Probability

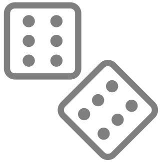
Holberton

Why Probability in Machine Learning

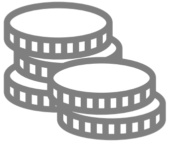
Why Probability for Machine Learning

- Probability = likelihood of an event occuring

- Framework to **reason** about uncertainties
  - Model and analyze complex systems



- Enhance predictions of ML models
  - Model complex systems in a principled manner

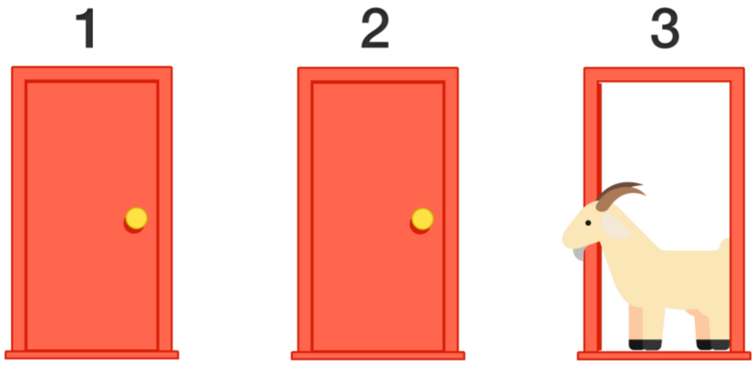


- Make informed decisions about data
  - Estimate likelihood of various outcomes

Fundamentals of Probability

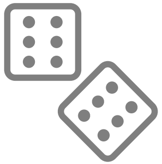
Fun facts

- Odds of winning Powerball jackpot: 1 in 292 million
  - Get struck by a lightning (1 in 15000); attacked by a shark (1 in 11.5 million)
- Odds of winning a single Blackjack game is about 42%
  - If playing with an optimal strategy
- Famous problems: the Monty Hall
  - Switching doors doubles the chance of success



Fundamentals of probability

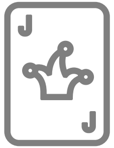
- Axioms of probability
  - $0 \leq P(w) \leq 1$
  - $\Sigma_w (w) = 1$



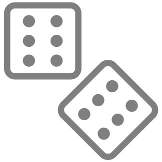
- **Random variable**: randomly takes a value from a sample space
- **Probability distribution**: describe likelihood of each possible value

Fundamentals of probability

- Probability of a single event
  - $probability = \frac{\#\ favorable\ outcomes}{\#\ all\ possible\ equally{-}like\ outcomes}$



- Example: you throw two dice? What is the probability that the sum of the two dice will be 7?
  - $\frac{6}{36}$ (favorable events (1,6; 6,1; 2,5; 5,2; 3,4; 4,3) out of 36 possible outcomes)



Fundamentals of probability

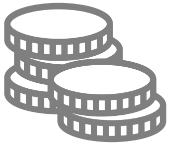
- Probability of two (or more) independent events

- A and B are independent events if probability of B occurring is not related to A

- **Probability of A and B**

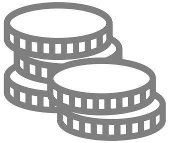
  - $P\ (A\ and\ B) = P(A) \cdot P(B)$

  - Example: flip a coin twice



Fundamentals of probability

- **Probability of A or B**
  - $P\ (A\ or\ B) = P(A) + P(B)\ - P\ (A\ and\ B)$
- Three possibilities
  - A occurs, but not B
  - B occurs, but not A
  - Both A and B occur

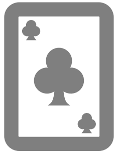


- Example: if you flip a coin twice, what are the odds of getting a head either the first time, or the second? (or both)

Fundamentals of probability

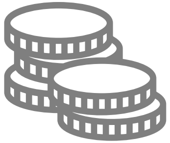
- **Conditional probabilities** $P\ (A \mid B)$
  - The probability of an event, given that another event has occurred
  - Example: $P(ace\ on\ the\ second\ draw \mid ace\ of\ the\ first\ draw)$
  - $P\ (A \mid B)\ = \dfrac{P\ (A\ and\ B)}{P\ (B)}$

- Example: if you draw a card from a deck what is the probability that you get a Ace of spades and a black card?



