Outline today

- Why probabilistic modeling?
- Basics of Probability
  - Random variable
  - Probability distribution
  - Dependent and independent variables
  - Correlation vs. Causation
- Bayes rule
- Probabilistic Graphical Models
- Probabilistic Robotics

Random Variable

- X is a random variable if it can take different values with different probabilities and can be observed with an experiment

Random Variable

- X is a random variable if it can take different values with different probabilities and can be observed with an experiment
- X can be continuous or discrete
  - Outcome of a dice → Discrete random variable
  - Height of adult population of a country → Continuous random variable

Random Variable

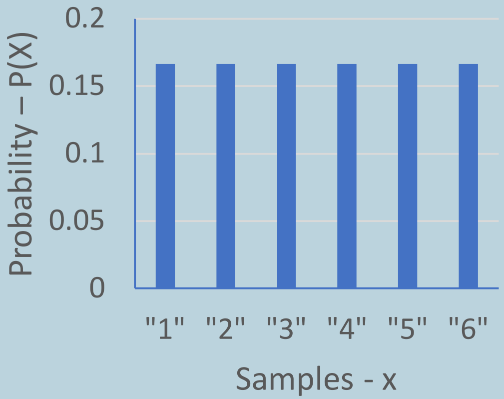
- X is a random variable if it can take different values with different probabilities and can be observed with an experiment
- X can be continuous or discrete
  - Outcome of a dice → Discrete random variable
  - Height of adult population of a country → Continuous random variable
- Sample space: all possible values that a random variable can take
  - X=outcome of a dice → S={1,2,3,4,5,6} → Discrete
  - X=Height of adult population → S=[1 feet, 9 feet] → Continuous

Random Variable

- X is a random variable if it can take different values with different probabilities and can be observed with an experiment
- X can be continuous or discrete
  - Outcome of a dice → Discrete random variable
  - Height of adult population of a country → Continuous random variable
- Sample space: all possible values that a random variable can take
  - X=outcome of a dice → S={1,2,3,4,5,6} → Discrete
  - X=Height of adult population → S=[1 feet, 9 feet] → Continuous
- P(X=x): Probability of an event (or probability of X taking the value x)
  - $P(X=x) \leq 1$
  - $\sum_x P(X = x) = 1$

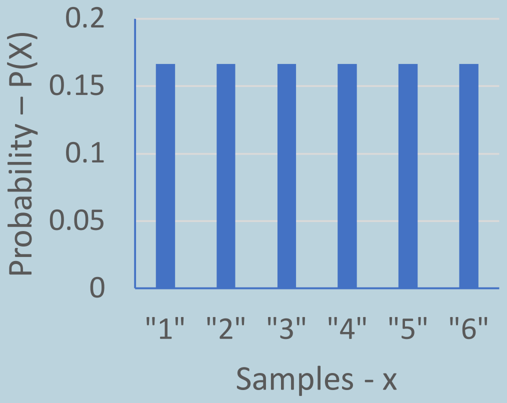
Probability distribution

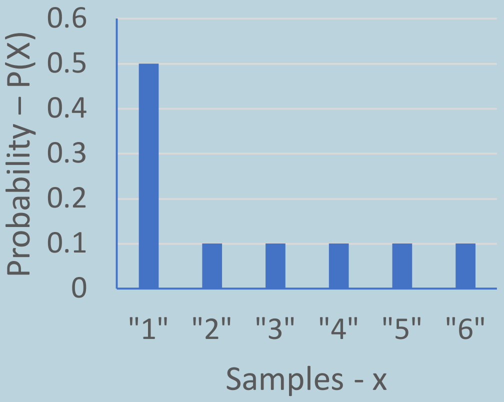
- Function P(X=x) (or directly P(x))
  - When we plot it:
    - x axis: the sample space
    - y axis: the probability of each sample, P(x)



