Dr. Denton Bobeldyk

CIS 365 Artificial Intelligence

Uncertainty in AI

Week in Review

Blackboard Check-in

Uncertainty in AI

- Methods to cope with uncertainty:
  - Fuzzy Logic
  - Probability Theory

Uncertainty in AI - Fuzzy Sets

- Crisp Sets:
  - Standing = {x | x is a student standing in your class}
  - Sitting = {y | y is a student sitting in your class}
- The intersection of the two above sets is an empty set. A student belongs to one or the other.

$$Standing \cap Sitting = \emptyset$$

Uncertainty in AI - Fuzzy Sets

Notation Note:
x such that x is a student standing in your class
used for set notation

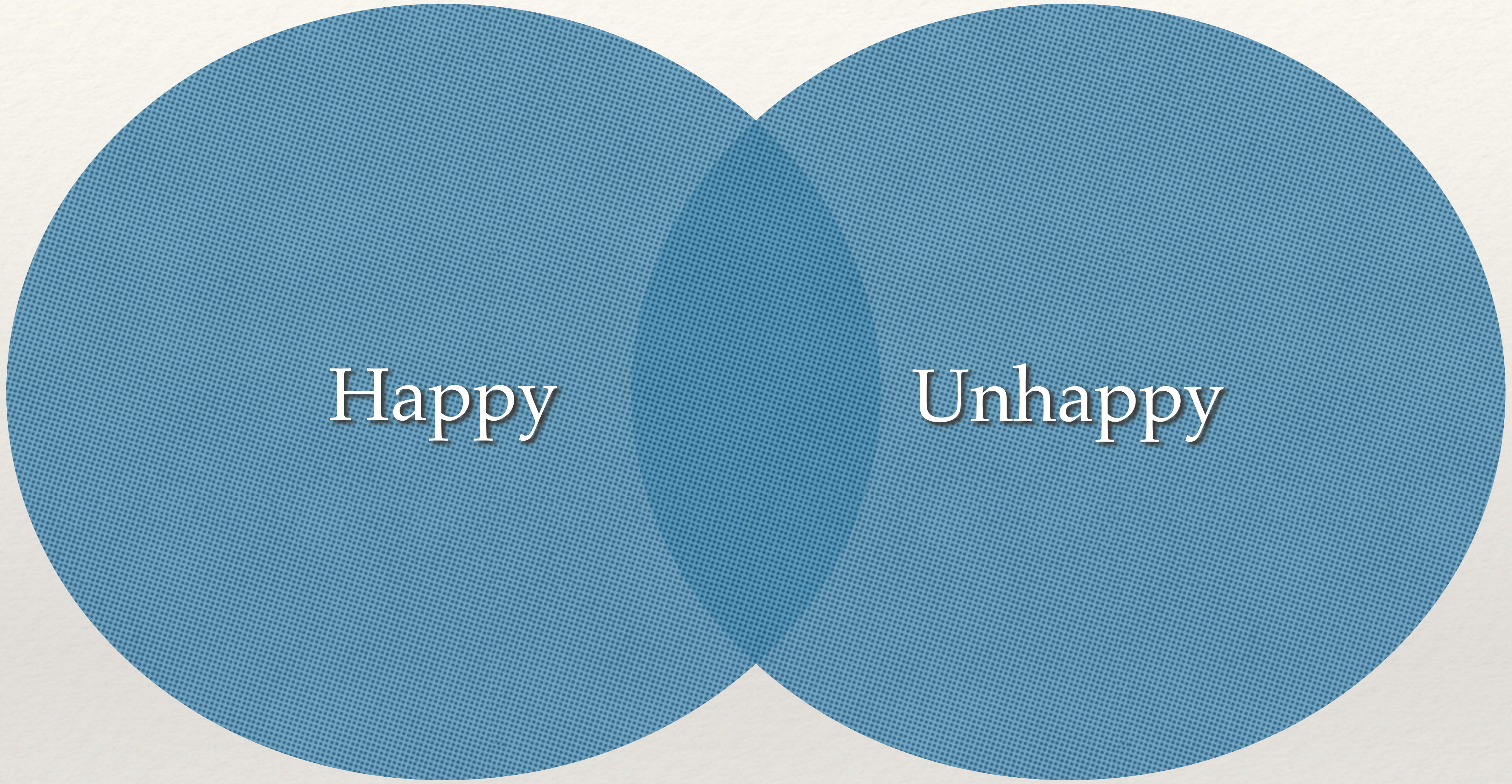
- Crisp Sets:
  - Standing = {x | x is a student standing in your class}
  - Sitting = {y | y is a student sitting in your class}
- The intersection of the two above sets is an empty set. A student belongs to one or the other.

$$Standing \cap Sitting = \varnothing$$

Fuzzy Sets

- Are you happy with your job?
- Are you unhappy with your job?