Fundamentals of probability

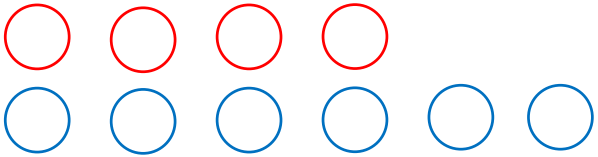
- **Birthday problem**
  - If there are 25 people in the room, what is the probability that at least two of them share the same birthday?
  - Not $\frac{25}{365}$!!!
- **Gambler‘s fallacy**
  - A fair coin is flipped 5 times and comes head each time. What is the probability it will come head on the sixth time?
  - Yes $\frac{1}{2}$!!!



Fundamentals of probability

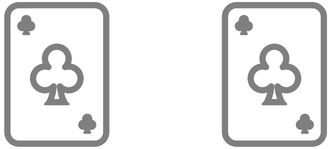
- **Disjoint probabilities**
  - Two events A and B are disjoint if they cannot occur together
  - $P\ (A\ or\ B) = P(A) + P(B)$

- **General addition rule = OR rule**
  - $P\ (A\ or\ B) = P(A) + P(B)\ - P\ (A\ and\ B)$
  - Example: suppose you have 4 red balls, and 6 blue balls. What is the probability of drawing a red or blue ball?



Fundamentals of probability

- **General multiplication rule (intersection rule)**
  - Intersection of two events (independent or not)
  - $P\ (A\ and\ B) = P(A) + P(B \mid A)$
  - Probability of two events occurring together
  - Example: suppose we have a deck of cards? What is the probability of drawing two spades?



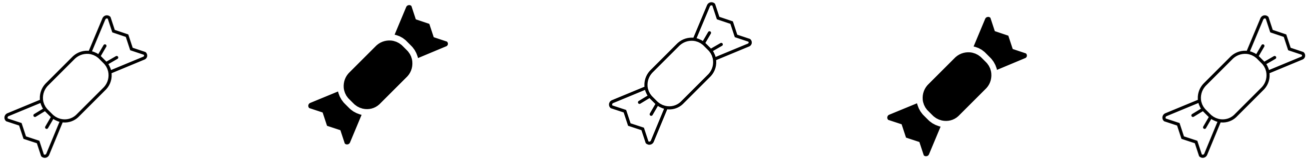
Permutations and Combinations

Permutations and combinations

- **Permutations**
  - Arrangements of objects in a specific order
  - The order matters
  - In a permutation, each object appears only once
  - $\# \, permutations = n!$ for n objects
  - Example: in how many ways can 5 candies be distributed?

Permutations and combinations

- **Permutations**
  - Suppose there are 5 candy and you must pick only 2. How many possible ways are there?



  - $P_{n,r} = \dfrac{n!}{(n-r)!}$
  - In other words: the number of ways r things can be selected from n things