Probability distribution

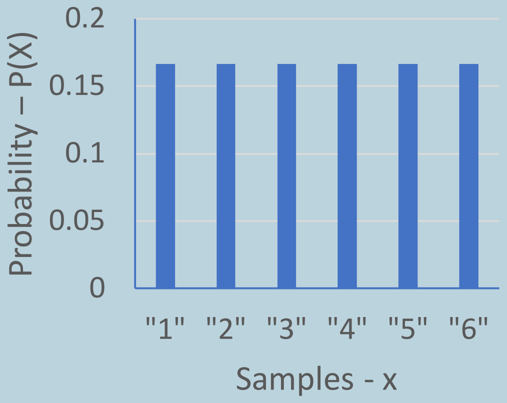
- Function P(X=x) (or directly P(x))
  - When we plot it:
    - x axis: the sample space
    - y axis: the probability of each sample, P(x)

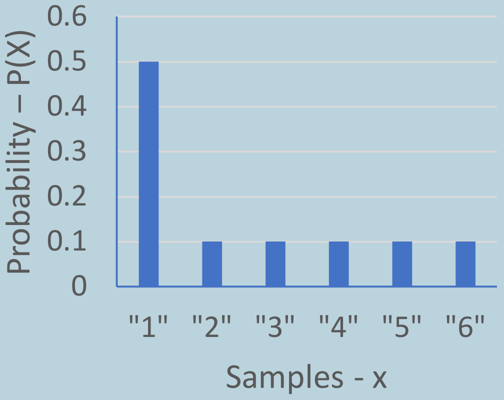


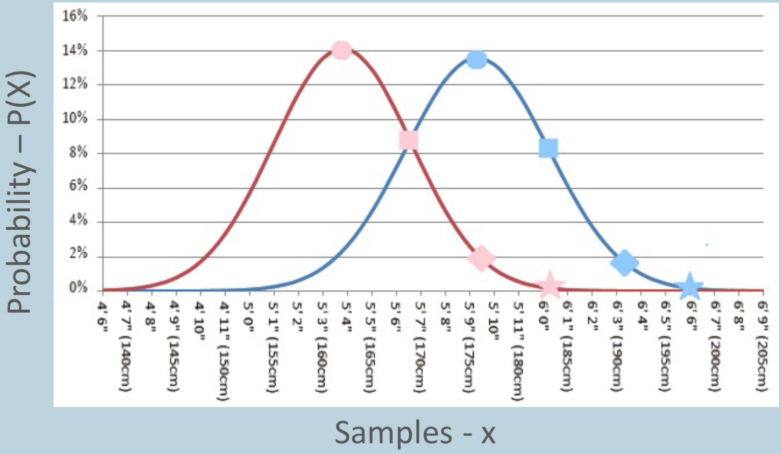


Probability distribution

- Function P(X=x) (or directly P(x))
  - When we plot it:
    - x axis: the sample space
    - y axis: the probability of each sample, P(x)





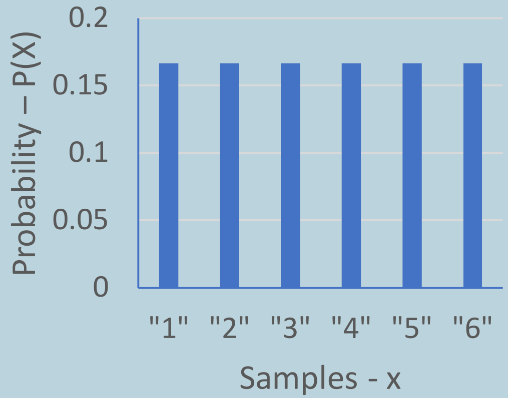


Mean of a Random Variable

- "Mean" is also called "Expected Value"
- The mean value depends on:
  - The numeric value of each sample
  - The probability of each sample
- Gives you the value that is most likely to happen if we repeat the experiment
- Formally: $E(X) = \sum_x x \cdot P(x)$

