

# DS250 (DAV) Class Module

DR.GAGAN RAJ GUPTA

January 2025

## 0.1 Introduction

Notes on Plotting Histograms and Heatmaps

### 1. HISTOGRAMS:-

A histogram is a graphical representation of the distribution of data. It divides the entire range of values into intervals (called bins) and counts the number of data points that fall into each bin.

*Mathematical Definition:-*

Let  $x$  be the data points. Divide the range of data into bins, where:

The bin width is:

$$w = \frac{\max(x) - \min(x)}{k}$$

$$b_i \leq x < b_{i+1}$$

Here,  $b_i$  and  $b_{i+1}$  are the edges of the  $i$ -th bin.

*Steps to Plot a Histogram:-*

1. Choose the number of bins: Use the rule of thumb or algorithms like the Sturges' formula:

$$k = \lceil \log_2(n) + 1 \rceil$$

2. Plot the frequencies using bar heights.

*Python Code Example:-*

```
import matplotlib.pyplot as plt
import numpy as np

data = np.random.normal(0, 1, 1000)

plt.hist(data, bins=30, edgecolor='black', alpha=0.7)
plt.title('Histogram')
plt.xlabel('Value')
plt.ylabel('Frequency')
plt.show()
```

*ACTIVITY ON HISTOGRAM:-*

### **0.1.1 Part A: Word Length Histogram**

#### **Objective**

Create a histogram that represents the frequency of words based on their length in any paragraph from your preferred book.

#### **Steps**

1. Choose a paragraph from your favorite book.
2. Count the number of letters in each word.
3. Group words by their lengths (e.g., words with 1 letter, 2 letters, etc.).
4. Plot the histogram with:
  - **X-axis:** Word lengths.
  - **Y-axis:** Frequency of words of each length.

### **0.1.2 Part B: Starting Letter Histogram**

#### **Objective**

Create a histogram to analyze the frequency of words starting with each letter of the alphabet.

#### **Steps**

1. Use the same paragraph as in Part A.
2. Count how many words start with each letter (A-Z).
3. Plot the histogram with:
  - **X-axis:** Starting letters (A-Z).
  - **Y-axis:** Frequency of words starting with each letter.