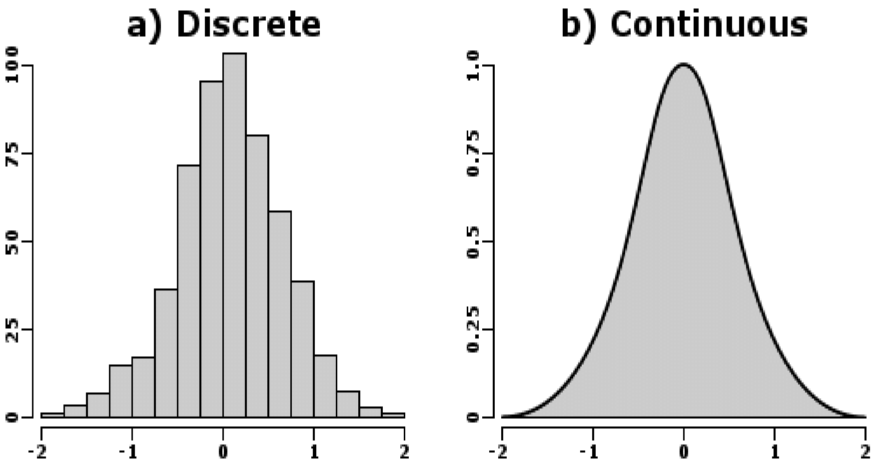
Permutations and combinations

- **Combinations**
  - Suppose you are only concerned with the final choices
  - Example: choosing red and yellow is the same as choosing yellow and red
  - $C_{n,r} = \frac{n!}{(n-r)! \cdot r!}$
  - In other words: the number of ways r things can be selected from n things if the order does not matter

Probability Distributions

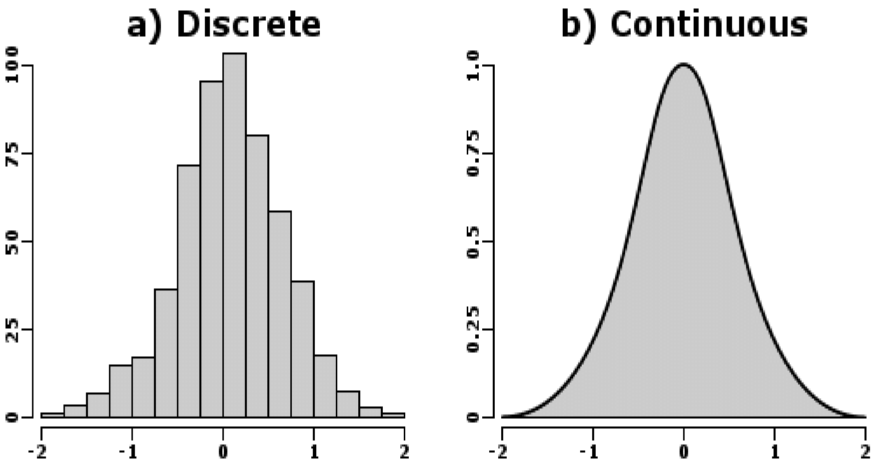
Probability distributions

- **Discrete distributions**
  - Discrete variables take only a countable number of values
  - A probability mass function (PMF) assigns probabilities to outcomes
  - Example: uniform distribution assign equal chance



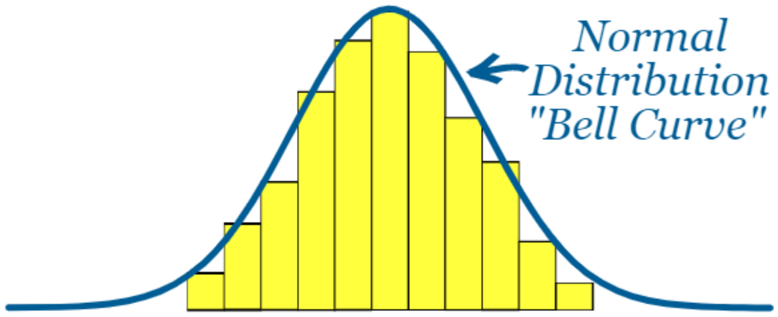
Probability distributions

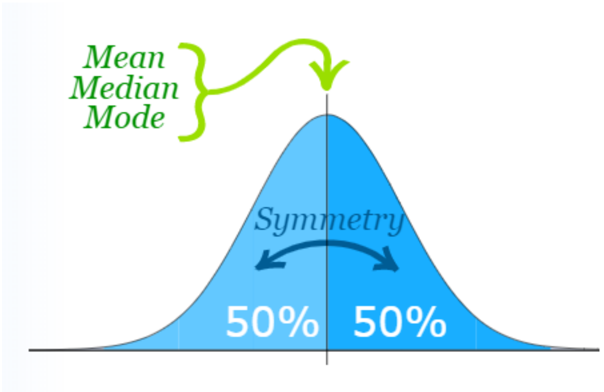
- **Continuous distributions**
  - Continuous variables take *any value* within a range
  - A probability density function (PDF) assigns probabilities to outcomes
  - Applications: hypotheses testing, regression, time-series