Fuzzy Sets



It is possible to be Happy and Unhappy about your job at the same time

Uncertainty in AI

- What about concepts like ‘tallness’? Is someone that is 5 foot, someone that would be considered tall?

Uncertainty in AI

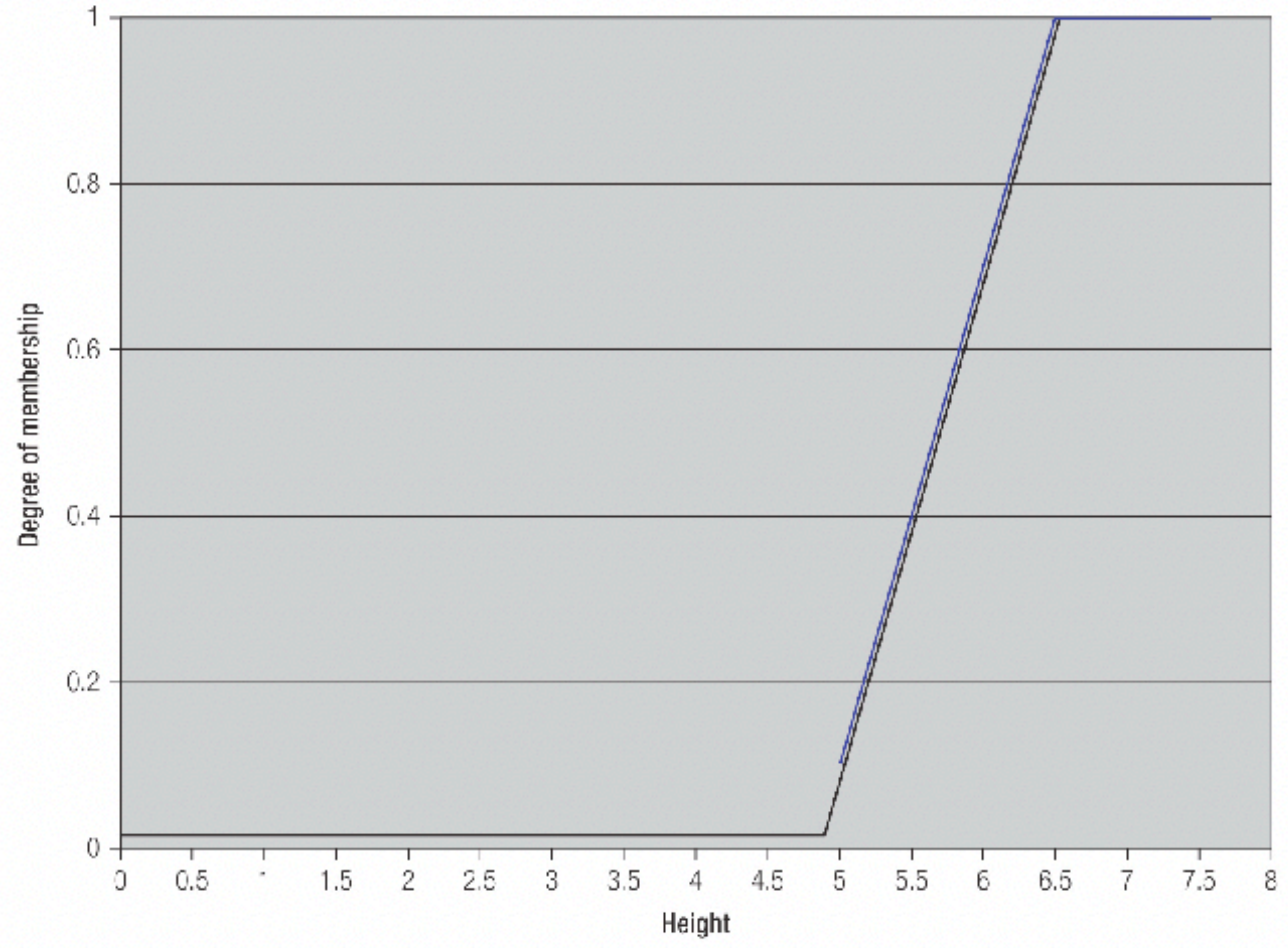


Figure 8.2

Membership function for the set of tall people.

