METIS

Lesson 4:
Random Variables

Introduction

METIS

Lecture Overview:

Goals of the lecture:

1. Differentiate between Discrete and Continuous Random Variables
2. Understand the Probability Mass and Density Function, Cumulative Distribution Function

Random Variables

METIS

Discrete Random Variables

Definition:

Discrete Random Variables are a specific type of Random Variable that can only take on finite number of values. All of the examples we have seen so far are cases of Discrete Random Variables

If we define the set of possible values for a Random Variable and find them to be a set of discrete points (i.e. $x_0, x_1, \dots, x_n$), this is an example of a Discrete Random Variable

Example 1:

1. Die
2. Deck of cards
3. Age

Continuous Random Variables

Definition:

Continuous Random Variables are a specific type of Random Variable that can take on infinitely many values over a continuous range**.** Unlike Discrete Random Variables, the probability of any given value for a Continuous Random Variables is 0 $\mathbb{P}(X = a) = 0$

Example 2:

1. Height
2. Temperature
3. Distance

Probability Mass Functions

Definition:

PMF is a function for discrete random variables that indicates that the probability of an event is exactly equal to some value

$$\mathbb{P}_X(x_i) = \mathbb{P}\{X = x_i\}$$

Cumulative Distribution Function

Definition

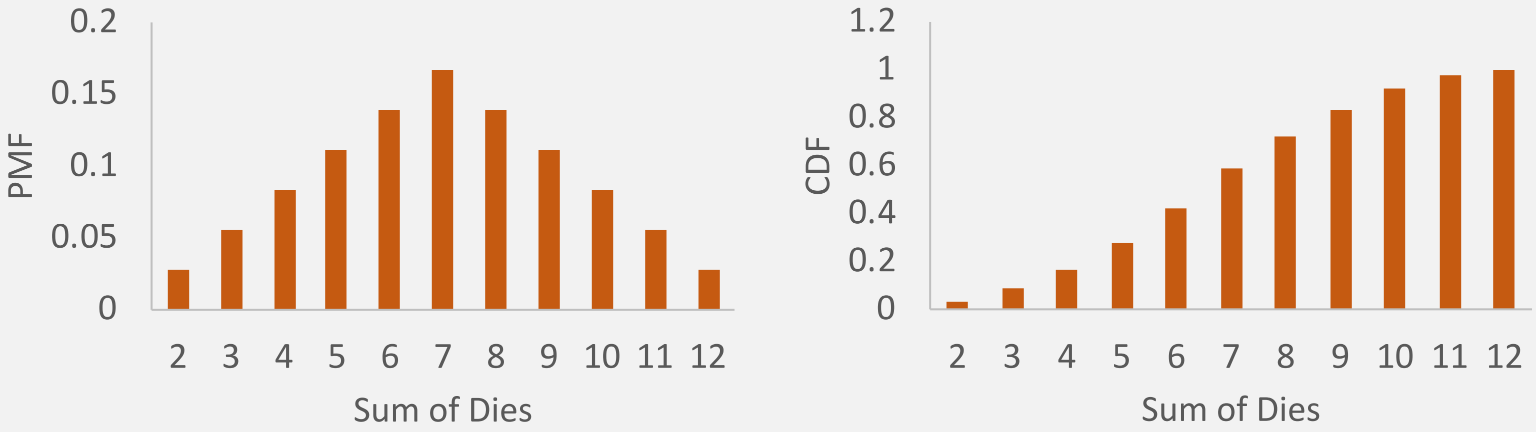
We can contrast the Probability Mass Function with a Cumulative Distribution Function (or just "Distribution Function"). The Cumulative Distribution Function is developed by summing probabilities in the Probability Mass Function in a cumulative fashion. In general, the formula for the CDF of a Random Variable X is:

$$F_x(x) = \mathbb{P}(X \le x) = \sum_{i=0}^{j} \mathbb{P}(X = x_i)$$

PMF and CMF

Example 3:

PMF and CDF of the sum of two dies

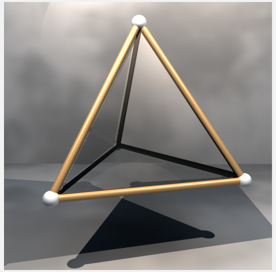


Problem 1

Problem 1:

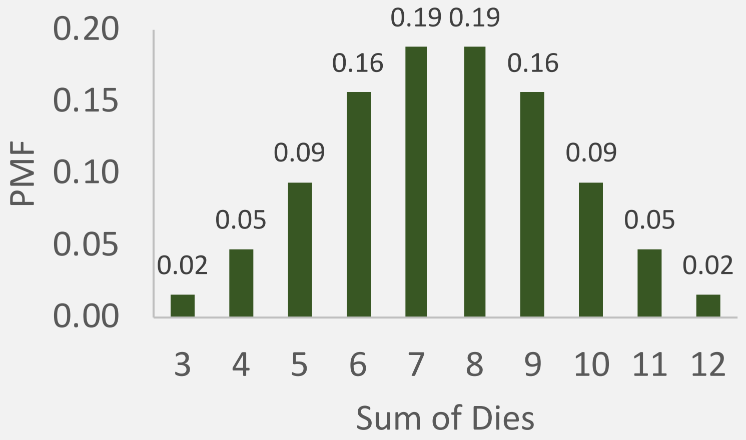
Assume you have a 4-sided die (S={1,2,3,4}), that is thrown three times.

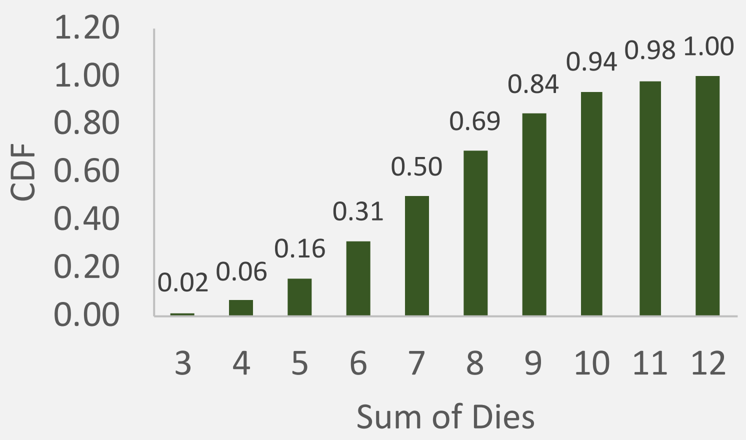
1. Draw the PMF and CDF for the sum of the 3 dies. What is the most likely outcome?
2. What is the probability for the sum of the three dies to be a 10 or higher?



Problem 1

Solution 1:





1. Most likely outcome is a 7 or 8
2. $\mathbb{P}(10,11,12)$ = 1-CDF(9) = 1-0.84 = 0.16

Probability Density Functions

Definition (1):

PDF is a function for continuous random variables, whose value at any given sample range can be interpreted as a relative likelihood that the value of the random value would equal that sample range. The (informal) formulation of the **pdf** is:

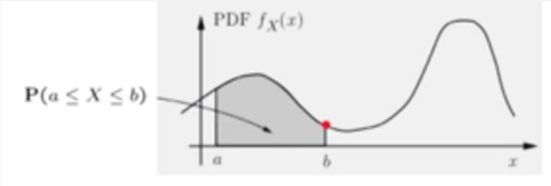
$$f_X(x)dx = \mathbb{P}(x \le X \le x + dx)$$

In the continuous case we measure **Density** (probability/unit of length) and NOT probability directly

Definition (2):

To calculate the probability of any range of x values, the area under the curve must be calculated as follows:

$$P(a \le X \le b) = \int_a^b f_X(x)dx$$



Cumulative Distribution Functions

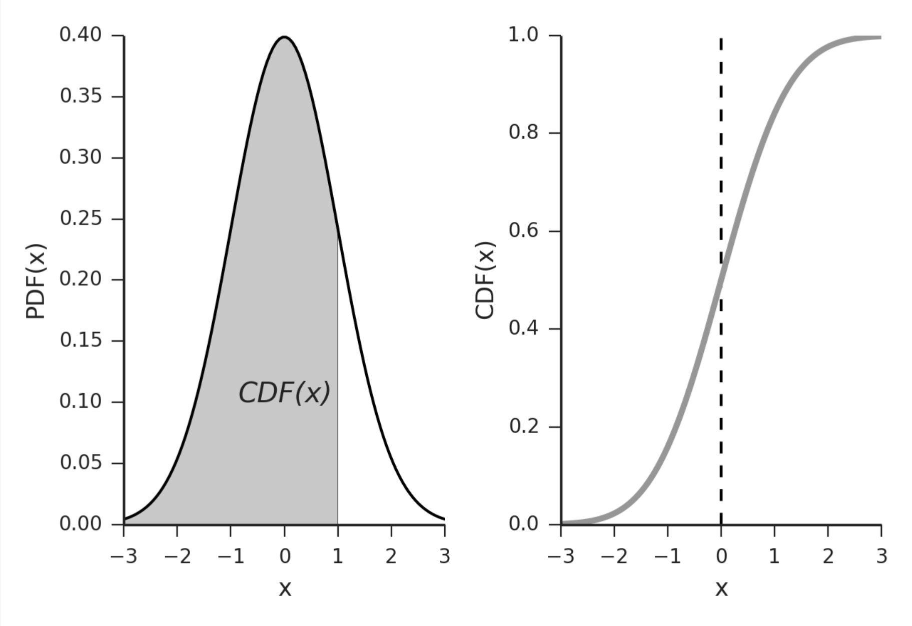
Definition

For Continuous Random Variables, we can no longer simply add probabilities in a cumulative fashion to develop the Cumulative Distribution Functions. Instead, we have to calculate the area under the curve from negative infinity to our desired point. Thus, for a Continuous Random Variable X, we have:

$$F_x(x_j) = \int_{-\infty}^{x} f(x)dx$$

Continuous RV and Distribution Functions

Example 17:



Problem 2

Problem 2:

If we have a Continuous Random Variable X with the following Probability Density Function:

$$f_X(x) = \begin{cases} 0 \text{ if } x \le 0, \\ 0.2 \; if \; 0 < x \le 5, \\ 0 \; if \; x > 5 \end{cases}$$

Then what is the probability of randomly selecting a value less than 4?

Problem 2

Solution 2:

$$F_x(4) = \int_{-\infty}^{4} f_X(x)dx = \int_{0}^{4} 0.2dx = [0.2x]_0^4 = 0.2(4) = 0.8$$

Common Discrete Distributions

METIS

Discrete Distributions

Common Discrete Distributions:

1. Bernoulli Distribution
2. Binomial Distribution
3. Poisson Distribution

Common Continuous Distributions

METIS

Continuous Distributions

Common Continuous Distributions:

1. Exponential Distribution
2. Gamma Distribution
3. Chi-Squared Distribution
4. Normal (Gaussian) Distribution
5. Log-Normal Distribution

More Problems

METIS

Problem 3

Problem 3:

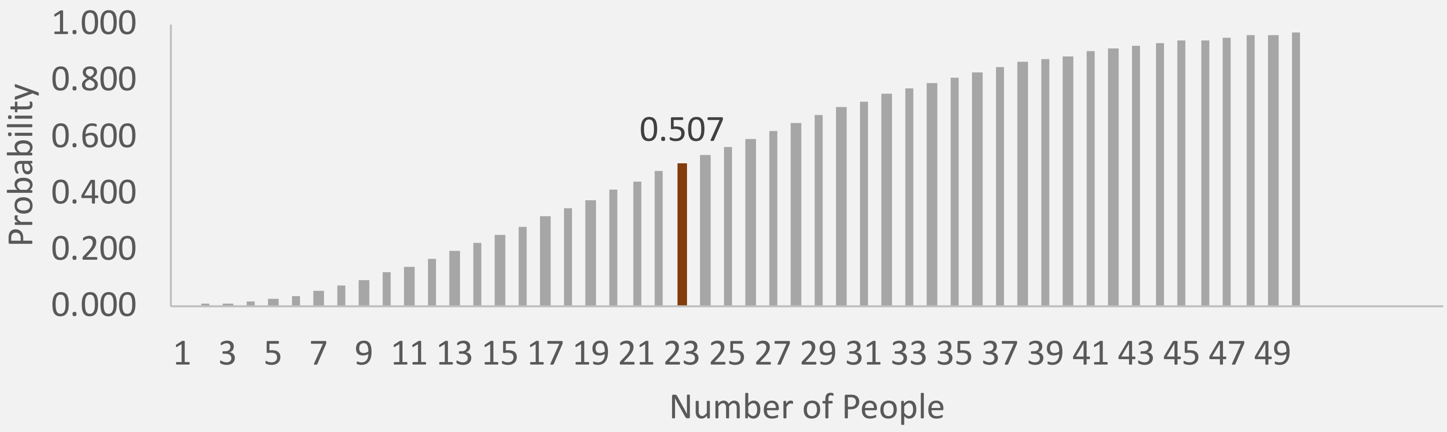
In a classroom of 23 people, what is the probability that at least two people have the same birthday? Assume 365 days in the calendar.

