random Variables

Chris Piech
CS109, Stanford University

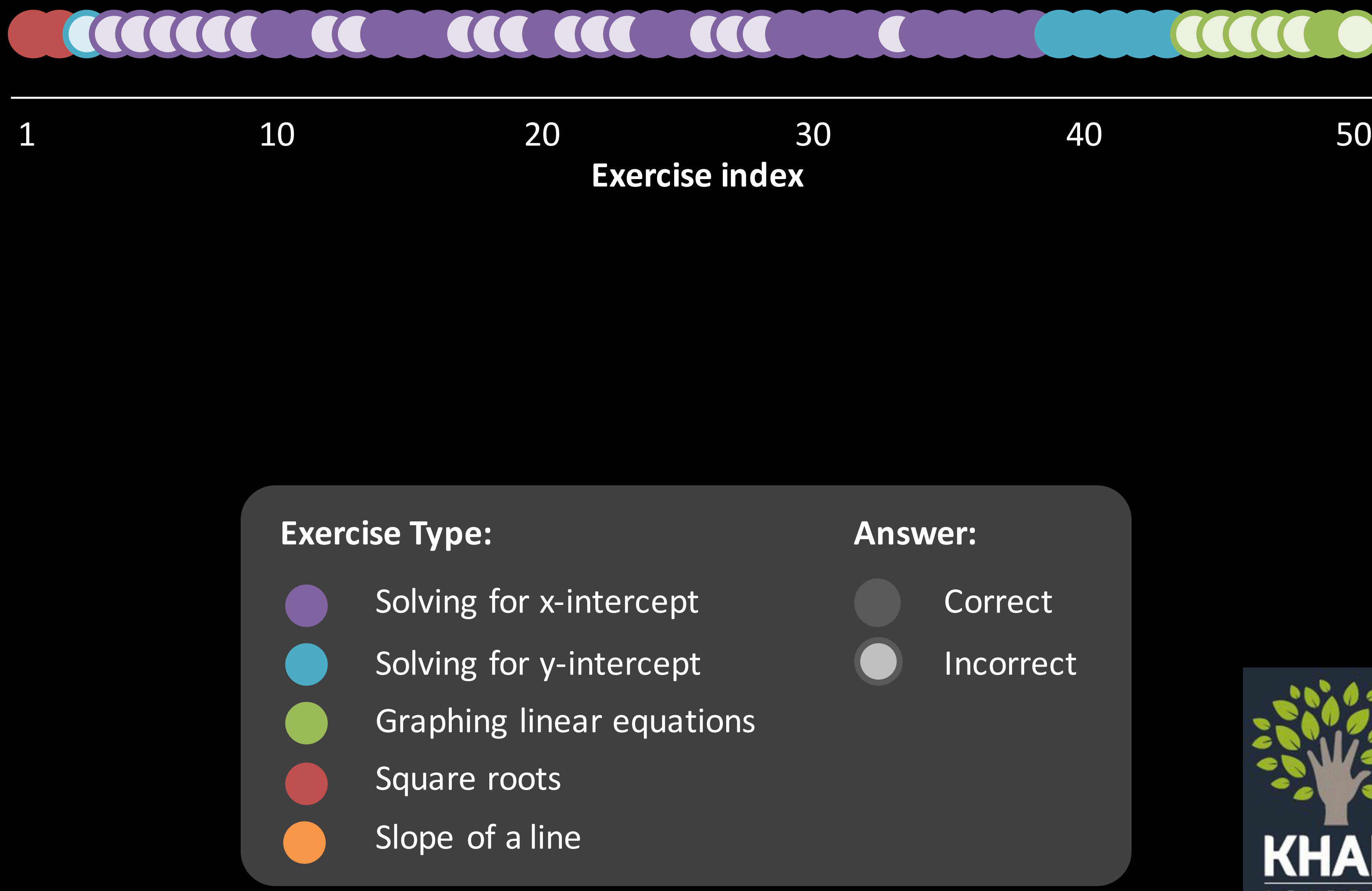
St Petersburg

- Game set-up
  - We have a fair coin (come up “heads” with p = 0.5)
  - Let n = number of coin flips (“heads”) before first “tails”
  - You win $\$2^n$

- How much would you pay to play?



Story of Riley



KHAN ACADEMY

Story of Riley

Exercise Type:

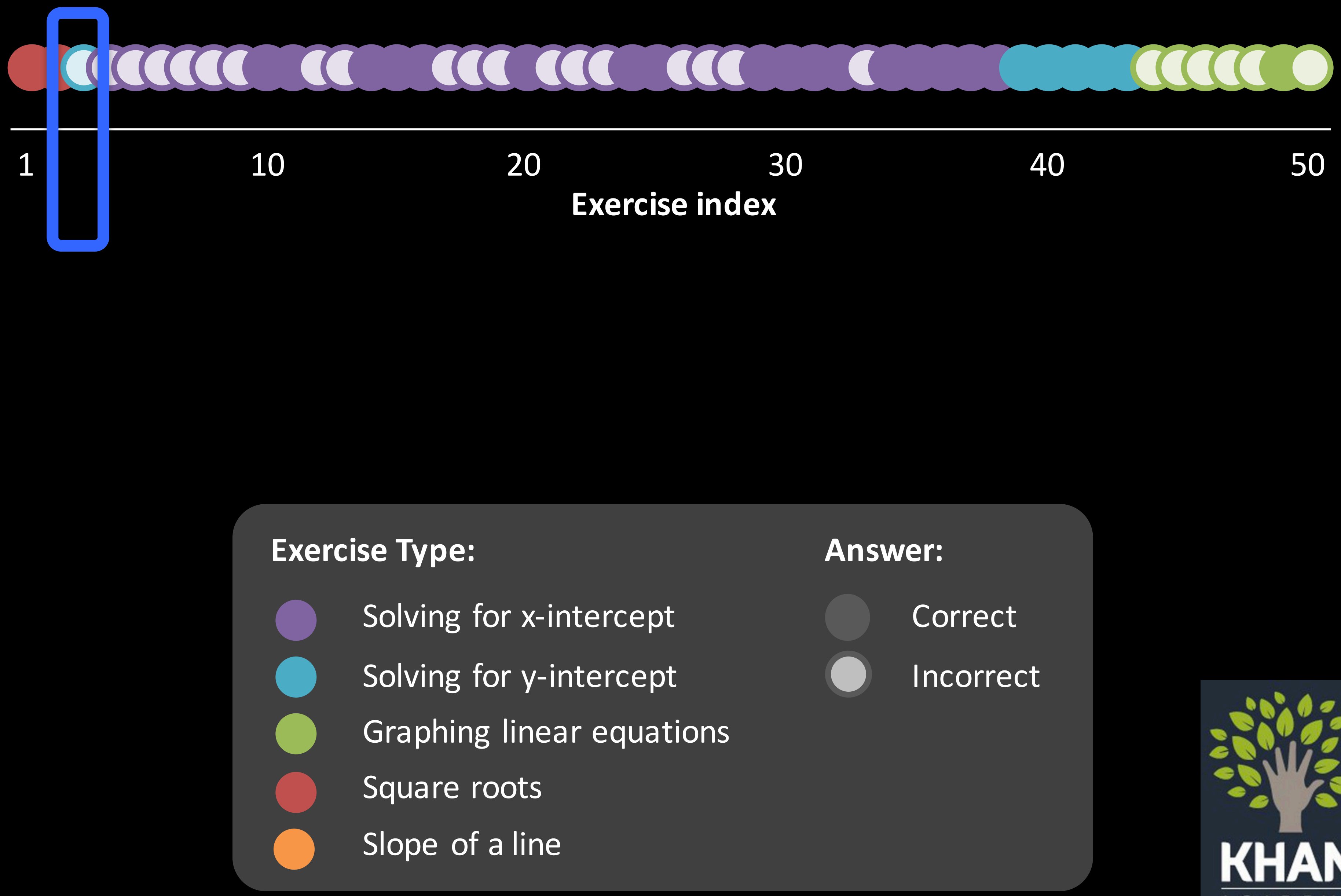
- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

Answer:

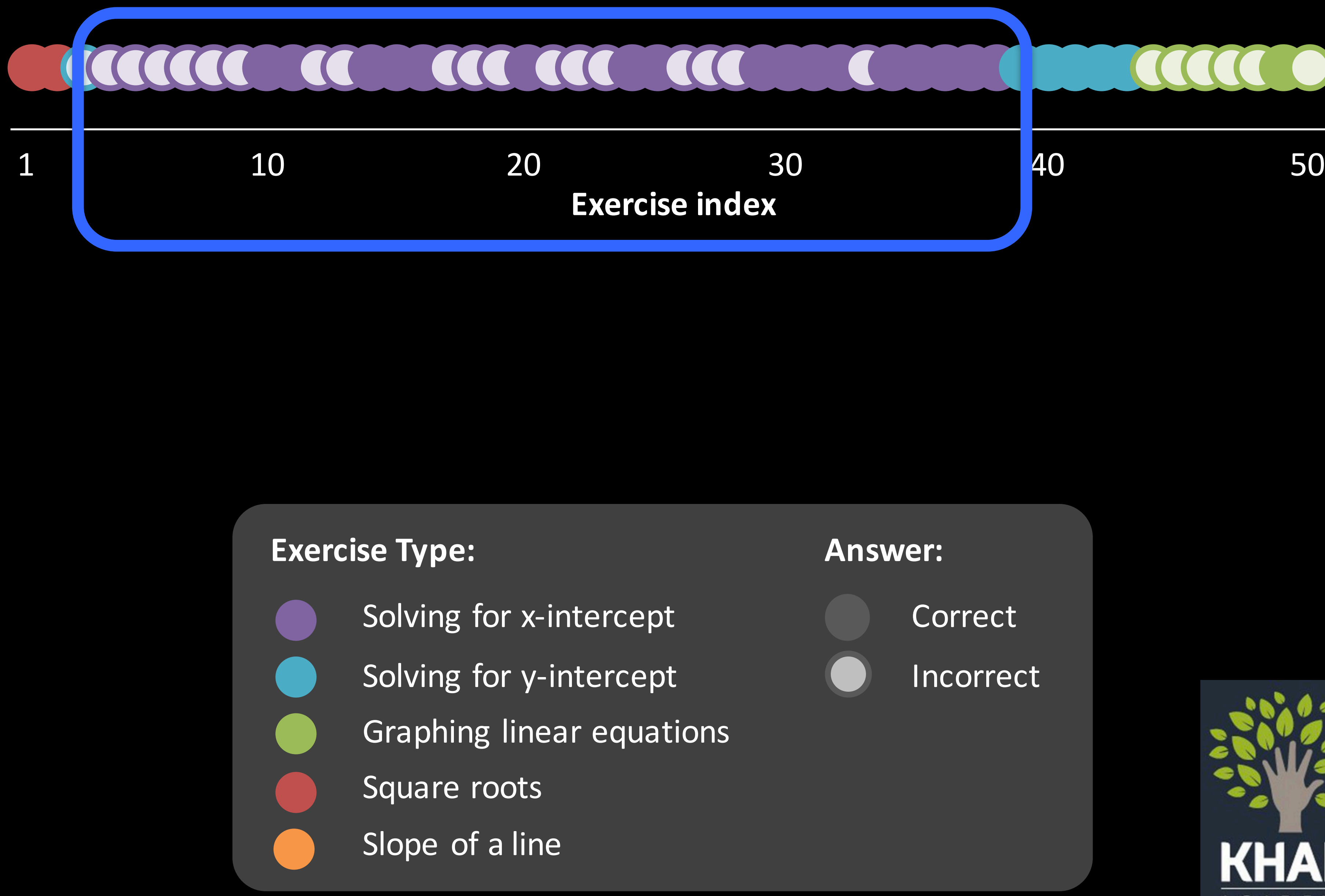
- Correct
- Incorrect

KHAN ACADEMY

Story of Riley



Story of Riley



KHAN ACADEMY

Story of Riley

Exercise Type:

- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

Answer:

- Correct
- Incorrect

KHAN ACADEMY

Story of Riley

Exercise Type:

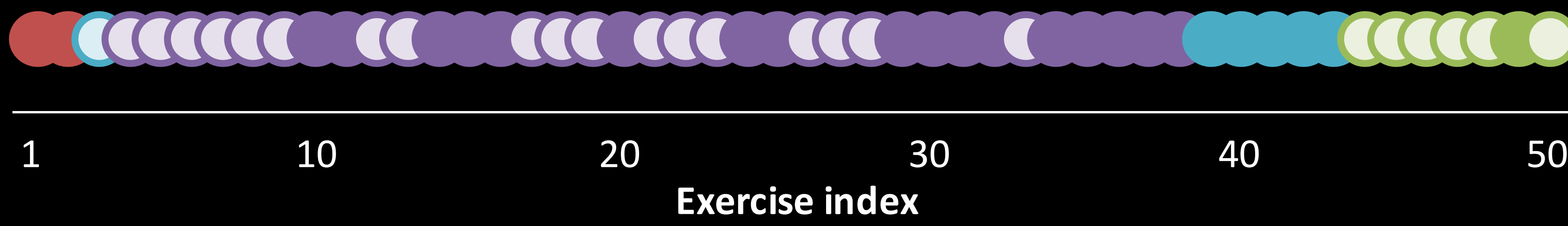
- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

Answer:

- Correct
- Incorrect

KHAN ACADEMY

Story of Riley



What should Riley do next?

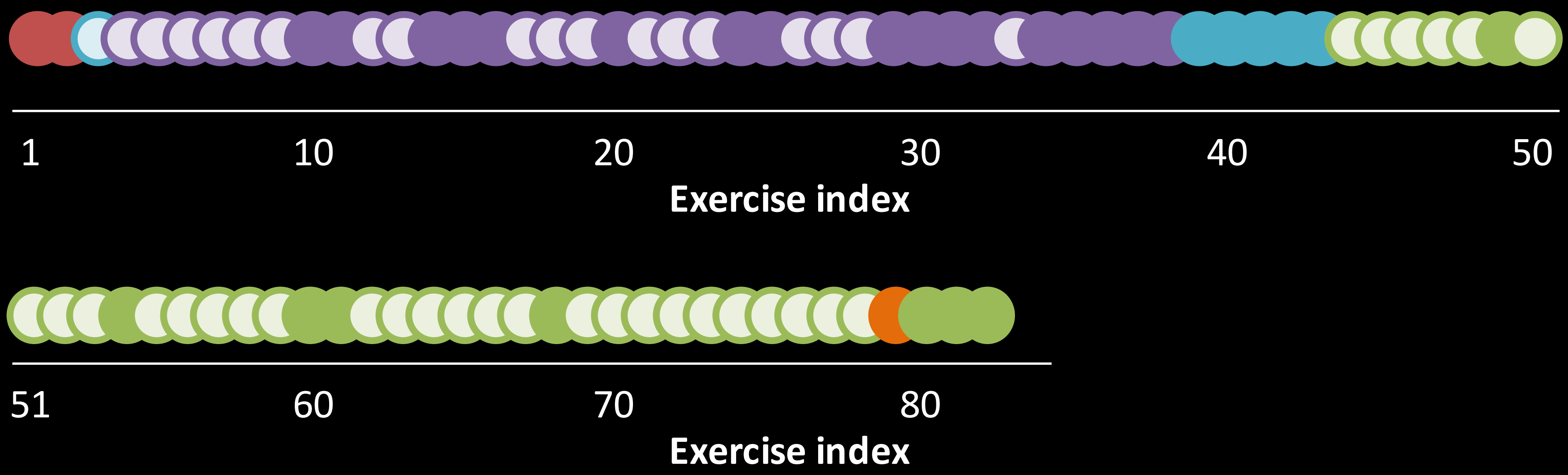
Exercise Type:

- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

Answer:

- Correct
- Incorrect

Story of Riley



Exercise Type:

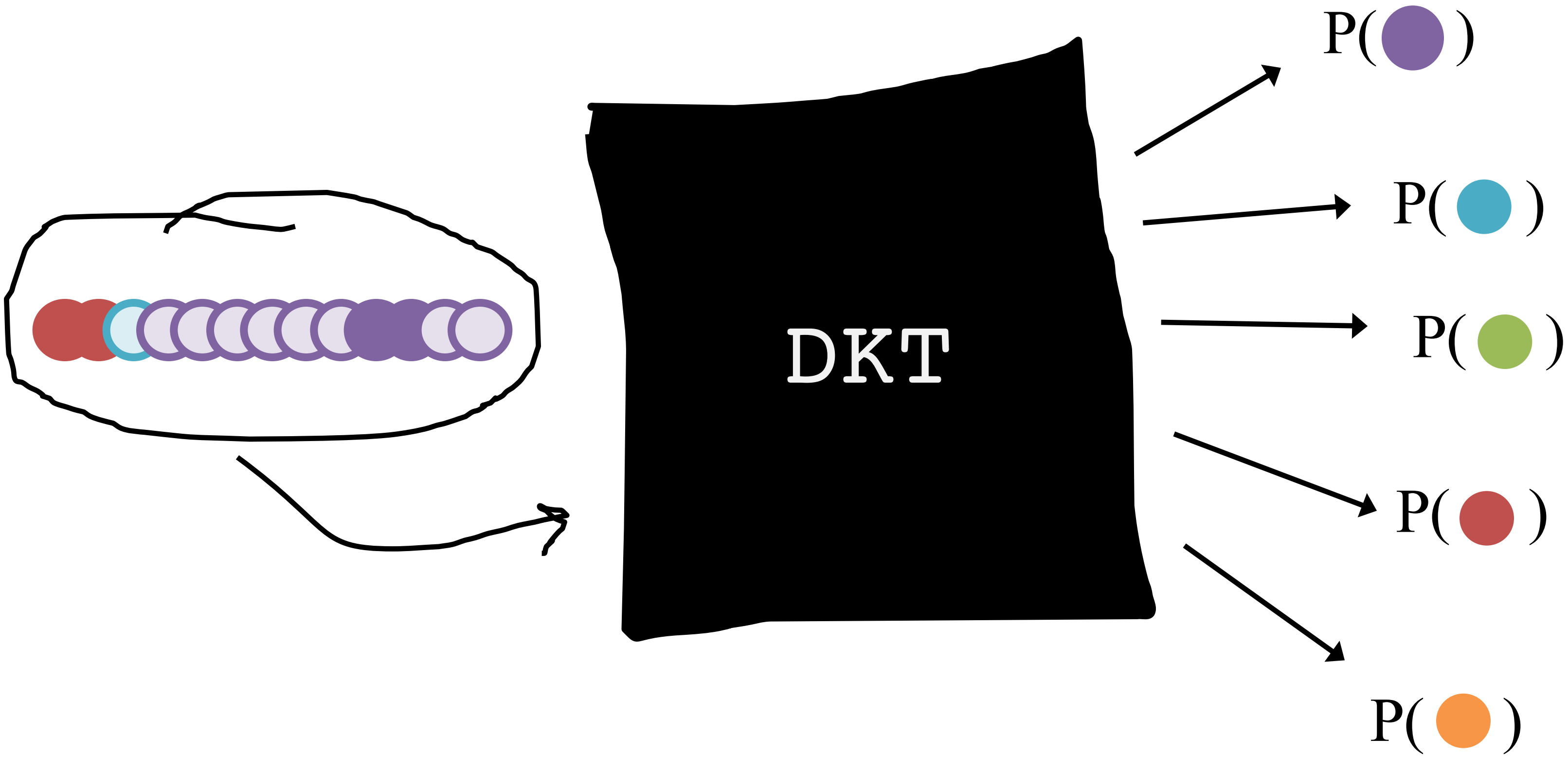
- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