Probability Theory

Probability Theory

- Sample Space for an experiment S is the set of all possible outcomes

Probability Theory

- Sample space of flipping one coin = {H, T}
- Sample space of flipping two coins = {(H,H), (H,T), (T, H), (T,T)}

Probability Theory

- Sample space of flipping one coin = {H, T}
- Probability of Heads = 0.5
  - P(H) = {H, T} = 0.5
- Probability of Tails = 0.5
  - P(T) = {H, T} = 0.5

Probability Theory

- Sample space of flipping two coins = {(H,H), (H,T), (T, H), (T,T)}
- Probability of outcomes:
  - P(H,H) = {(H,H), (H,T), (T, H), (T,T)} = 1/4 = 0.25
  - P(H, T) = {(H,H), (H,T), (T, H), (T,T)} = 1/4 = 0.25
  - P(T, H) = {(H,H), (H,T), (T, H), (T,T)} = 1/4 = 0.25
  - P(T, T) = {(H,H), (H,T), (T, H), (T,T)} = 1/4 = 0.25

Probability Theory

- Rolling a single die
  - What’s the sample set?
  - What’s the probability that each event will occur?

Probability Theory

- Rolling a single die
  - What's the sample set? {1, 2, 3, 4, 5, 6}
  - What's the probability that each event will occur?
    - P(1) = {1, 2, 3, 4, 5, 6} = 1/6 = 16.67%
    - P(2) = {1, 2, 3, 4, 5, 6} = 1/6 = 16.67%
    - …

Probability Theory

- Rolling 2 dice
    - What's the sample set?
    - What's the probability that each event will occur?

Probability Theory

- Rolling 2 dice
  - What's the sample set?
    - {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
    - (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
    - (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}
  - What's the probability that each event will occur?

Probability Theory

- Rolling 2 dice:
  - P(1,1) =
    - {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
    - (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
    - (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}
  - P(1,1) = P(Event/Number of items in sample space) = 1/36

Probability Theory

- Rolling a 2 dice:
  - P(Sum of two dice = 4)
    - {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
    - (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
    - (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}
  - P(Sum of two dice = 4) = ??