Normal distribution

- Data is spread around a central value with no bias





- mean = median = mode
- symmetrical

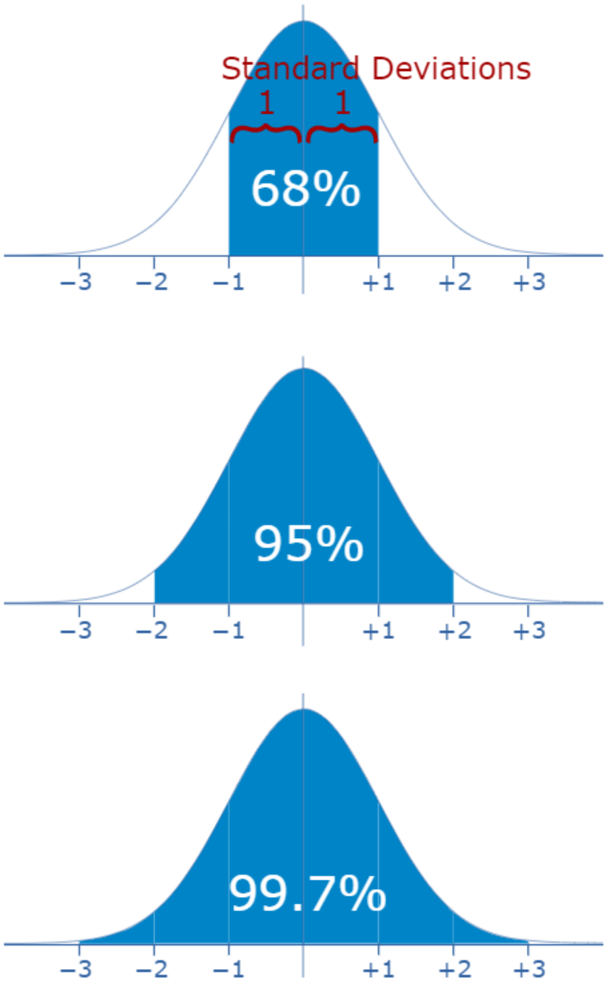
Normal distribution

- **Standard deviation**

- A measure of how spread out the numbers are

- $\sigma = \sqrt{\frac{\Sigma(x - mean)^2}{n}}$

- Standard deviation measures teh variability of data



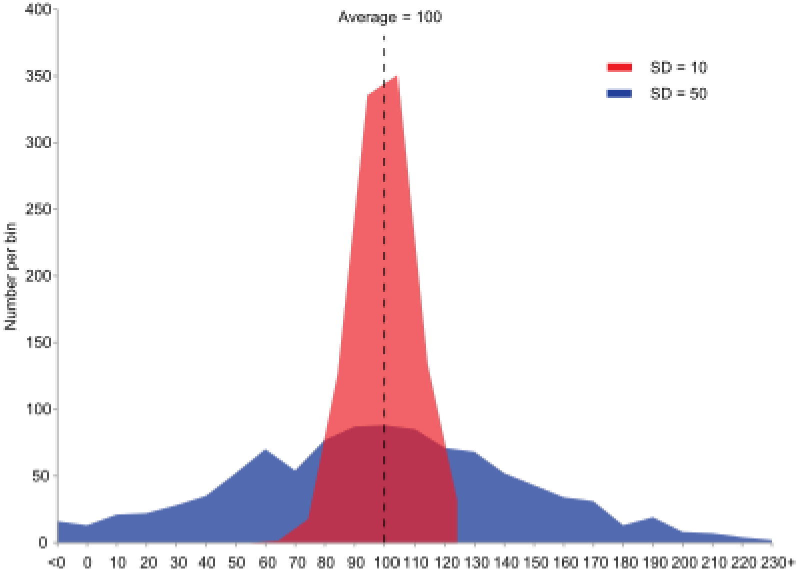
Normal distribution

- **Variance**

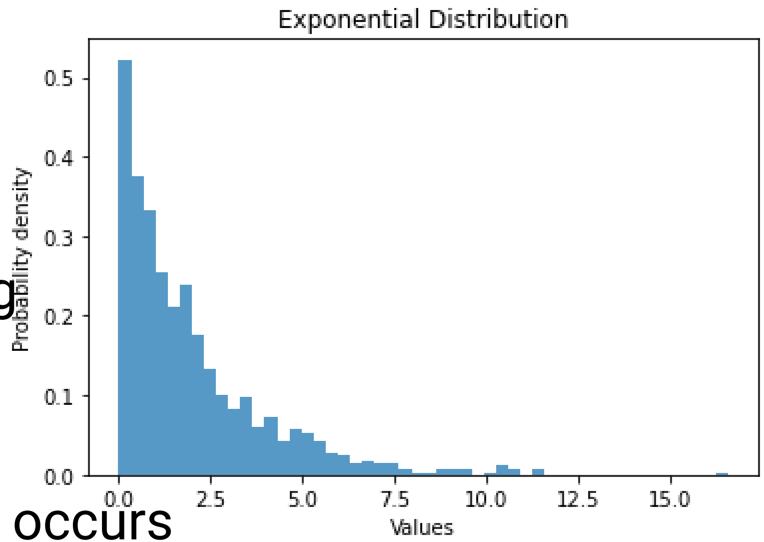
- Measures how far data points are from the mean

- $\sigma^2 = \frac{\Sigma(x - mean)^2}{n}$

- Standard deviation and variance differ in units of measurement or calculation



Exponential distribution