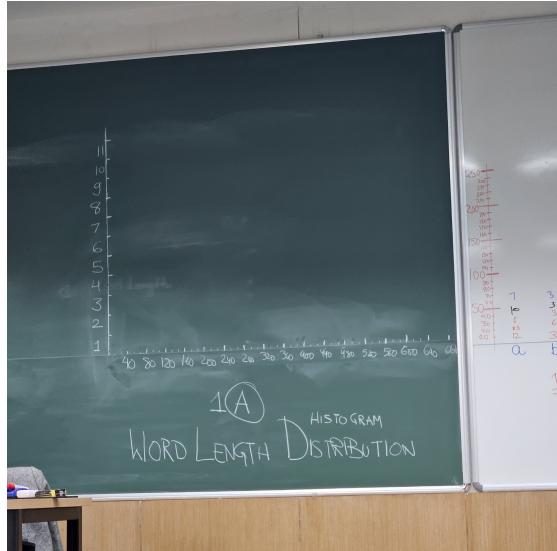


Figure 1: word length distribution

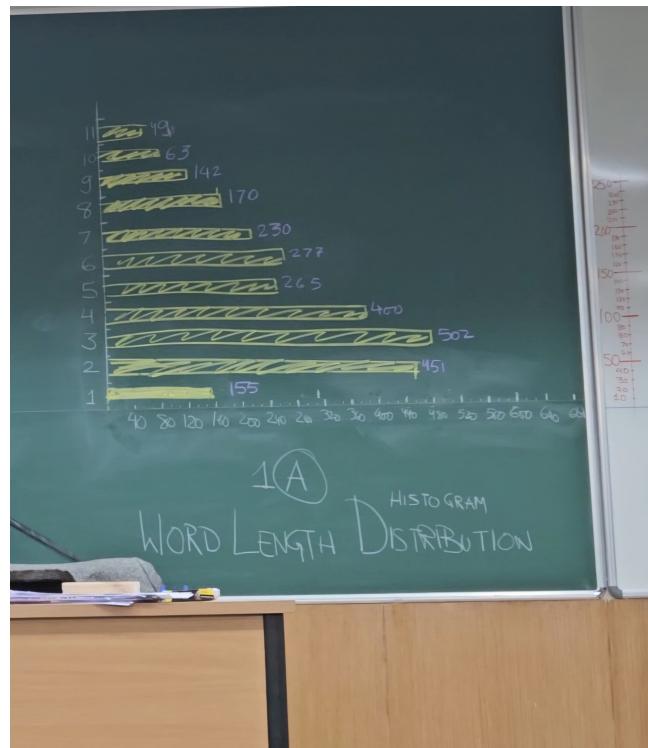


Figure 2: Word length Distribution

### 0.1.3 Part C: Forecasting with Weight and Height Functions

#### Objective

Plot histograms using functions for weight and height over time.

#### Functions

Define the following functions:

$$\begin{aligned} w(t) &= t + 2, \quad (\text{Weight function}) \\ h(t) &= 3t + 5, \quad (\text{Height function}) \end{aligned}$$

where  $t$  represents time (from the beginning to today).

#### Steps

1. Create data points for different values of  $t$  (e.g.,  $t = 0, 1, 2, \dots$ ).
2. Calculate  $w(t)$  and  $h(t)$  for each  $t$ .
3. Plot the histogram using:
  - **X-axis:** Time ( $t$ ).
  - **Y-axis:** Weight ( $w(t)$ ) or Height ( $h(t)$ ).

**Click here to watch the video:** Watch Video

*ANSWER THE FOLLOWING QUESTIONS*

Q1. A. PLOT the histogram on the basis of word length of any Paragraph of your preferred book.

B. PLOT the Starting letter of the Word.

C. Use the following function to PLOT the same:-

Weight =  $w(t)$ , Height =  $h(t)$  where :-

time =  $t$  (beginning, ..... , Today)

forecast the following in Histogram Plotting.

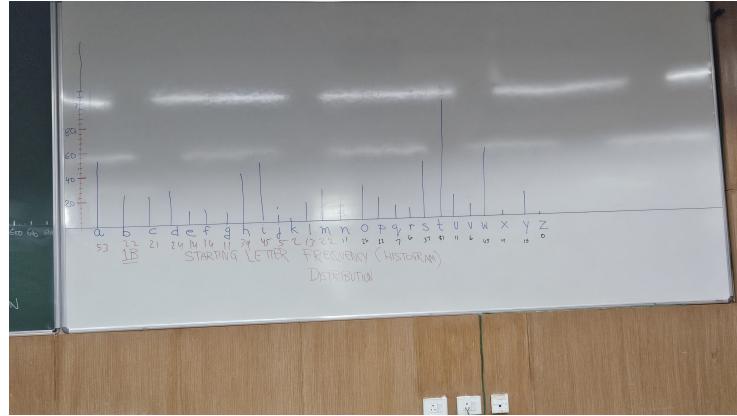


Figure 3: Starting letter distribution

D. Plot a histogram for the dataset [1, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4] with customizations:

Change color, labels, and title.

Add gridlines and annotate the highest frequency bin.

E. Import a real-world dataset (e.g., Iris dataset).

Plot histograms for numerical columns like petal length or sepal width.

—

## 2. HEAT MAPS

A heatmap is a data visualization technique used to represent data in a matrix format with varying intensities of color to show magnitude.

Mathematical Definition:-

For a matrix with elements , a heatmap assigns a color to each cell based on a colormap :

$$c_{ij} = C(a_{ij})$$

Steps to Create a Heatmap:-

1. Normalize the data: Rescale values if needed.

2. Choose a colormap: Maps numerical values to colors.
3. Visualize the data as a grid.

**Python Code Example:-**

```
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
data = np.random.rand(10, 10)
sns.heatmap(data, annot=True, cmap='viridis')
plt.title('Heatmap')
```

**Part 1: Basics of Vector Algebra**

**1.1 Vectors**

A vector is a quantity that has both magnitude and direction. It can be represented in the Cartesian coordinate system with components along the x, y, and z axes.

For example, a vector  $\vec{A}$  in two-dimensional space can be written as:

$$\vec{A} = A_x \hat{i} + A_y \hat{j}$$

where: -  $A_x$  and  $A_y$  are the components of the vector along the x and y axes, respectively. -  $\hat{i}$  and  $\hat{j}$  are the unit vectors along the x and y axes.