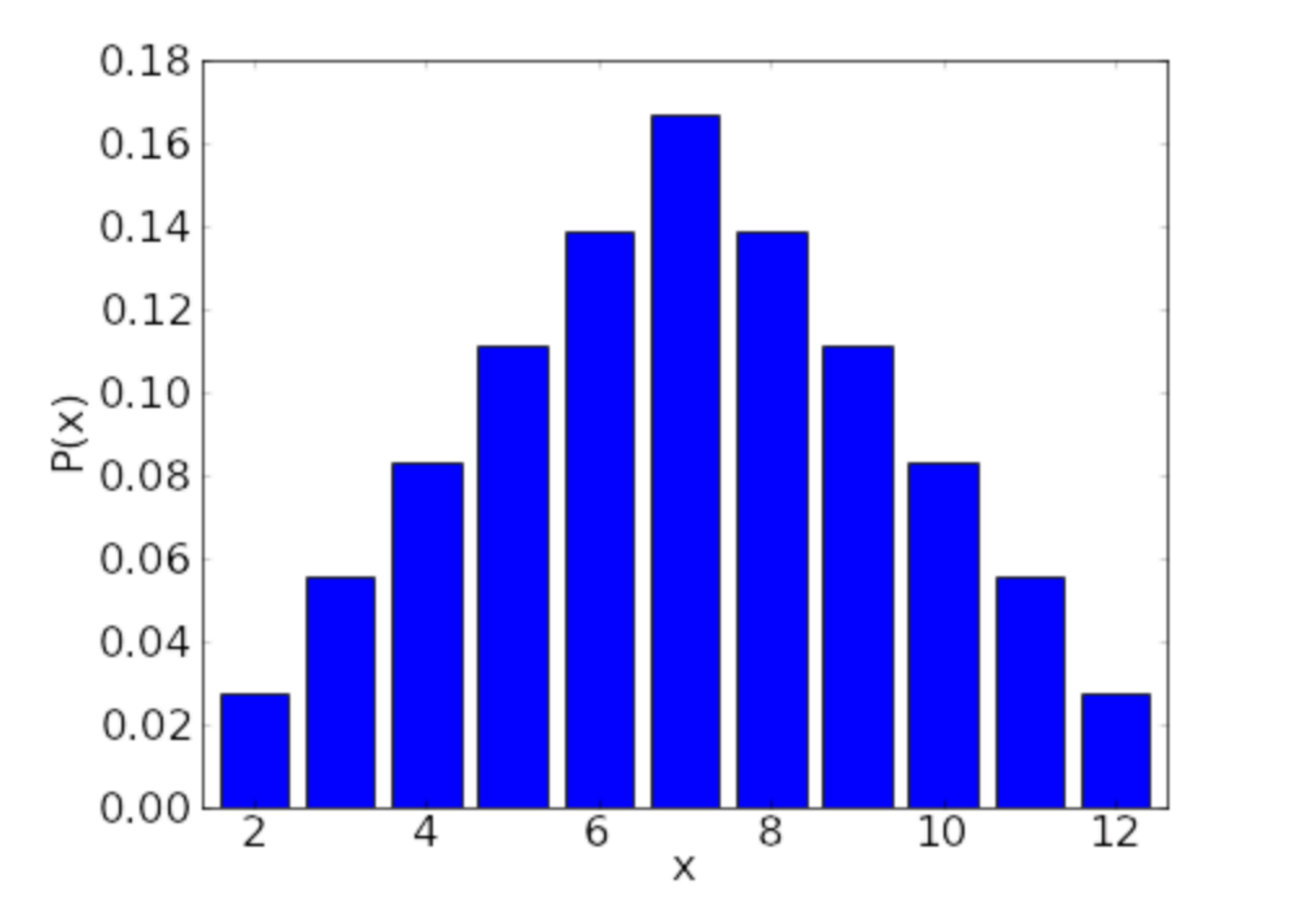
Probability Theory

- Rolling a 2 dice:
  - P(Sum of two dice = 4)
    - {(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6),
    - (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6),
    - (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}
  - P(Sum of two dice = 4) = 3/36 = 8.3%

Probability Theory

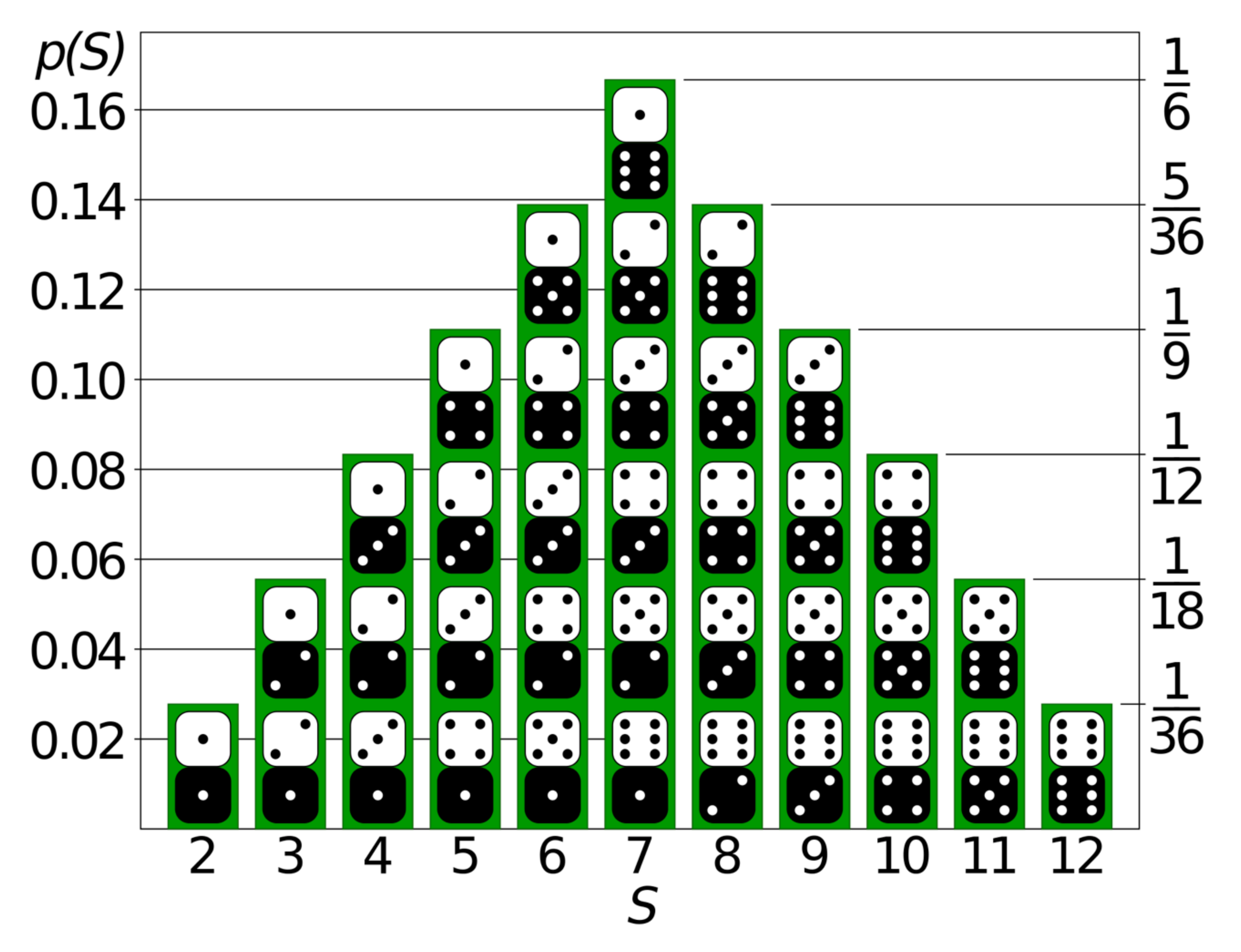
- Probability Distribution for the sum of two dice is?
- If we know the probability distribution, we can determine:
  - What sum has the highest probability?
  - What sum has the lowest probability?
  - Cumulative distribution (i.e., what's the probability of rolling less than 8)