Mean of a Random Variable

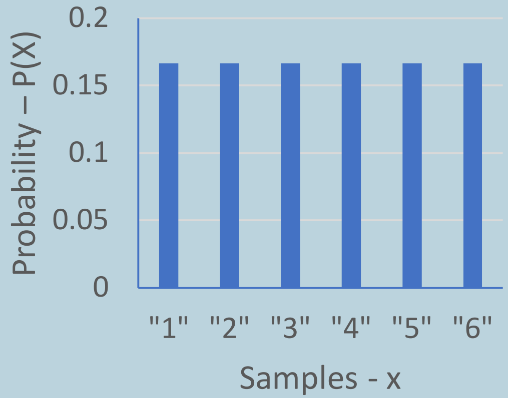
- $E(X) = \sum_x x \cdot P(x)$



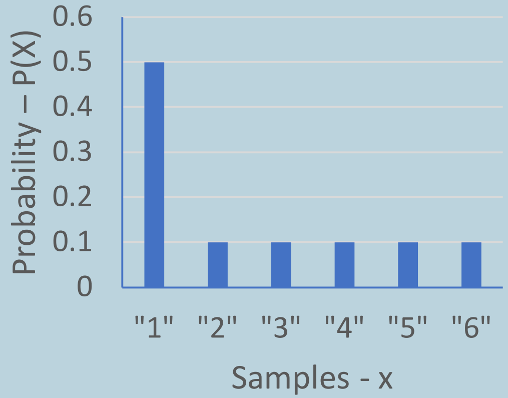
$$E(x) = \sum_x x \cdot P(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

Mean of a Random Variable

- $E(X) = \sum_x x \cdot P(x)$

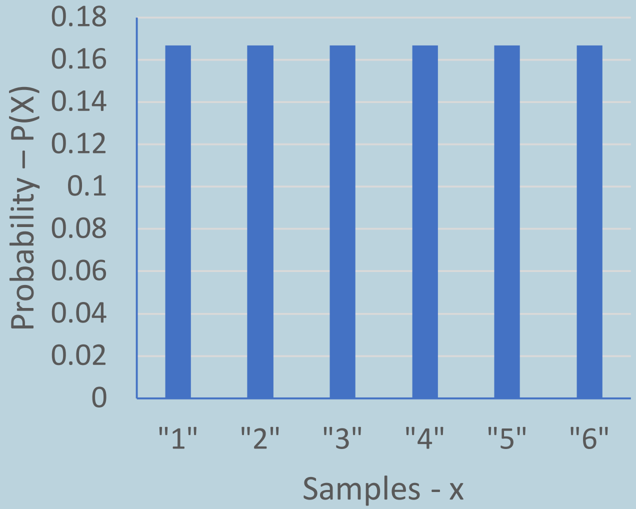


$$E(x) = \sum_x x \cdot P(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$



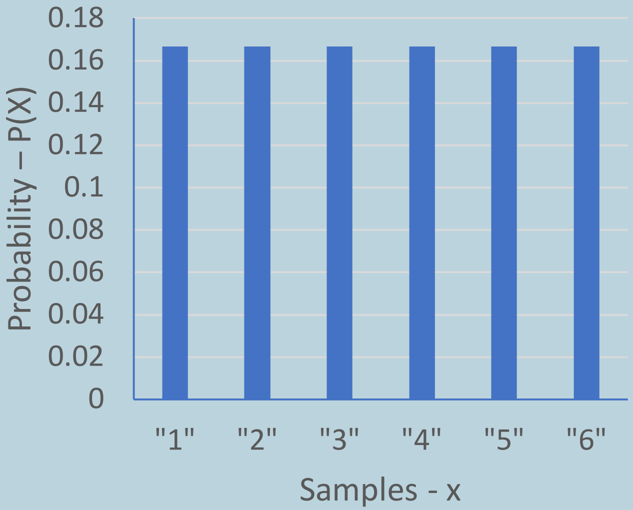
$$E(x) = \sum_x x \cdot P(x) = 1 \cdot 0.5 + 2 \cdot 0.1 + 3 \cdot 0.1 + 4 \cdot 0.1 + 5 \cdot 0.1 + 6 \cdot 0.1 = 2.5$$