Answer:

- Correct
- Incorrect

KHAN ACADEMY

We Have a Model that Can Predict the Future



There are millions of Rileys

#iamriley

Summary of Independence

Two events A and B are called **<u>independent</u>** if:

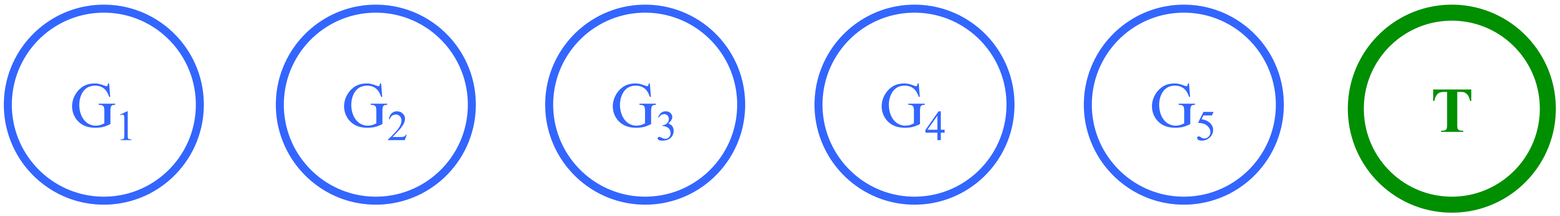
$$P(AB) = P(A)P(B) \qquad P(A|B) = P(A)$$

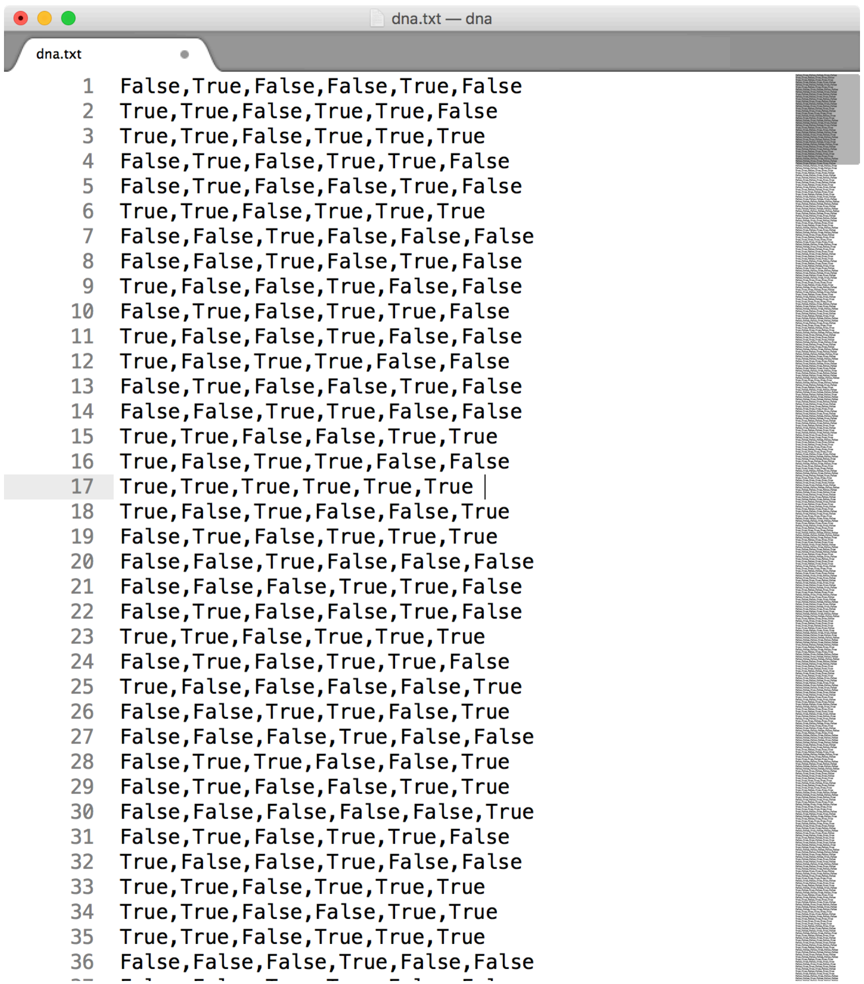
Otherwise, they are called **<u>dependent</u>** events

Two events A and B are
<u>conditionally independent</u> on C if:

$$P(AB|C) = P(A|C)P(B|C)$$

$$P(A|BC) = P(A|C)$$



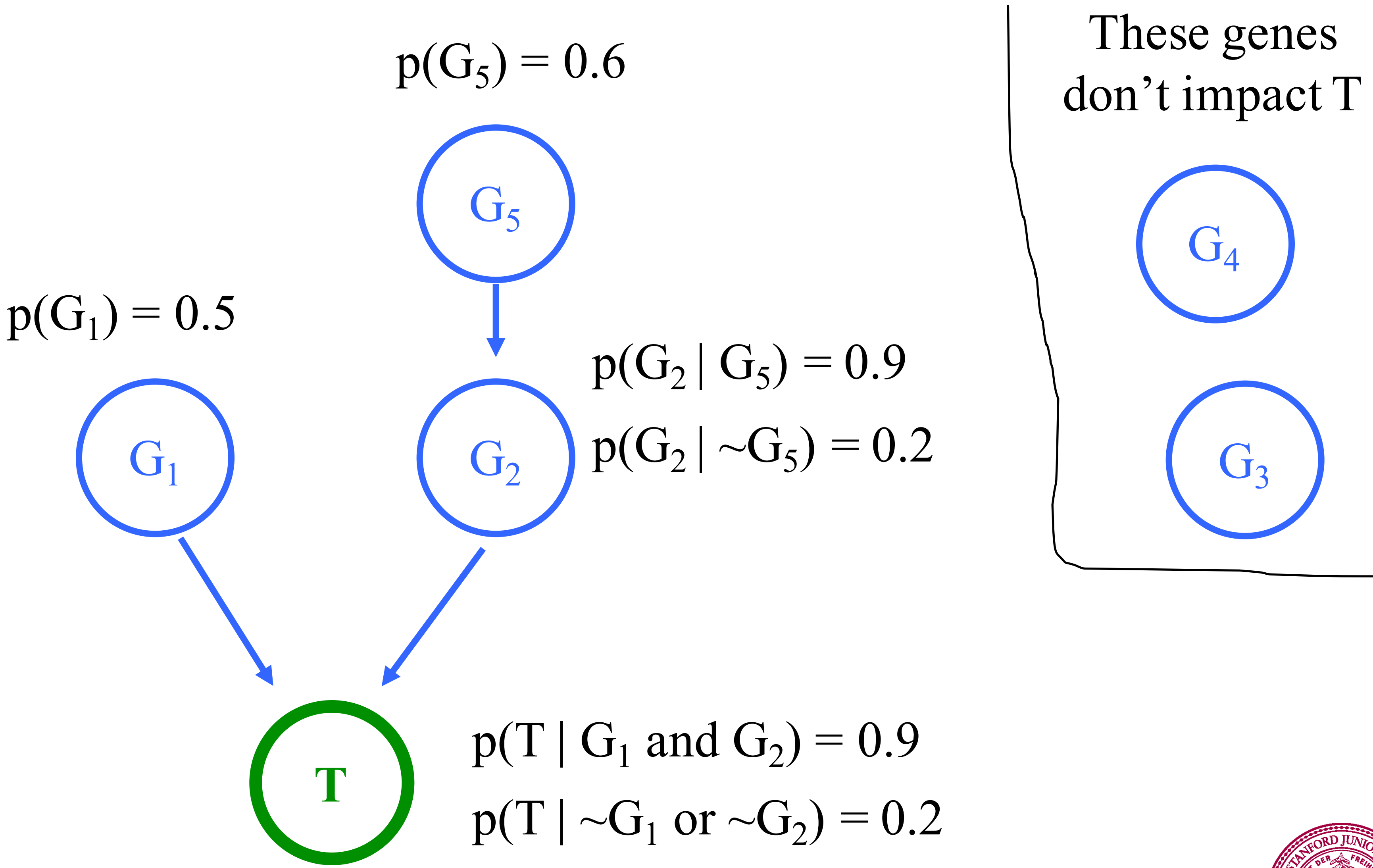



```
False,True,False,False,True,False
True,True,False,True,True,False
True,True,False,True,True,True
False,True,False,True,True,False
False,True,False,False,True,False
True,True,False,True,True,True
False,False,True,False,False,False
False,False,True,False,True,False
True,False,False,True,False,False
False,True,False,True,True,False
True,False,False,True,False,False
True,False,True,True,False,False
False,True,False,False,True,False
False,False,True,True,False,False
True,True,False,False,True,True
True,False,True,True,False,False
True,True,True,True,True,True
True,False,True,False,False,True
False,True,False,True,True,True
False,False,True,False,False,False
False,False,False,True,True,False
False,True,False,False,True,False
True,True,False,True,True,True
False,True,False,True,True,False
True,False,False,False,False,True
False,False,True,True,False,True
False,False,False,True,False,False
False,True,True,False,False,True
False,True,False,False,True,True
False,False,False,False,False,True
False,True,False,True,True,False
True,False,False,True,False,False
True,True,False,True,True,True
True,True,False,False,True,True
True,True,False,True,True,True
False,False,False,True,False,False
```