Probability Distribution



https://mathinsight.org/image/two_dice_distribution

Probability Distribution



https://math.stackexchange.com/questions/1204396/why-is-the-sum-of-the-rolls-of-two-dices-a-binomial-distribution-what-is-define

Probability Distribution

Three basic axioms of probability:

1. For any event E: P(E) >= 0
2. P(S) = 1 where S is the sample space of an experiment
3. If the events E_1 and E_2 are mutually exclusive then $P(E_1 \cap E_2) = P(E_1) + P(E_2)$

Class Exercise

- Probability Distribution for 1d6 and 2d6 have been covered
- What is the sample space for:
  - Rolling 1d4
  - Rolling 2d4
- What is the probability distribution:
  - Rolling 1d4
  - Rolling 2d4
  - Rolling 3d4

Bayesian Classifier

- Probability distribution
  - Rolling 1d4
  - Rolling 2d4
  - Rolling 3d4
- Given the following numbers, what is the probability that it came from each of the above distributions?
  - 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

Bayesian Classifier

- Distributions: 1d4, 2d4, 3d4
- Given the following numbers, what is the probability that it came from each of the above distributions?
  - 1:
    - Probability of 1 Given 1d4 = ?
    - Probability of 1 Given 2d4 = ?
    - Probability of 1 Given 3d4 = ?

Bayesian Classifier

- Distributions: 1d4, 2d4, 3d4
- Given the following numbers, what is the probability that it came from each of the above distributions?
  - 2:
    - Probability of 2 Given 1d4 = ?
    - Probability of 2 Given 2d4 = ?
    - Probability of 2 Given 3d4 = ?

Bayesian Classifier

- Distributions: 1d4, 2d4, 3d4
- Given the following numbers, what is the probability that it came from each of the above distributions?
  - Complete for each of the following numbers: 1 - 12

Completed Answer Next Slide

Probability Distribution

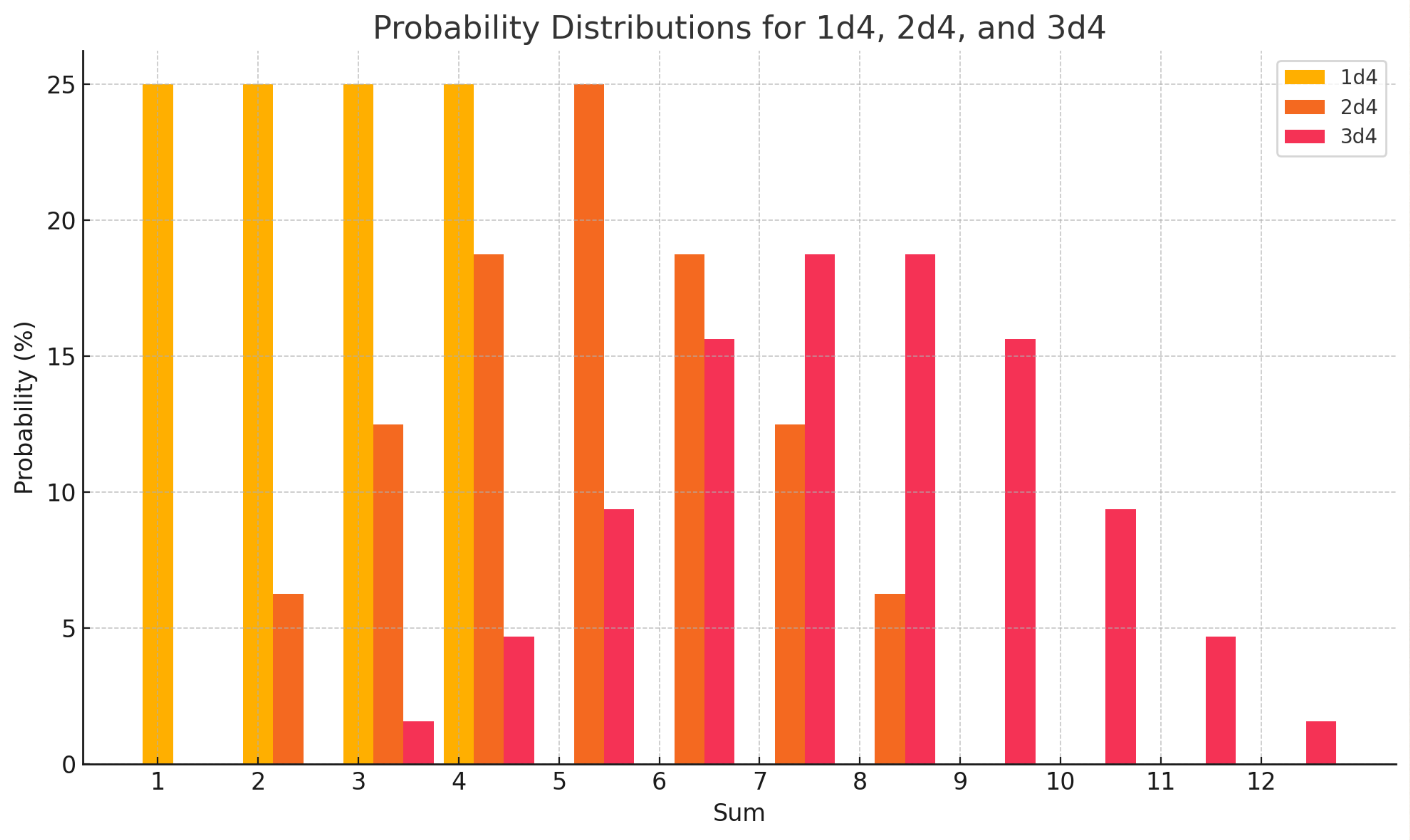
	1d4	2d4	3d4
1	25	0	0
2	25	6.25	0
3	25	12.5	1.6
4	25	18.75	4.7
5	0	25	9.4
6	0	18.75	15.6
7	0	12.5	18.8
8	0	6.25	18.8
9	0	0	15.6
10	0	0	9.4
11	0	0	4.7
12	0	0	1.6