Informed Decisions based on Probabilistic Modeling



Instapoll

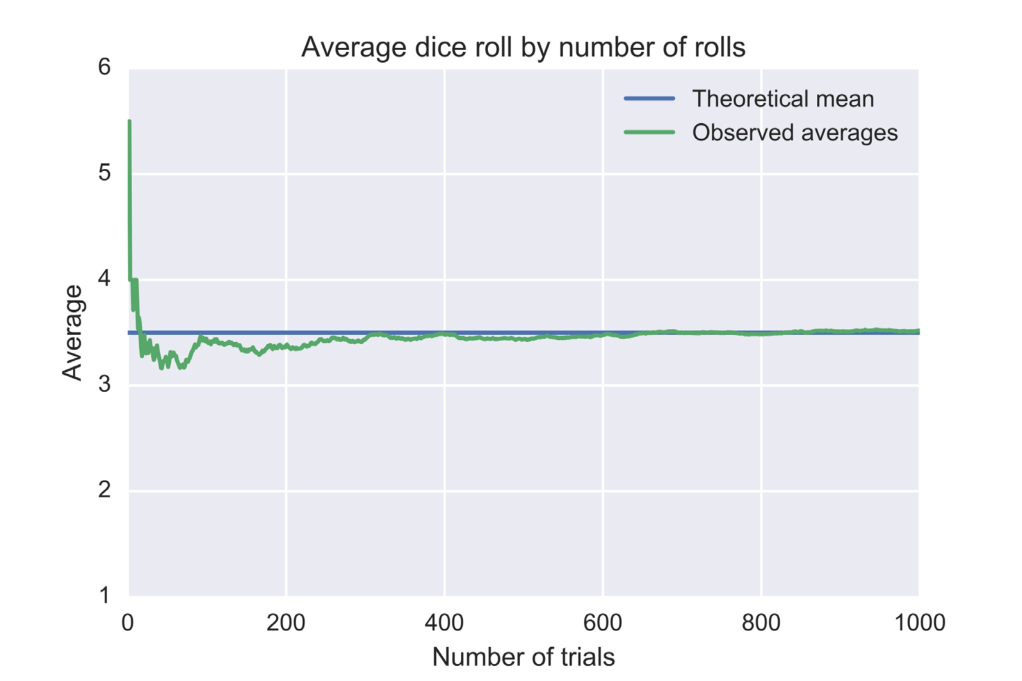
Informed decisions based on probabilistic modeling



Mean: $3.5

Law of Large Numbers

- The average of the outcome of an experiment repeated many times tends to the expected value (mean)
- It gets closer, the more we repeat the experiment



$$\lim_{n\to\infty}\sum_{i=1}^{n}\frac{X_i}{n} = E(X)$$

Dependent and Independent Variables

- Let's assume X and Y are two random variables
  - X → outcome of a coin toss, Y → outcome of a dice roll
  - X → develop lung cancer, Y → be a smoker
- P(X,Y) is called the joint probability
  - P(X="tails", Y="6")
  - P(X="develop lung cancer", Y="smoker")

Dependent and Independent Variables

- Let's assume X and Y are two random variables
  - X → outcome of a coin toss, Y → outcome of a dice toss
  - X → develop lung cancer, Y → be a smoker
- P(X,Y) is called the joint probability
  - P(X="tails", Y="6")
  - P(X="develop lung cancer", Y="smoker")
- P(X|Y) is called a conditional probability, probability of X conditioned on knowing Y
  - P(X|Y="6") → Probability of outcome of a coin toss given that the dice toss was "6"
  - P(X|Y="smoker) → Probability of developing lung cancer given that is a smoker

Dependent and Independent Variables

- X and Y are said to be independent variables if the distribution of X is not influenced by the value taken by Y and vice versa. In that case:
  - the probability of a joint event is the product of the probability of one event and the probability of the other
    - P(X="tails", Y="6") = P(X="tails") · P(Y="6")
  - the conditional probability of one knowing the other is the same as the original probability
    - P(X | Y="6") = P(X)

Dependent and Independent Variables

- X and Y are said to be independent variables if the distribution of X is not influenced by the value taken by Y and vice versa. In that case:
  - the probability of a joint event is the product of the probability of one event and the probability of the other
    - P(X="tails", Y="6") = P(X="tails") ▪ P(Y="6")
  - the conditional probability of one knowing the other is the same as the original probability
    - P(X | Y="6") = P(X)
- X and Y are said to be dependent variables if knowing the value of one of the events affects the probability distribution over the other event. In that case:
    - P(X | Y="smoker") ≠ P(X)