**1.2 Vector Operations:-**

\*Addition of Vectors\*: The sum of two vectors is obtained by adding their corresponding components.

$$\vec{A} + \vec{B} = (A_x + B_x) \hat{i} + (A_y + B_y) \hat{j}$$

\*Scalar Multiplication\*: When a vector is multiplied by a scalar, each component of the vector is multiplied by that scalar.

$$k \cdot \vec{A} = k \cdot A_x \hat{i} + k \cdot A_y \hat{j}$$



Figure 4: VECTOR MATRIX

**\*Dot Product\***: The dot product of two vectors is a scalar quantity given by:

$$\vec{A} \cdot \vec{B} = A_x \cdot B_x + A_y \cdot B_y$$

**\*Cross Product\***: The cross product of two vectors gives a vector perpendicular to both vectors. For two 2D vectors  $\vec{A} = A_x \hat{i} + A_y \hat{j}$  and  $\vec{B} = B_x \hat{i} + B_y \hat{j}$ , the cross product is:

$$\vec{A} \times \vec{B} = (A_x B_y - A_y B_x) \hat{k}$$

where  $\hat{k}$  is the unit vector along the z-axis.

### — Part 2: Tug of War Problem

In this problem, we will model a tug of war scenario where three people are pulling a rope from different directions. The goal is to determine the resulting force on the rope by calculating the Vector that is sum of all the individual forces.

#### 2.1 Problem Description

Three people are pulling on a rope from different directions:

- Person 1 applies a force of  $\vec{F}_1 = 10\hat{i} + 5\hat{j}$  Newtons.
- Person 2 applies a force of  $\vec{F}_2 = -8\hat{i} + 7\hat{j}$  Newtons.
- Person 3 applies a force of  $\vec{F}_3 = 3\hat{i} - 6\hat{j}$  Newtons.

We will calculate the net force on the rope by adding these vectors together and analyze the direction and magnitude of the result.

### 2.2 Steps to solve:-

1. Add the vectors  $\vec{F}_1$ ,  $\vec{F}_2$ , and  $\vec{F}_3$ .
2. Compute the magnitude of the resulting vector.
3. Analyze the direction of the resultant vector.

### ACTIVITY:-

#### 0.1.4 Activity: Understanding Vectors through Tug of War

##### Objective

The goal of this activity is to understand the concept of vectors, their operations (addition, magnitude, and direction), by modeling a tug of war scenario where three people are pulling a rope from different directions.

##### Problem Description

Three people are pulling on a rope from different directions:

- Person 1 applies a force of  $\vec{F}_1 = 10\hat{i} + 5\hat{j}$  Newtons.
- Person 2 applies a force of  $\vec{F}_2 = -8\hat{i} + 7\hat{j}$  Newtons.
- Person 3 applies a force of  $\vec{F}_3 = 3\hat{i} - 6\hat{j}$  Newtons.

We will determine the net force on the rope by calculating the vector sum of  $\vec{F}_1$ ,  $\vec{F}_2$ , and  $\vec{F}_3$ . Additionally, we will compute the magnitude and analyze the direction of the resultant vector.

### Steps to Solve

1. Add the vectors  $\vec{F}_1$ ,  $\vec{F}_2$ , and  $\vec{F}_3$ :

$$\begin{aligned}\vec{F}_{\text{net}} &= \vec{F}_1 + \vec{F}_2 + \vec{F}_3 \\ &= (10\hat{i} + 5\hat{j}) + (-8\hat{i} + 7\hat{j}) + (3\hat{i} - 6\hat{j}) \\ &= (10 - 8 + 3)\hat{i} + (5 + 7 - 6)\hat{j} \\ &= 5\hat{i} + 6\hat{j}.\end{aligned}$$

2. Compute the magnitude of the resulting vector  $\vec{F}_{\text{net}}$ :

$$\begin{aligned}|\vec{F}_{\text{net}}| &= \sqrt{(5)^2 + (6)^2} \\ &= \sqrt{25 + 36} \\ &= \sqrt{61} \approx 7.81 \text{ N}.\end{aligned}$$

3. Analyze the direction of the resultant vector:

$$\begin{aligned}\theta &= \tan^{-1} \left( \frac{\text{Vertical component}}{\text{Horizontal component}} \right) \\ &= \tan^{-1} \left( \frac{6}{5} \right) \\ &= \tan^{-1}(1.2) \\ &\approx 50.19^\circ \text{ above the positive } \hat{i}\text{-axis.}\end{aligned}$$

### Discussion Points

- **Addition of Vectors:** Demonstrates how forces combine in two dimensions.
- **Magnitude of the Resultant Vector:** Shows how to calculate the strength of the net force.
- **Direction of the Resultant Vector:** Explains the use of trigonometry to find the orientation of the net force.