Problem 3

Solution 3:

$$\mathbb{P}(\text{Different Birthdays}) = \frac{365}{365}\frac{364}{365}\frac{363}{365}\cdots\frac{343}{365} = 0.4927$$

$$\mathbb{P}(\text{at least 2 Birthday}) = 1 - \mathbb{P}(\text{Different Birthdays}) = 1 - 0.4927 = 0.5072$$



Problem 4

Problem 4:

A survey determined the following probabilities for a group of students:
$\mathbb{P}$(at least play piano) = 0.3
$\mathbb{P}$(at least play guitar) = 0.2
$\mathbb{P}$(play piano and guitar) = 0.1

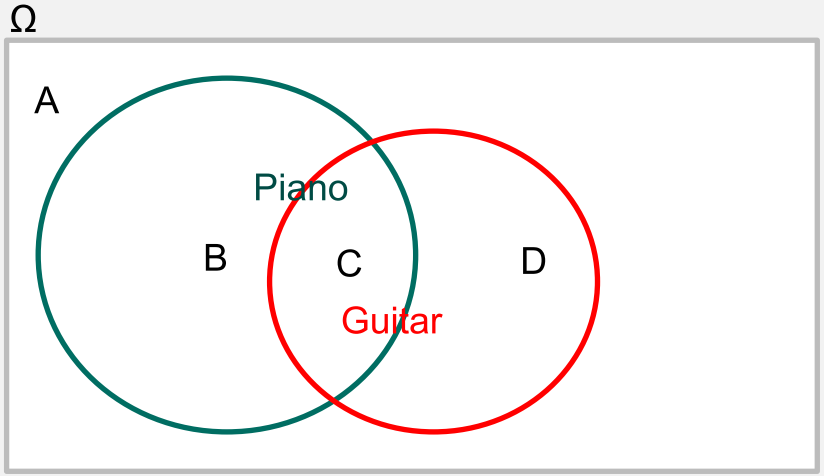
Find the probability that a student plays guitar, given that they play piano.

Problem 4

Solution 4:

$\mathbb{P}$(at least play piano) = 0.3
$\mathbb{P}$(at least play guitar) = 0.2
$\mathbb{P}$(play piano and guitar) = 0.1

$$\mathbb{P}(\mathrm{G}|\mathrm{P}) = \frac{\mathbb{P}(\mathrm{P}+\mathrm{G})}{\mathbb{P}(\mathrm{P})} = \frac{0.1}{0.3} = 0.33$$



Problem 5

Problem 5:

An ice cream shop noticed the following:
$\mathbb{P}$(only chocolate) = 0.3
$\mathbb{P}$(only vanilla) = 0.1
$\mathbb{P}$(chocolate and vanilla) = 0.4

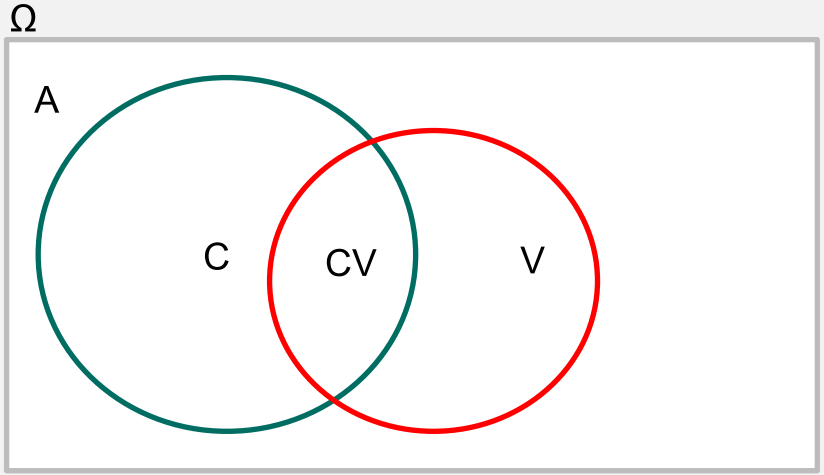
What is the probability that a random person ordered chocolate given that they ordered vanilla?

Problem 5

Problem 5:

An ice cream shop noticed the following:
$\mathbb{P}(\mathrm{C}) =$ $\mathbb{P}$(only chocolate) = 0.3
$\mathbb{P}(\mathrm{V}) =$ $\mathbb{P}$(only vanilla) = 0.1
$\mathbb{P}(\mathrm{CV}) =$ $\mathbb{P}$(chocolate and vanilla) = 0.4

$$\mathbb{P}(\mathrm{C}V|\mathrm{CV} \cup \mathrm{V}) = \frac{\mathrm{P(CV)}}{\mathrm{P(CV \cup V)}} = \frac{0.4}{0.5} = 0.8$$



QUESTIONS?