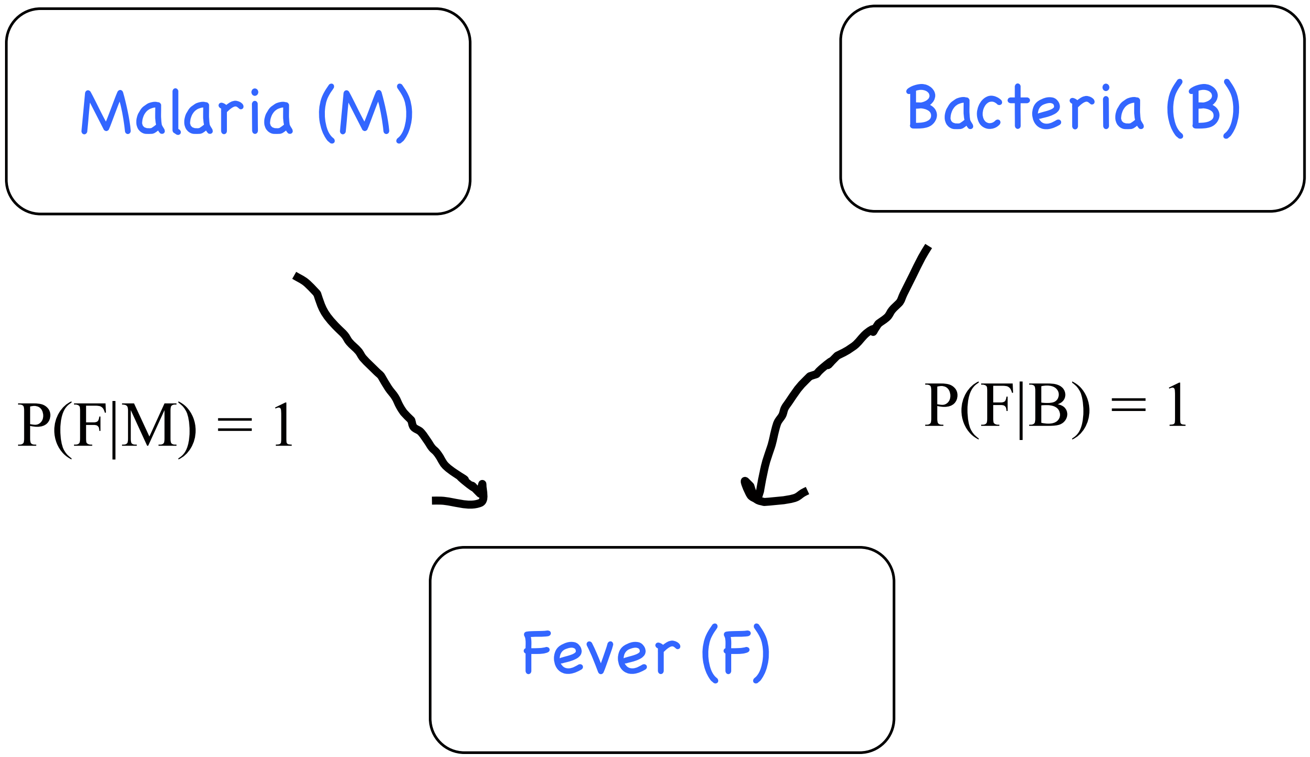
100,000 samples

6 observations per sample

Use Independence to Hypothesize



Causality



*This is a “causal” diagram. It helps explain why things are independent

Learning Goals

1. Be able to define a random variable (R.V.)
2. Be able to use + produce a PMF of a R.V.
3. Be able to use + produce a CDF of a R.V.
4. Be able to calculate the expectation of the R.V.

Remember Learning to Code?

name

value

type

```
int a = 5;
double b = 4.2;
bit c = 1;
choice d = medium;
```

$$z \in \{\text{high}, \text{medium}, \text{low}\}$$

Random Variable

- A **Random Variable** is a real-valued function defined on a sample space
- Example:
  - 3 fair coins are flipped.
  - Y = number of “heads” on 3 coins
  - Y is a random variable
  - $P(Y = 0) = 1/8$ (T, T, T)
  - $P(Y = 1) = 3/8$ (H, T, T), (T, H, T), (T, T, H)
  - $P(Y = 2) = 3/8$ (H, H, T), (H, T, H), (T, H, H)
  - $P(Y = 3) = 1/8$ (H, H, H)
  - $P(Y \geq 4) = 0$

It is confusing that both random variables and events use the same notation

Random variables are vars that probabilistically take on values.

Random variables and events are two different things

We can define an event to be a particular assignment to a random variables

Binary Random Variable

- A **binary** random variable is a random variable with 2 possible outcomes (e.g., coin flip).
- Let C be the outcome of a coin flip {Heads, Tails}
- C is a binary random variable!

- What is: $P(C = \text{Heads})$?
- What is $P(C = \text{Tails})$?
- What is: $P(C = \text{Heads} \cup C = \text{Tails})$?

Indicator Random Variable

- A variable I is called an indicator variable for event *A* if

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

- This is a binary random variable that takes on the values 1 and 0.

Indicator Random Variables
are awesome.

Sum of Random Variable

- A coin flip has 2 possible outcomes
  - Now consider *n* coin flips, each which independently come up heads with probability *p*
  - Y = number of “heads” on *n* flips
  - P(Y = k) = $\binom{n}{k} p^k (1-p)^{n-k}$, where *k* = 0, 1, 2, …, n
  - So, $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$
  - Proof: $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$

* Pro tip: no coin works like this… but many real world binary events do

Simple Game

- Urn has 11 balls (3 blue, 3 red, 5 black)
  - 3 balls drawn. +\$1 for blue, -\$1 for red, \$0 for black
  - Y = total winnings
  - P(Y = 0) = $\left[\binom{5}{3}+\binom{3}{1}\binom{3}{1}\binom{5}{1}\right]\bigg/\binom{11}{3}=\frac{55}{165}$
  - P(Y = 1) = $\left[\binom{3}{1}\binom{5}{2}+\binom{3}{2}\binom{3}{1}\right]\bigg/\binom{11}{3}=\frac{39}{165}$ = P(Y = -1)
  - P(Y = 2) = $\binom{3}{2}\binom{5}{1}\bigg/\binom{11}{3}=\frac{15}{165}$ = P(Y = -2)
  - P(Y = 3) = $\binom{3}{3}\bigg/\binom{11}{3}=\frac{1}{165}$ = P(Y = -3)

Fun with Random Variables

- Probability Mass Function:

$$P(X = a)$$

- Cumulative Distribution Function:

$$\left[\binom{5}{3} + \binom{3}{1}\binom{3}{1}\binom{5}{1}\right] \Big/ \binom{11}{3} = \frac{55}{165}$$

$$P(X \leq a)$$

$$\left[\binom{3}{1}\binom{5}{2} + \binom{3}{2}\binom{3}{1}\right] \Big/ \binom{11}{3} = \frac{39}{165}$$

- Expectation:

$$E[X] \quad \binom{3}{2}\binom{5}{1} \Big/ \binom{11}{3} = \frac{15}{165}$$

- Variance:

$$\binom{3}{3} \Big/ \binom{11}{3} = \frac{1}{165}$$

$$\mathrm{Var}(X)$$

Learning goals for today

1. Probability Mass Function

All the different assignments to a random variable make a function

$$P(Y = k)$$

This is a function

For example Y is the number of heads in 5 coin flips

Random Variables -> Functions

$$P(Y = k)$$

$$k = 5$$

$$0.03125$$

For example Y is the number of heads in 5 coin flips

Random Variables -> Functions

$$P(Y = k)$$

```
private double eventProbability(int k) {
        int ways = choose(N, k);
        double a = Math.pow(P, k);
        double b = Math.pow(P, N-k);
        return ways * a * b;
}

private static final int N = 5;
private static final double P = 0.6;
```

For example Y is the number of heads in 5 coin flips

If a random variable is iscrete we call this function the Probability Mass Function

Probability Mass Function

- A random variable *X* is discrete if it has countably many values (e.g., x_1, x_2, x_3, ...)