### Extension Activity

Challenge students to:

- Add more participants with different forces and calculate the new net force.



Figure 5: VECTOR

- Use graphical methods (parallelogram or head-to-tail) to verify the calculations.
  - Explore real-world scenarios where vectors play a role, such as forces in sports or physics problems involving motion.
- 

### Part 3: *Python Implementation*

Let's implement the solution in Python using numpy to handle vector operations efficiently.

Python:-

```
import numpy as np
```

Define the force vectors

```
F1 = np.array([10, 5]) Force from person 1
```

```
F2 = np.array([-8, 7]) Force from person 2
```

```
F3 = np.array([3, -6]) Force from person 3
```

Step 1: Calculate the net force (vector sum)  $\text{net}_{\text{force}} = F1 + F2 + F3$

Step 2: Calculate the magnitude of the net force  $\text{magnitude} = \text{np.linalg.norm}(\text{net}_{\text{force}})$

Step 3: Find the direction of the net force (angle with respect to x-axis)

```
direction = np.arctan2(net_force[1], net_force[0]) Angle in radians
```

```

directiondegrees = np.degrees(direction)Converttodegrees
print("Net Force Vector (Fnet) : ", netforce)
print("Magnitude of Net Force:", magnitude, "Newtons")
print("Direction of Net Force:", directiondegrees, "degrees")
—

```

### **Output:**

Net Force Vector (F<sub>net</sub>) : [56]  
 Magnitude of Net Force: 7.810249675906654 Newtons  
 Direction of Net Force: 50.19442890773457 degrees.

**Click here to watch the video:** Watch Video

**Click here to watch the video:** Watch Video

**Click here to watch the video:** Watch Video

### **CONDITIONAL PROBABILITY:-**

Conditional probability is the probability of an event occurring given that another event has already occurred. In other words, it quantifies the likelihood of an event  $A$  happening, under the condition that event  $B$  is known to have occurred. The conditional probability of  $A$  given  $B$  is denoted as  $P(A|B)$ , and it is calculated using the formula:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

where:

- $P(A \cap B)$  is the probability that both events  $A$  and  $B$  occur (i.e., the intersection of  $A$  and  $B$ ).
- $P(B)$  is the probability that event  $B$  occurs.

### **Dice Experiment and Conditional Probability**

Let's consider a dice experiment. We roll a standard 6-sided die, and let's define the following events:

- \*Event A\*: The die shows an even number (2, 4, or 6).
- \*Event B\*: The die shows a number greater than 3 (4, 5, or 6).

We want to calculate the conditional probability of event A given that event B has occurred, i.e.,  $P(A|B)$ .

*Step-by-Step Calculation:-*

Step 1: Define the Sample Space The sample space  $S$  for a 6-sided die is:

$$S = \{1, 2, 3, 4, 5, 6\}$$

Step 2: Calculate the Probability of Event B,  $P(B)$  Event B occurs when the die shows a number greater than 3. So,  $B = \{4, 5, 6\}$ . The probability of event B is:

$$P(B) = \frac{\text{Number of outcomes in } B}{\text{Total number of outcomes in } S} = \frac{3}{6} = 0.5$$

Step 3: Calculate the Probability of the Intersection of Events A and B,  $P(A \cap B)$  The intersection of events A and B,  $A \cap B$ , occurs when the die shows a number that is both even (from event A) and greater than 3 (from event B). The numbers that satisfy both conditions are  $\{4, 6\}$ .

So,  $A \cap B = \{4, 6\}$ , and the probability of this intersection is:

$$P(A \cap B) = \frac{\text{Number of outcomes in } A \cap B}{\text{Total number of outcomes in } S} = \frac{2}{6} = 0.3333$$

Step 4: Apply the Conditional Probability Formula Now, we can apply the formula for conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$$

So, the conditional probability of event A (the die shows an even number) given that event B (the die shows a number greater than 3) has occurred is:

$$P(A|B) = \frac{2}{3}$$

### *Interpretation*

Given that the die has shown a number greater than 3, the probability that it is also an even number (either 4 or 6) is  $\frac{2}{3}$ . This conditional probability helps us understand how the knowledge of event B influences the likelihood of event A.