Correlation vs. Causality

- Correlation:
  - Measurement of “how much two random variables are dependent on each other”
  - Correlation = 0 → Variables are independent
  - Correlation $\neq$ 0 → Variables are dependent

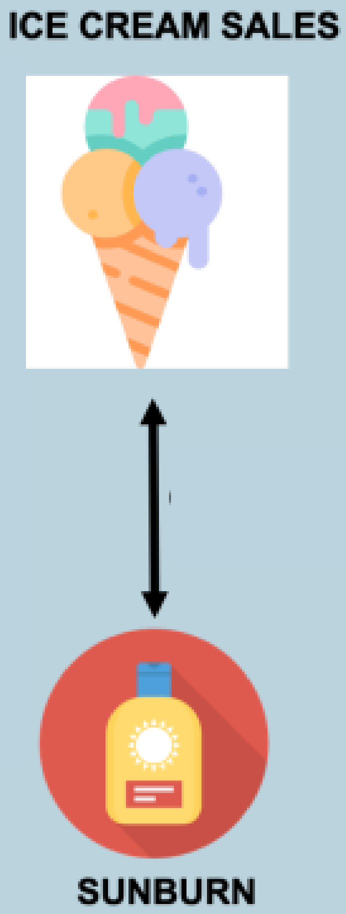
- <u>Correlation is not causation</u>
  - We can mathematically proof that two variables are correlated but that doesn’t mean one is causing the other!
  - Careful! This is sometimes the door to “biases” and “prejudices”

Correlation vs. Causality

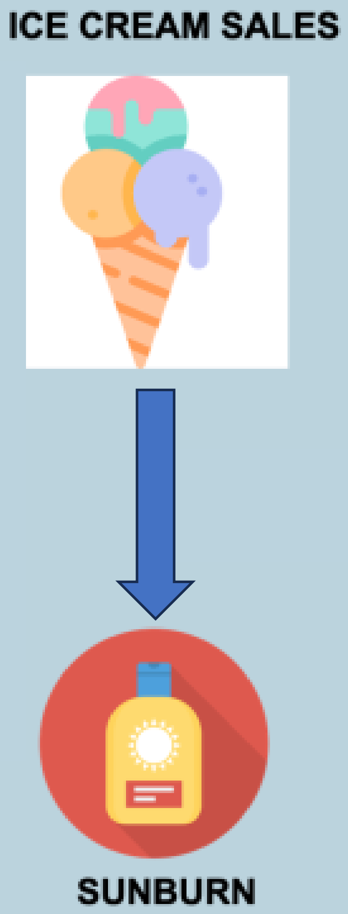
ICE CREAM SALES

SUNBURN

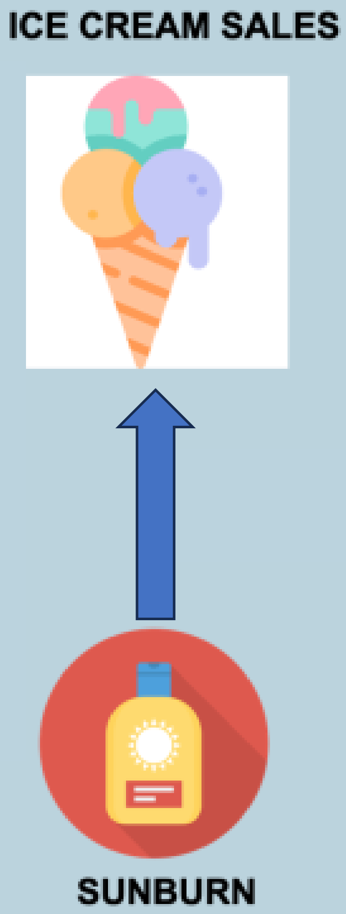
Correlation vs. Causality



Correlation vs. Causality



Correlation vs. Causality



Correlation vs. Causality