- Probability Mass Function (PMF) of a discrete random variable is:

$$p(a) = P(X = a)$$

- Since $\sum_{i=1}^{\infty} p(x_i) = 1$, it follows that:

$$P(X = a) = \begin{cases} p(x_i) \geq 0 \text{ for } i = 1, 2, \ldots \\ p(x) = 0 \;\; \text{otherwise} \end{cases}$$

where X can assume values x_1, x_2, x_3, …

Probability Mass Function for a Single Dice

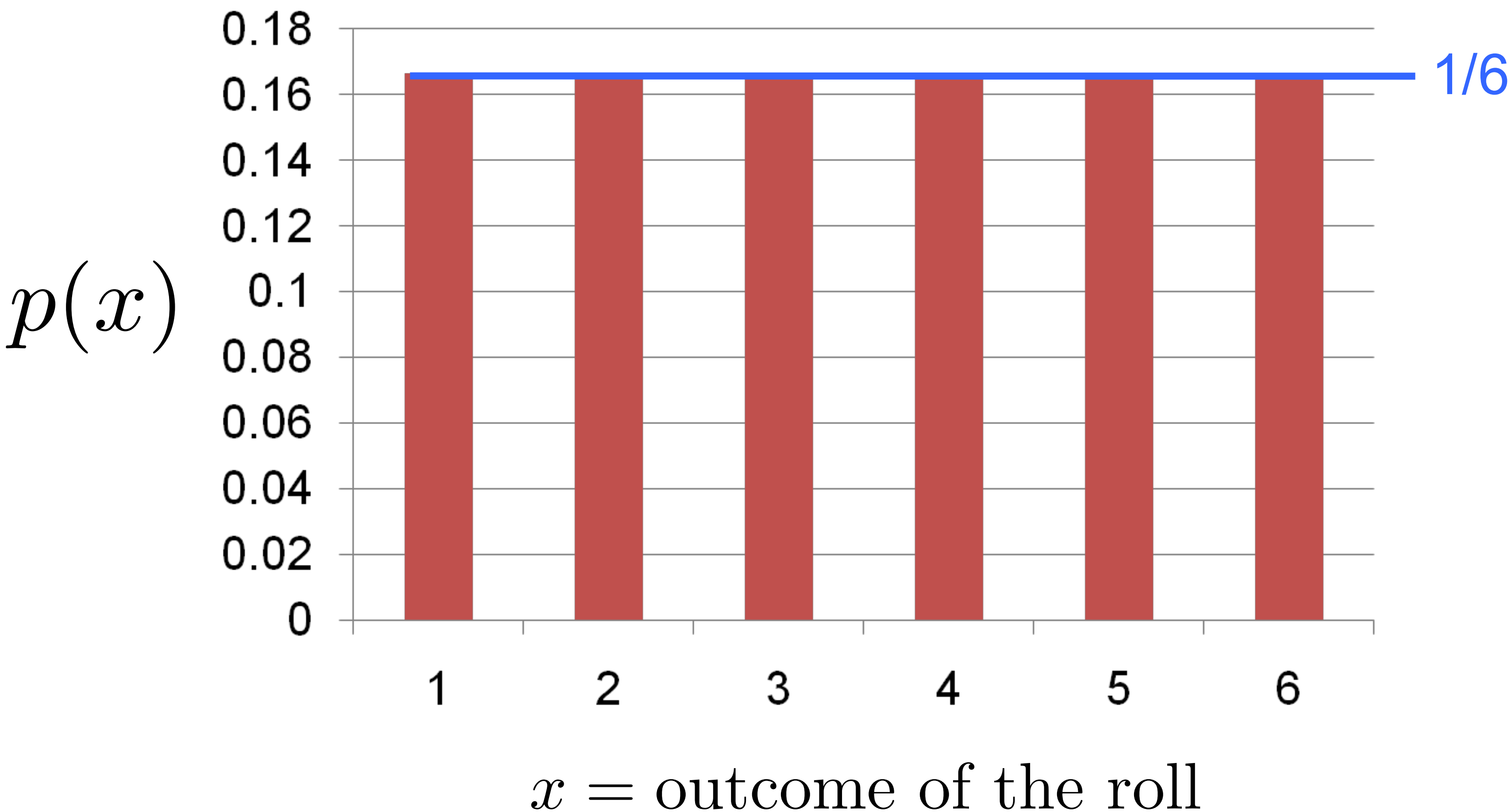
Let X be a random variable that represents the result of a **single dice roll**. X can take on the values {1, 2, 3, 4, 5, 6}

$$P(X = x)$$

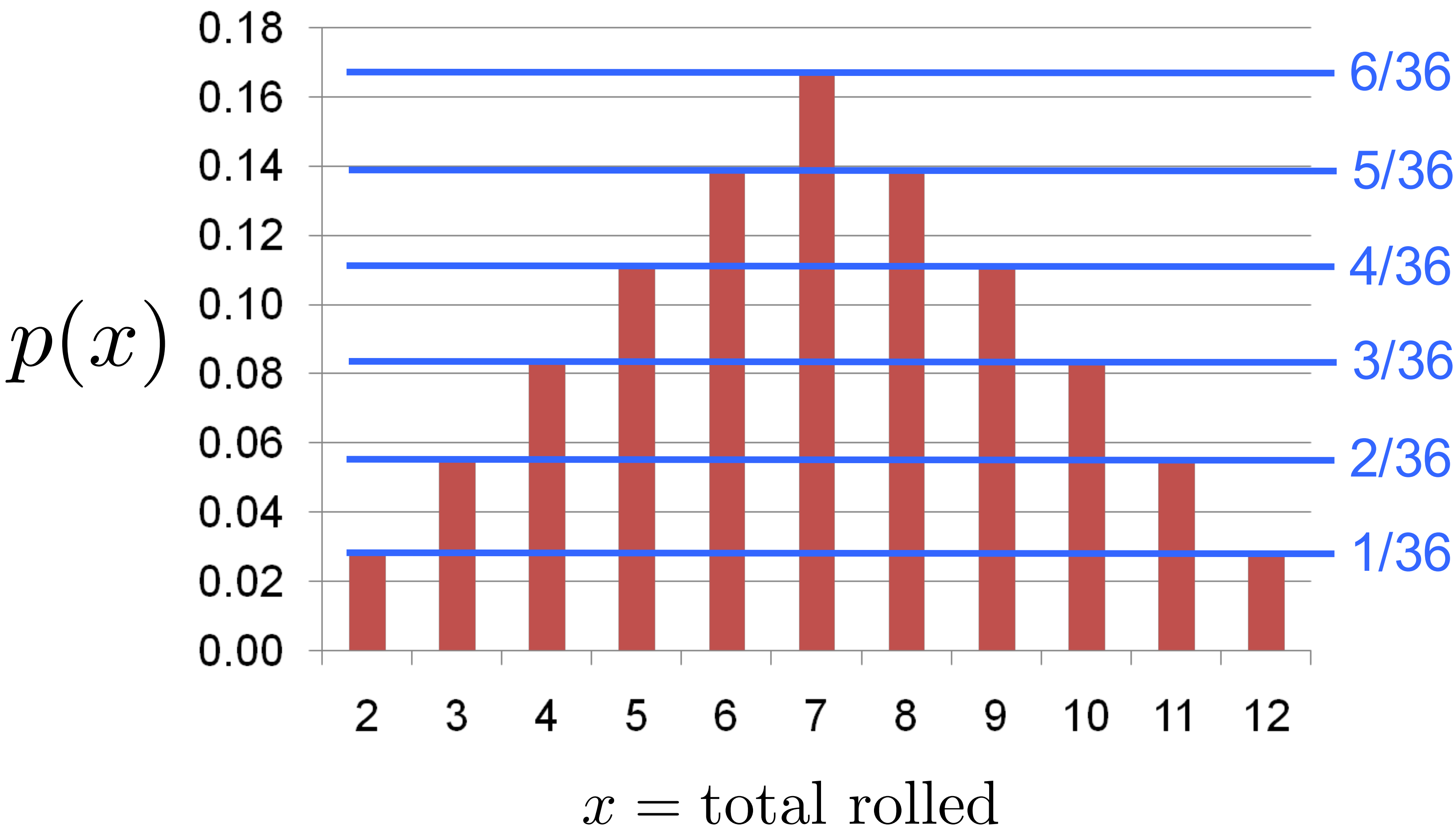
$$p(x)$$

This is shorthand notation for the PMF

PMF For a Single 6 Sided Dice



PMF for the sum of two dice



2. Cumulative Distribution Function

The event that a random variable takes on a value less than *a* also defines a function

Cumulative Distribution Function

- For a random variable X, the Cumulative Distribution Function (CDF) is defined as:

$$F(a) = P(X \leq a) \quad \text{where } -\infty < a < \infty$$

- The CDF of a <u>discrete</u> random variable is:

$$F(a) = P(X \leq a) = \sum_{\text{all } x \leq a} p(x)$$

Cumulative Distribution Function for a Single Dice

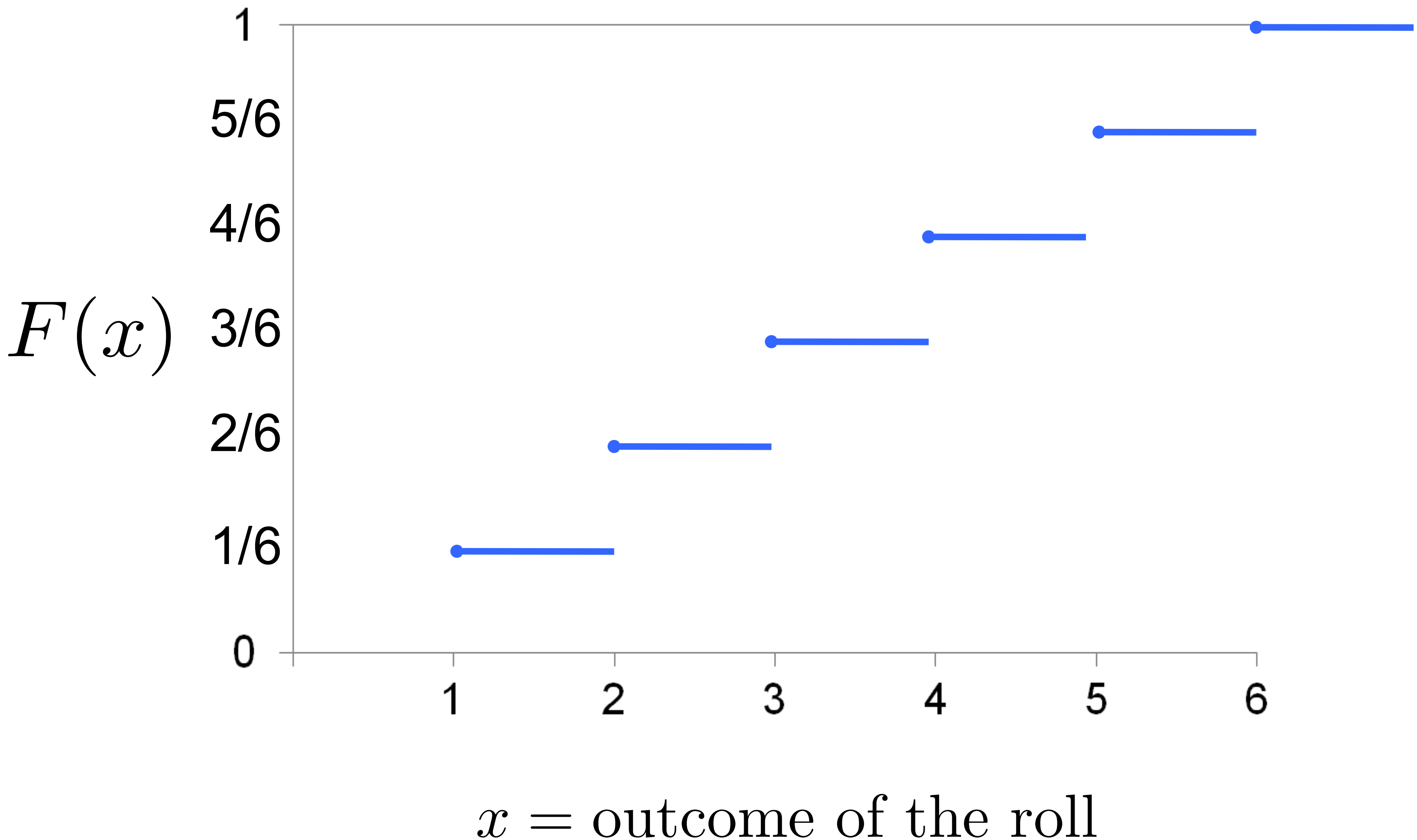
Let X be a random variable that represents the result of a **single dice roll**. X can take on the values {1, 2, 3, 4, 5, 6}

$$P(X \leq a)$$

$$F(x)$$

This is shorthand notation for the PMF

CDF for a 6 sided dice



3. Expectation

Expected Value

- The Expected Values for a discrete random variable X is defined as:

$$E[X] = \sum_{x:p(x)>0} x\,p(x)$$

- Note: sum over all values of x that have p(x) > 0.

- Expected value also called: *Mean*, *Expectation*, *Weighted Average*, *Center of Mass*, *1st Moment*

Expected Value

- Roll a 6-Sided Die. X is outcome of roll
  - p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = 1/6
- E[X] = $1\left(\frac{1}{6}\right)+2\left(\frac{1}{6}\right)+3\left(\frac{1}{6}\right)+4\left(\frac{1}{6}\right)+5\left(\frac{1}{6}\right)+6\left(\frac{1}{6}\right)=\frac{7}{2}$

- Y is random variable
  - P(Y = 1) = 1/3,   P(Y = 2) = 1/6,   P(Y = 3) = 1/2
- E[Y] = 1 (1/3) + 2 (1/6) + 3 (1/2) = 13/6

Indicator Variable

- A variable I is called an indicator variable for event A if

$$I = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

- What is E[I]?
  - p(I =1) = P(A), p(I = 0) = 1 – P(A)
  - E[I] = 1 P(A) + 0 (1 – P(A)) = P(A)

We'll use this property frequently!

Lying with Statistics

"There are three kinds of lies:
lies, damned lies, and statistics"

– *Mark Twain*

- School has 3 classes with 5, 10 and 150 students
- Randomly choose a <u>class</u> with equal probability
- X = size of chosen class
- What is E[X]?
  - $E[X] = 5\,(1/3) + 10\,(1/3) + 150\,(1/3)$
    $= 165/3 = 55$

Lying with Statistics

"There are three kinds of lies:
lies, damned lies, and statistics"
– *Mark Twain*

- School has 3 classes with 5, 10 and 150 students
- Randomly choose a <u>student</u> with equal probability
- Y = size of class that student is in
- What is E[Y]?
  - $E[Y] = 5\,(5/165) + 10\,(10/165) + 150\,(150/165)$
    $= 22635/165 \approx 137$
- Note: E[Y] is students' perception of class size
  - But E[X] is what is usually reported by schools!

More on Expectation

Properties of Expectation

- Linearity:

$$E[aX + b] = aE[X] + b$$

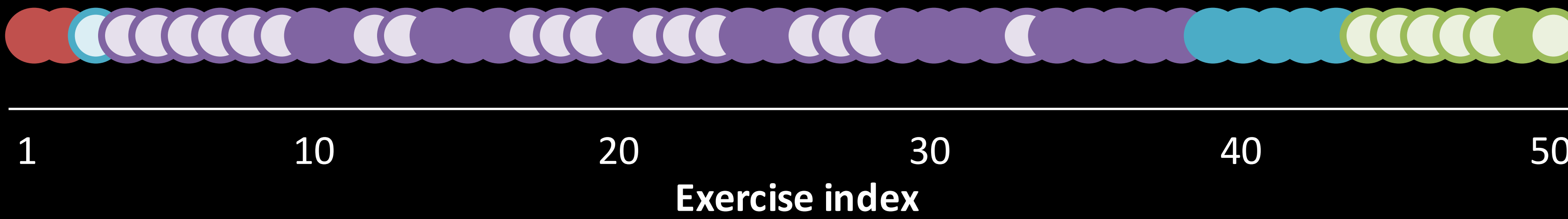
  - Consider X = 6-sided die roll, Y = 2X – 1.
  - E[X] = 3.5 E[Y] = 6

- *N*-th Moment of X:

$$E[X^n] = \sum_{x:p(x)>0} x^n \cdot p(x)$$

  - We'll see the 2nd moment soon...

Story of Riley



What should Riley do next?

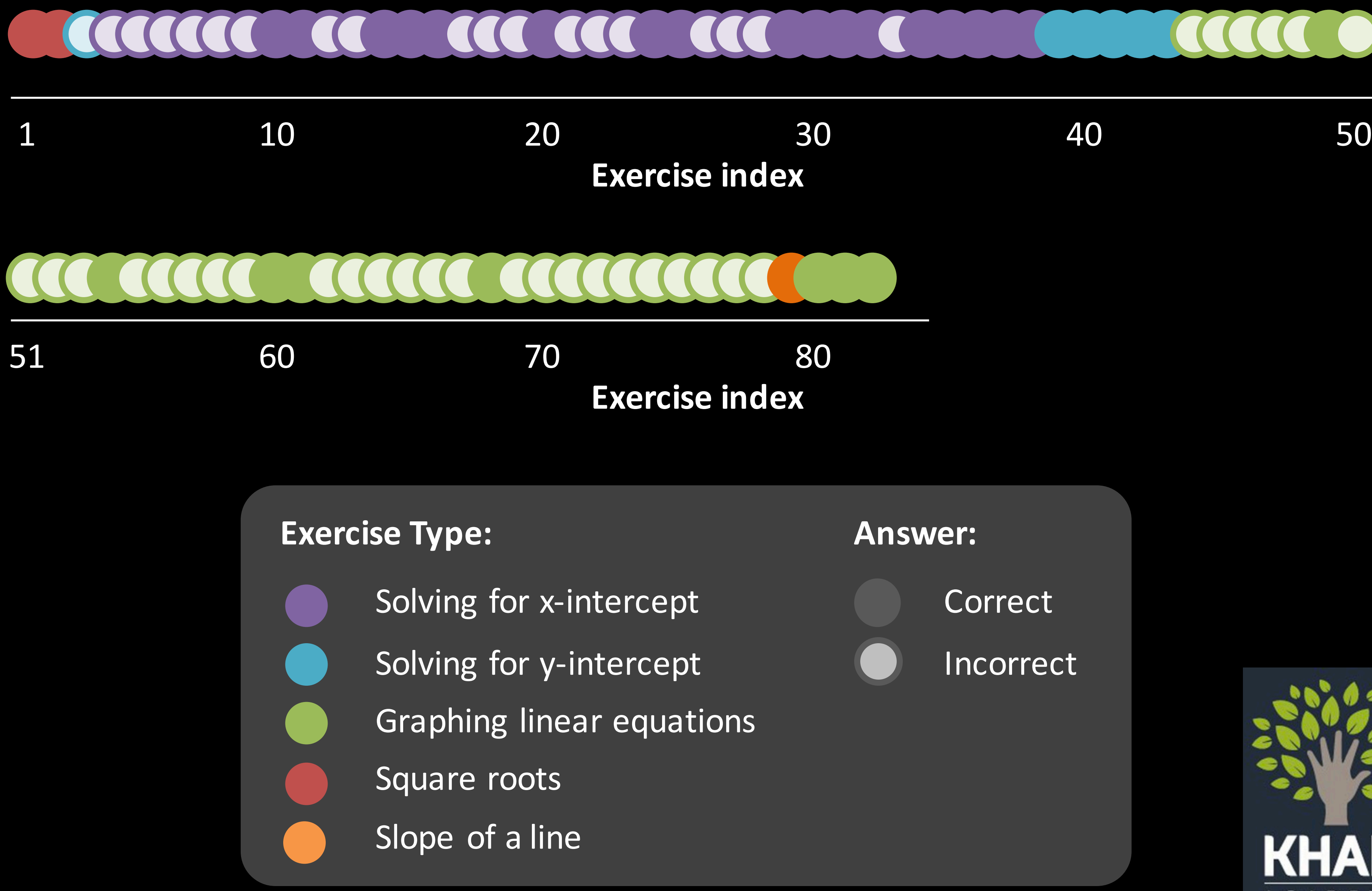
Exercise Type:

- Solving for x-intercept
- Solving for y-intercept
- Graphing linear equations
- Square roots
- Slope of a line

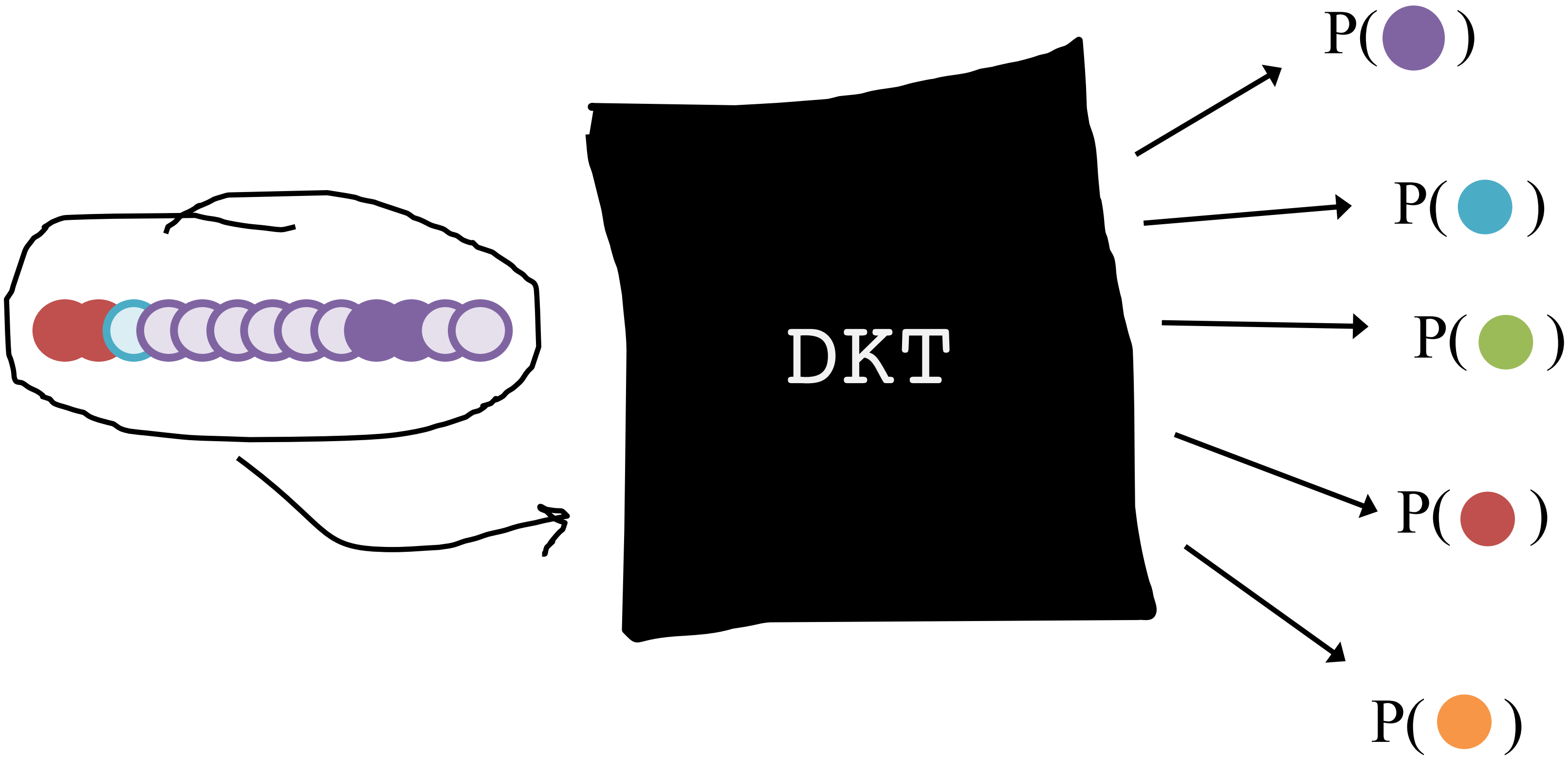
Answer:

- Correct
- Incorrect

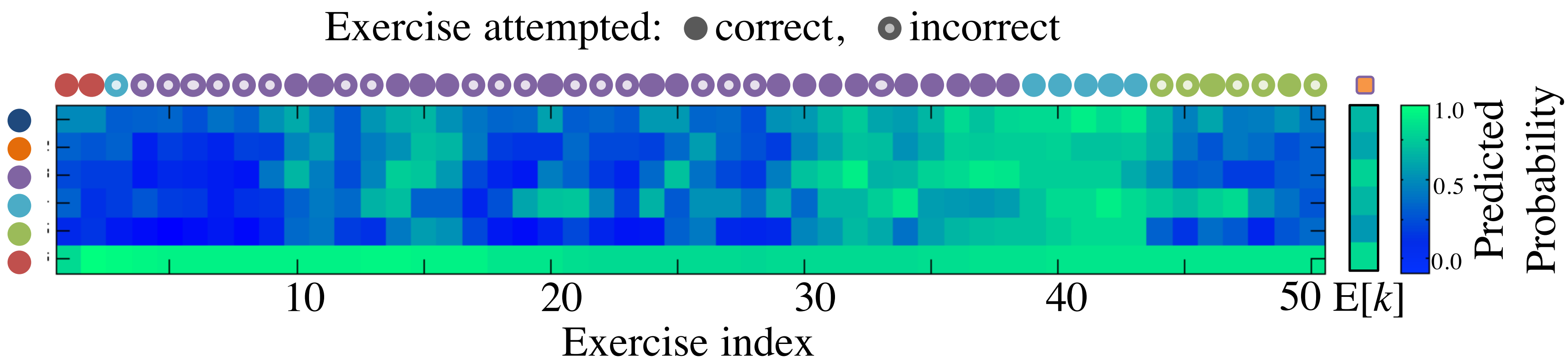
Story of Riley



We Have a Model that Can Predict the Future



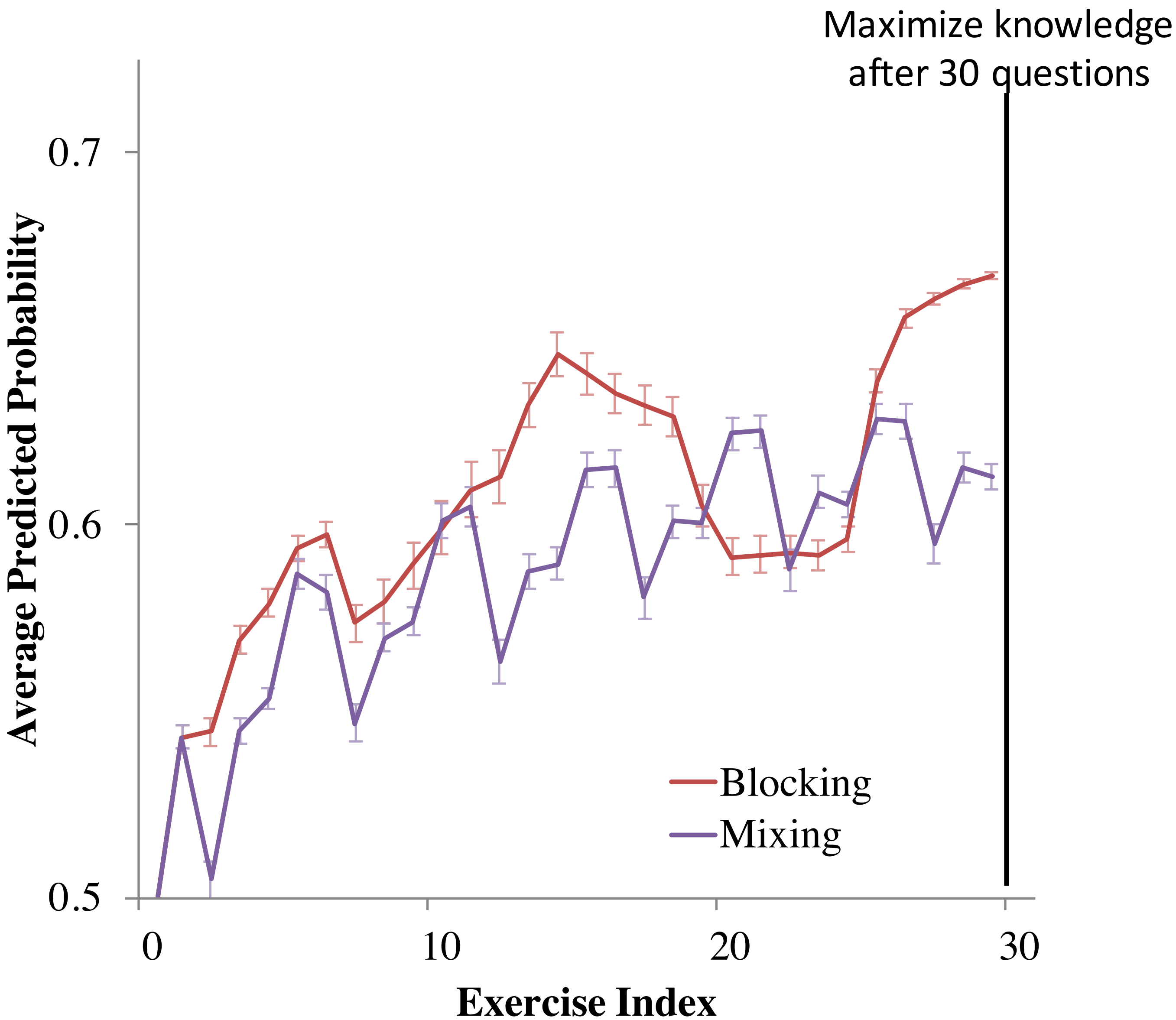
Predicting Next Item



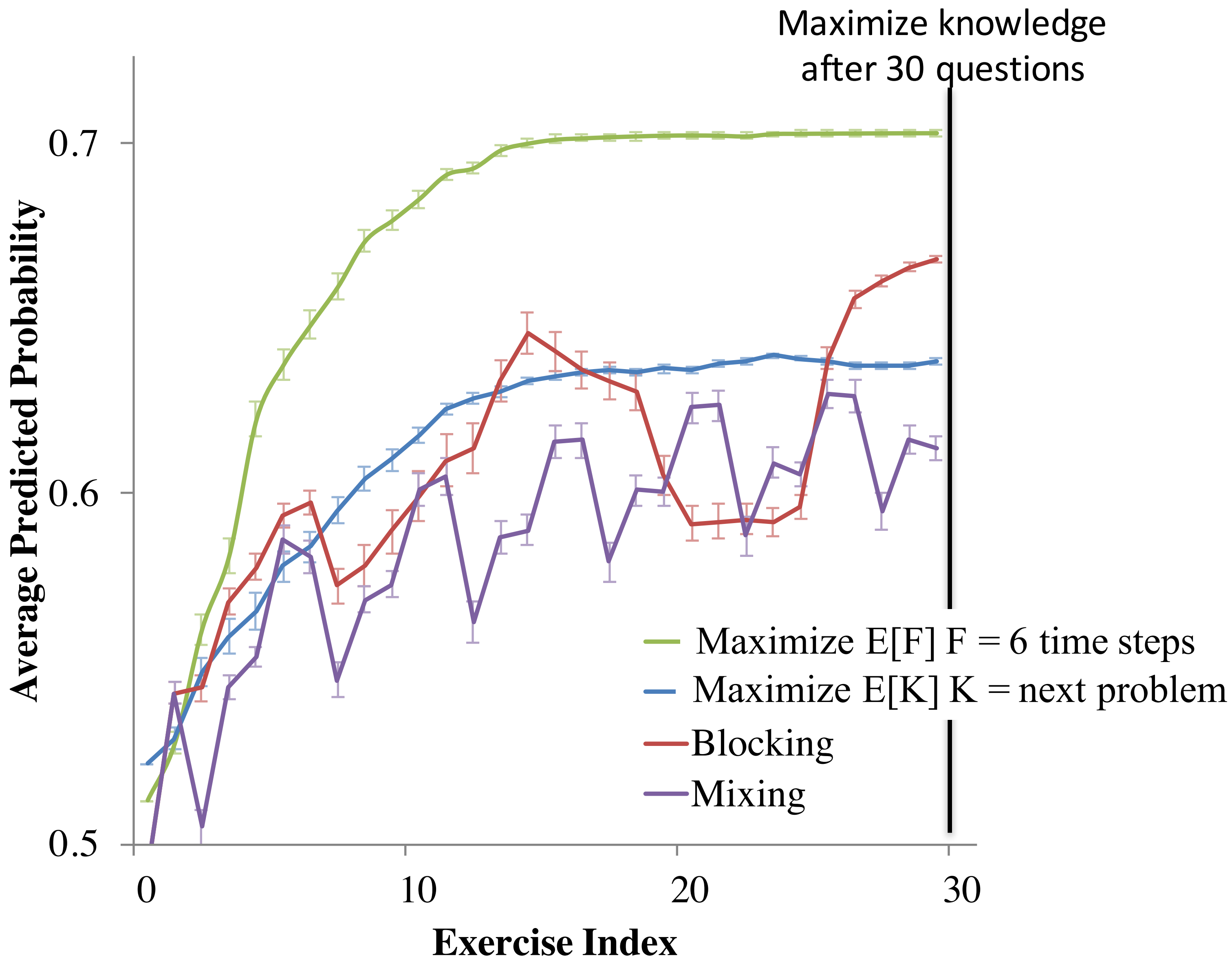
- Line graph intuition
- Slope of a line
- Solving for x
- Solvi
- quations
- ots

We can find what is the "deal maker"

Optimal Curriculum



Optimal Curriculum



Wonderful

St Petersburg

- Game set-up
  - We have a fair coin (come up "heads" with p = 0.5)
  - Let n = number of coin flips ("heads") before first "tails"
  - You win $\$2^n$

- How much would you pay to play?

St Petersburg

- Game set-up
  - We have a fair coin (come up "heads" with p = 0.5)
  - Let n = number of coin flips ("heads") before first "tails"
  - You win $\$2^n$
- How much would you pay to play?
- Solution
  - Let X = your winnings
  - E[X] = $\left(\frac{1}{2}\right)^1 2^0 + \left(\frac{1}{2}\right)^2 2^1 + \left(\frac{1}{2}\right)^3 2^2 + \left(\frac{1}{2}\right)^4 2^3 + \ldots \;\; = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^{i+1} 2^i$

    $= \sum_{i=0}^{\infty} \frac{1}{2} = \infty$
  - I'll let you play for $1 thousand… but just once! Takers?

St Petersburg + Reality

- What if Chris has only $65,536?
  - Same game
  - If you win over $65,536 I leave the country.
- Solution
  - Let X = your winnings
  - E[X] $= \left(\frac{1}{2}\right)^1 2^0 + \left(\frac{1}{2}\right)^2 2^1 + \left(\frac{1}{2}\right)^3 2^2 + \left(\frac{1}{2}\right)^4 2^3 + \ldots$

$$= \sum_{i=0}^{k} \left(\frac{1}{2}\right)^{i+1} 2^i \text{ s.t. } k = \log_2(65,536)$$

$$= \sum_{i=0}^{16} \frac{1}{2} = 8.5$$

Utility

- Utility is value of some choice
  - 2 options, each with n consequences: $c_1, c_2, \ldots, c_n$
  - One of c_i will occur with probability p_i
  - Each consequence has some value (utility): $U(c_i)$
  - Which choice do you make?
- Example: Buy a $1 lottery ticket (for $1M prize)?
  - Probability of winning is $1/10^7$
  - **Buy**: c_1 = win, c_2 = lose, $U(c_1) = 10^6 - 1$, $U(c_2) = -1$
  - **Don't Buy**: c_1 = lose, $U(c_1) = 0$
  - $E(\text{buy}) = 1/10^7 (10^6 - 1) + (1 - 1/10^7)(-1) \approx -0.9$
  - $E(\text{don't buy}) = 1 (0) = 0$
  - *"You can't lose if you don't play!"*

And Then There's This

Recall, Geometric Series

$$a^0 + a^1 + a^2 + \ldots$$

$$= \sum_{i=0}^{\infty} a^i$$

$$= \frac{1}{1-a}$$

$$\text{where } 0 < a < 1$$

Breaking Vegas

- Consider even money bet (e.g., bet "Red" in roulette)
  - p = 18/38 you win \$Y, otherwise (1 – p) you lose \$Y
  - Consider this algorithm for one series of bets:
    1. Y = \$1
    2. Bet Y
    3. If Win then stop
    4. If Loss then Y = 2 * Y, goto 2
  - Let Z = winnings upon stopping
  - E[Z] $= \left(\frac{18}{38}\right)1 + \left(\frac{20}{38}\right)\left(\frac{18}{38}\right)(2-1) + \left(\frac{20}{38}\right)^2\left(\frac{18}{38}\right)(4-2-1) + \ldots$

$$= \sum_{i=0}^{\infty}\left(\frac{20}{38}\right)^i\left(\frac{18}{38}\right)\left(2^i - \sum_{j=0}^{i-1} 2^j\right) = \left(\frac{18}{38}\right)\sum_{i=0}^{\infty}\left(\frac{20}{38}\right)^i = \left(\frac{18}{38}\right)\frac{1}{1-\frac{20}{38}} = 1$$

  - Expected winnings ≥ 0. Use algorithm infinitely often!

Vegas Breaks You

- Why doesn't everyone do this?
  - Real games have maximum bet amounts
  - You have finite money
    - Not able to keep doubling bet beyond certain point
  - Casinos can kick you out
- But, if you had:
  - No betting limits, and
  - Infinite money, and
  - Could play as often as you want...
- Then, go for it!
  - And tell me which planet you are living on