## **Activity: Understanding Conditional Probability Using a Dice Experiment**

### **Objective**

To understand the concept of conditional probability through a simple dice-rolling experiment.

### **Materials Needed**

- A six-sided die (standard die with faces numbered 1 to 6).
- Pen and paper (or a whiteboard) for recording outcomes.

### **Instructions**

#### **Step 1: Define the Experiment**

1. Roll a standard six-sided die multiple times (e.g., 20 trials).
2. Record the outcome of each roll.

#### **Step 2: Define Events**

Let:

- **Event A:** The die roll results in an even number ( $\{2, 4, 6\}$ ).
- **Event B:** The die roll results in a number greater than 3 ( $\{4, 5, 6\}$ ).

### **Step 3: Conduct the Experiment**

1. Roll the die and record the outcomes.
2. For each roll, determine whether it satisfies Event A, Event B, or both.

### **Step 4: Analyze the Results**

1. Count the number of rolls where:
  - Event A occurs ( $n(A)$ ).
  - Event B occurs ( $n(B)$ ).
  - Both Event A and Event B occur ( $n(A \cap B)$ ).
2. Use the frequency of these events to estimate probabilities:

$$P(A) = \frac{n(A)}{\text{total rolls}}, \quad P(B) = \frac{n(B)}{\text{total rolls}}, \quad P(A \cap B) = \frac{n(A \cap B)}{\text{total rolls}}.$$

### **Step 5: Calculate Conditional Probabilities**

Using the formula for conditional probability:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

1. Substitute the experimental values of  $P(A \cap B)$  and  $P(B)$  into the formula.
2. Interpret the result:
  - $P(A | B)$ : Probability of rolling an even number, given that the number is greater than 3.

### **Example of Recorded Data (Sample Table)**

Trial	Outcome	Event A (Even)	Event B (> 3)	Both A and B
1	4	Yes	Yes	Yes
2	2	Yes	No	No
3	5	No	Yes	No
4	6	Yes	Yes	Yes
5	3	No	No	No
...	...	...	...	...

## Discussion Points

### 1. Theoretical Probability:

- $P(A) = \frac{3}{6} = 0.5$  (since there are 3 even numbers out of 6).
- $P(B) = \frac{3}{6} = 0.5$  (since there are 3 numbers greater than 3).
- $P(A \cap B) = \frac{2}{6} = 0.333$  (only 4 and 6 satisfy both conditions).

Using these values:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0.333}{0.5} \approx 0.666$$

### 2. Connection to Real-World Applications:

Conditional probability is used in fields like medicine (e.g., probability of disease given a positive test result) and weather forecasting.

## Extension Activity

- Modify the conditions of the experiment. For example:
  - Event A: Rolling a multiple of 3 ( $\{3, 6\}$ ).
  - Event B: Rolling an odd number ( $\{1, 3, 5\}$ ).
- Explore the relationship between the events when they are independent or mutually exclusive.

### *Python Code Implementation:-*

We can simulate this dice experiment and compute the conditional probability using Python:

```
import numpy as np
```

Define the sample space for a 6-sided die

```
sample_space = [1, 2, 3, 4, 5, 6]
```

Define the events A and B

```
A = [2, 4, 6] Even numbers
```

```
B = [4, 5, 6] Numbers greater than 3
```

Calculate  $P(A \cap B)$  - intersection of A and B

```
intersection_AB = list(set(A) & set(B))
```

Calculate  $P(B)$

$$P_B = \text{len}(B) / \text{len}(\text{sample\_space})$$

Calculate  $P(A \cap B)$

$$P_{A \cap B} = \text{len}(\text{intersection}_{AB}) / \text{len}(\text{sample\_space})$$

Calculate  $P(A - B)$  using the formula

$$P_{A|B} = P_{A \cap B} / P_B$$

Display the result

```
print("P(A - B) =", P_A|B)
```

Output:

text

$$P(A - B) = 0.6666666666666666$$

This Python code confirms our theoretical calculation that the conditional probability of event A given event B is  $\frac{2}{3}$ .

**Click here to watch the video:** Watch Video

**Click here to watch the video:** Watch Video

**Click here to watch the video:** Watch Video