Which distribution did it come from?

- Can we determine if given an outcomes what distribution it came from?

Bayesian Classifier - Which Distribution Generated the outcome?

Outcome	1d4	2d4	3d4
1	25	0	0
2	25	6.25	0
3	25	12.5	1.6
4	25	18.75	4.7
5	0	25	9.4
6	0	18.75	15.6
7	0	12.5	18.8
8	0	6.25	18.8
9	0	0	15.6
10	0	0	9.4
11	0	0	4.7
12	0	0	1.6



Please note the X axis is slightly misaligned

D4 Distributions

- Does a formula exist, if we know the number of d4 we are using, what the probability distribution is?

Please note the X axis is slightly misaligned

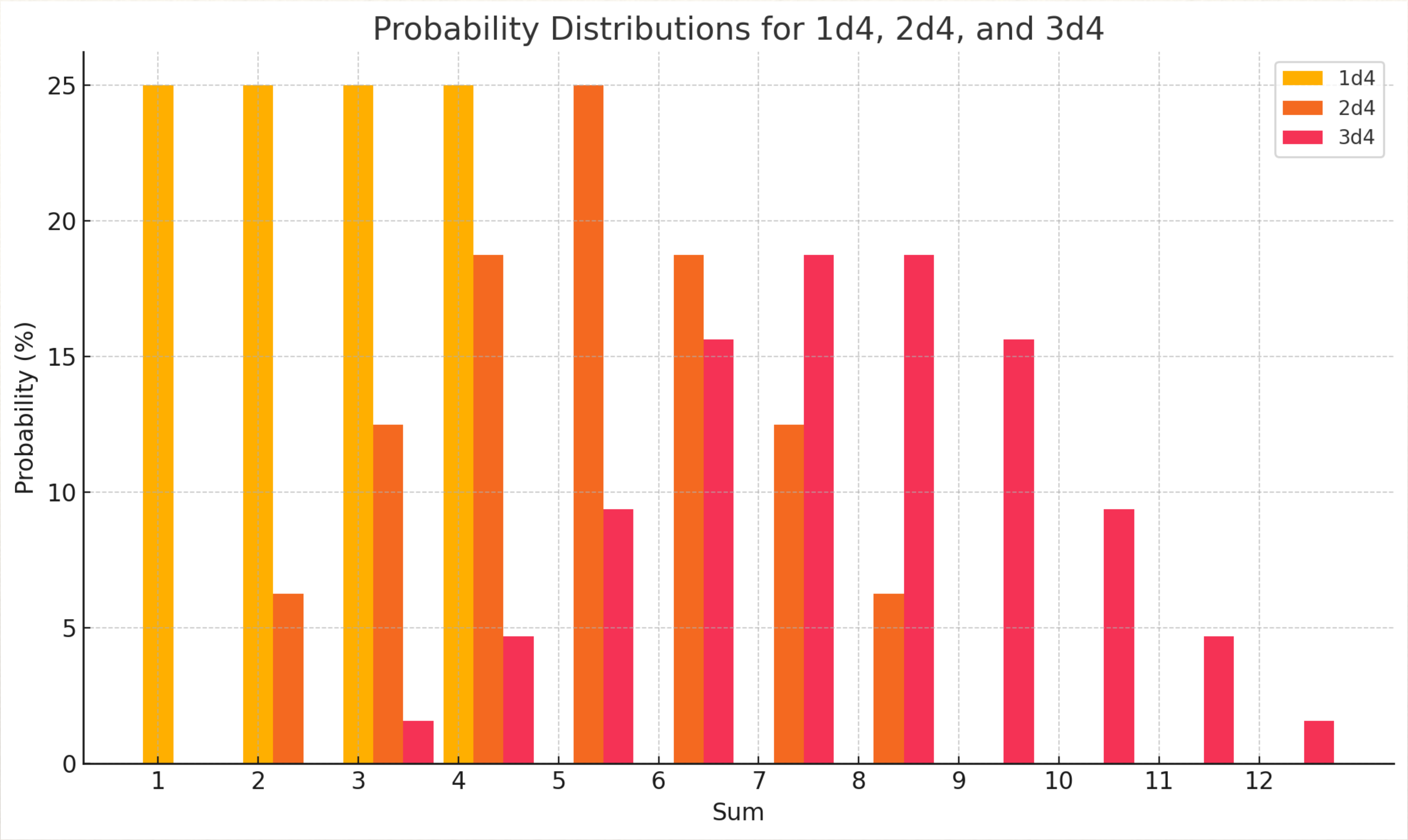
D4 Distributions

- If a formula did exist, we could easily calculate the probability of an outcome given the number of d4's that were used.
- For example, calculateD4Distribution(X, numberOfD4s)
  - calculateD4Distribution(1, 1) = .25
  - calcuateD4Distribution(1,2) = 0
  - calcuateD4Distribution(2,2) = 6.25

Please note the X axis is slightly misaligned

The formula would return the values to the right

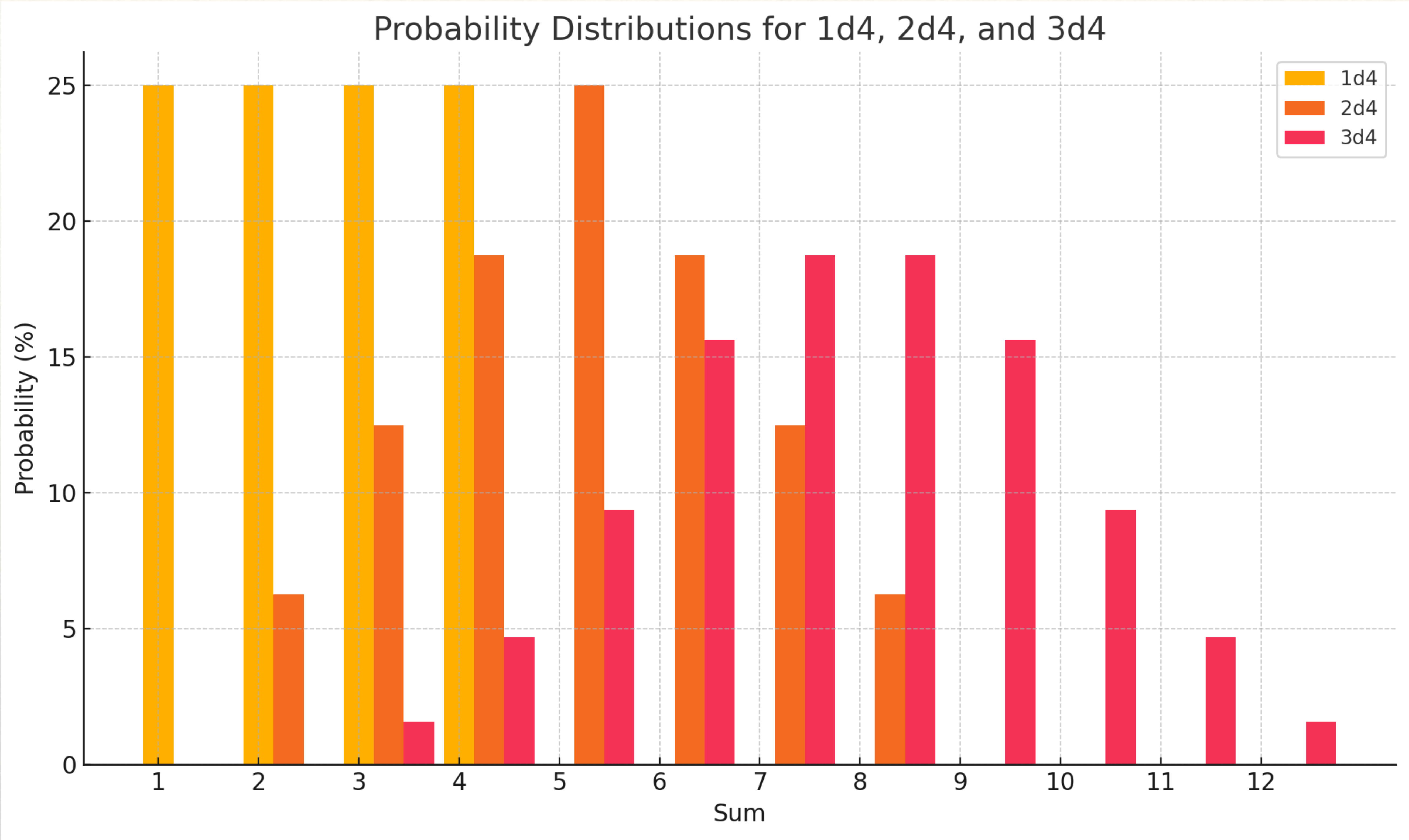
- calculateD4Distribution(1, 1) = .25
- calcuateD4Distribution(1,2) = 0
- calcuateD4Distribution(2,2) = 6.25