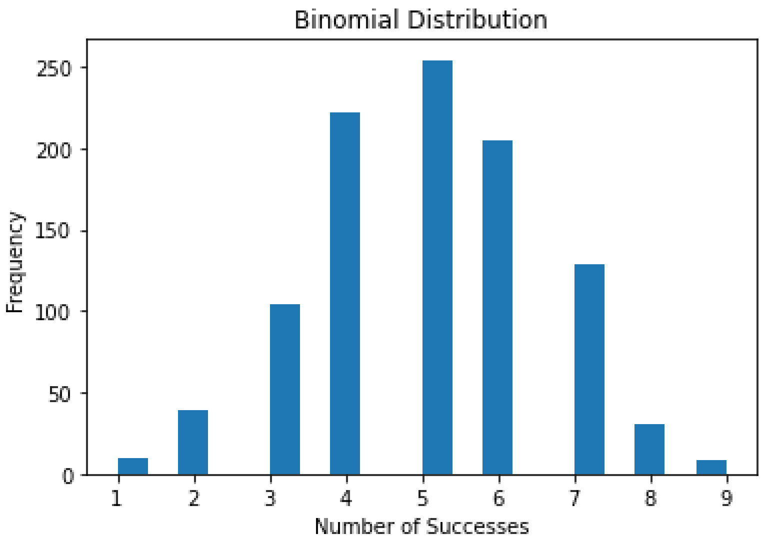
- Models the time until an event occurs where the probability of event occurring is constant over time
- Parameter λ denotes the rate the event occurs
- $f(x) = \lambda \cdot e^{-\lambda \cdot x}$ for $x \geq 0$; or 0 for $x < 0$
- Memoryless distribution
- Applications: survival analysis & reliability engineering

Binomial distribution

- Models the numbers of successes in a fixed number of Bernoulli tries

- $P(x) = C_{x,n}\, p^{x} \cdot (1-p)^{n-x}$

- Applications: probability of binary outcomes (spam or not; will customer purchase or not)



“The only certainty
is that nothing
is certain”

Any questions?