Please note the X axis is slightly misaligned

The formula would return the values to the right

- calculateD4Distribution(1, 1) = .25
- calcuateD4Distribution(1,2) = 0
- calcuateD4Distribution(2,2) = 6.25



Please note the X axis is slightly misaligned

No Parametric distribution like this exists!!
We could however write a program that statically returns a value based on a static mapping

Gaussian Distribution

- No formula exists to determine the parametric distribution based on the number of d4's, however parametric distributions do exist, and one of those is the Gaussian distribution.
- Gaussian distribution is determined by the mean (average sample) and standard deviation (how spread out the values are)

Please note the X axis is slightly misaligned

Gaussian Distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

x is the variable
σ is the standard deviation for the distribution
μ is the mean for the distribution
π is a pre-defined constant
e is a pre-defined constant

Please note the X axis is slightly misaligned

Gaussian Distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Can relatively easily convert the above into a function that programmatically determines the probability of x occurring given a gaussian distribution defined by 2 parameters: μ, σ

Please note the X axis is slightly misaligned

Gaussian Distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

calculateDistribution(x, average, standardDeviation)

Please note the X axis is slightly misaligned

Gaussian Distribution

What if we have more than one feature?

Pack all of the features into a 'feature vector'

Please note the X axis is slightly misaligned

Gaussian Distribution

What if we have more than one feature?

Pack all of the features into a 'feature vector'

For example:
Height of an athlete

Just One feature!!

More than one feature would be:
Height, Weight, Shoe size

Three Features!!

Please note the X axis is slightly misaligned

Gaussian Distribution

What if we have more than one feature?

Pack all of the features into a 'feature vector'

Data Representation

For example:
Height of an athlete

[70]

More than one feature would be:
Height, Weight, Shoe size

[70,200,13]

Please note the X axis is slightly misaligned

Multivariate Gaussian Distribution

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^k |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mu)\right)$$

Please note the X axis is slightly misaligned

```python
import numpy as np

def multivariate_gaussian(x, mu, covarianceMatrix):
    k = len(mu)
    covarianceMatrix_det = np.linalg.det(covarianceMatrix)
    covarianceMatrix_inv = np.linalg.inv(covarianceMatrix)
    norm_factor = 1.0 / np.sqrt((2 * np.pi) ** k * covarianceMatrix_det)
    diff = x - mu
    exponent = -0.5 * np.dot(np.dot(diff.T, covarianceMatrix_inv), diff)
    return norm_factor * np.exp(exponent)
```

Parameters:
x (np.array): The k-dimensional data point (column vector).
mu (np.array): The mean vector (column vector).
covarianceMatrix (np.array): The covariance matrix (k x k matrix).

```python
import numpy as np

def multivariate_gaussian(x, mu, covarianceMatrix):
    k = len(mu)
    covarianceMatrix_det = np.linalg.det(covarianceMatrix)
    covarianceMatrix_inv = np.linalg.inv(covarianceMatrix)
    norm_factor = 1.0 / np.sqrt((2 * np.pi) ** k * covarianceMatrix_det)
    diff = x - mu
    exponent = -0.5 * np.dot(np.dot(diff.T, covarianceMatrix_inv), diff)
    return norm_factor * np.exp(exponent)
```

```python
# Example usage:
x = np.array([1.0, 2.0])
mu = np.array([0.0, 0.0])
covMatrix = np.array([[1.0, 0.5], [0.5, 1.0]])


# Calculate the multivariate Gaussian value
result = multivariate_gaussian(x, mu, covMatrix)
print(result)
```

End Slide