Introduction Statistical Methods- Session-4

Bayes' Theorem Explained

Bayes' Theorem Formula

Bayes' Theorem is a way to calculate the probability of an event A, given that another event B has occurred. It's expressed as:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Where:

- $P(A \mid B)$: Probability of A occurring given B has occurred.
- $P(B \mid A)$: Probability of B occurring given A has occurred.
- $P(A)$: Prior probability of A occurring.
- $P(B)$: Total probability of B occurring.

Expanded Formula for Multiple Hypotheses

If there are multiple mutually exclusive events ($A_1, A_2, \ldots, A_n$) that could lead to B, the formula for $P(B)$ is:

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i)$$

Then:

$$P(A_k \mid B) = \frac{P(B \mid A_k) \cdot P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i)}$$

Example Problem

Scenario

A factory has three machines M_1, M_2, and M_3 that produce items. The proportion of items produced by each machine and the defect rates are as follows:

- M_1: Produces 30% of the items, with a defect rate of 2%.
- M_2: Produces 50% of the items, with a defect rate of 3%.
- M_3: Produces 20% of the items, with a defect rate of 5%.

An item is randomly selected, and it is found to be defective. What is the probability it was produced by M_1?

Step 1: Define the Events

- A_1: Item is produced by M_1.
- A_2: Item is produced by M_2.
- A_3: Item is produced by M_3.
- B: Item is defective.

We need $P(A_1 \mid B)$, the probability the defective item came from M_1.

Step 2: Use Bayes' Theorem

$$P(A_1 \mid B) = \frac{P(B \mid A_1) \cdot P(A_1)}{P(B)}$$

Step 3: Calculate Components

1. $P(A_1)$: Probability item is produced by M_1:

$$P(A_1) = 0.3$$

2. $P(B \mid A_1)$: Probability item is defective given it is from M_1:

$$P(B \mid A_1) = 0.02$$

3. $P(B)$: Total probability of a defective item:

$$P(B) = P(B \mid A_1) \cdot P(A_1) + P(B \mid A_2) \cdot P(A_2) + P(B \mid A_3) \cdot P(A_3)$$

Substitute values:

$$P(B) = (0.02 \cdot 0.3) + (0.03 \cdot 0.5) + (0.05 \cdot 0.2)$$

$$P(B) = 0.006 + 0.015 + 0.01 = 0.031$$

Step 4: Apply Bayes' Theorem

Substitute into the formula:

$$P(A_1 \mid B) = \frac{P(B \mid A_1) \cdot P(A_1)}{P(B)}$$

$$P(A_1 \mid B) = \frac{0.02 \cdot 0.3}{0.031}$$

$$P(A_1 \mid B) = \frac{0.006}{0.031} \approx 0.1935$$

Example:1

The Probabilities of X, Y *and* Z becoming managers are $\frac{4}{9}, \frac{2}{9}$, *and* $\frac{1}{3}$ respectively. The probabilities that the bonus schemes will be introduced if X ,Y and Z becomes managers are $\frac{3}{10}, \frac{1}{2}$ and $\frac{4}{5}$ respectively.

(i) What is the probability that bonus scheme will be introduced ?

(ii) If the bonus scheme has been introduced, What is the probability that the manager appointed was X?

Given Information

1. Probabilities of becoming a manager:
   - $P(X) = \frac{4}{9}$
   - $P(Y) = \frac{2}{9}$
   - $P(Z) = \frac{1}{3} = \frac{3}{9}$
2. Conditional probabilities of introducing the bonus scheme:
   - $P(B \mid X) = \frac{3}{10}$
   - $P(B \mid Y) = \frac{1}{2} = \frac{5}{10}$
   - $P(B \mid Z) = \frac{4}{5} = \frac{8}{10}$
3. B: Event that the bonus scheme is introduced.

We are solving for: (i) The probability that the bonus scheme will be introduced ($P(B)$). (ii) Given that the bonus scheme is introduced, the probability that X was appointed as manager ($P(X \mid B)$).

Part (i): Probability that the bonus scheme will be introduced ($P(B)$)

Step 1: Total Probability Formula

The total probability of B is given by:

$$P(B) = P(B \mid X) \cdot P(X) + P(B \mid Y) \cdot P(Y) + P(B \mid Z) \cdot P(Z)$$

Step 2: Substitute the Values

Substitute the given probabilities:

$$P(B) = \left(\frac{3}{10} \cdot \frac{4}{9}\right) + \left(\frac{5}{10} \cdot \frac{2}{9}\right) + \left(\frac{8}{10} \cdot \frac{3}{9}\right)$$

Step 3: Calculate Each Term

1. $\frac{3}{10} \cdot \frac{4}{9} = \frac{12}{90}$
2. $\frac{5}{10} \cdot \frac{2}{9} = \frac{10}{90}$
3. $\frac{8}{10} \cdot \frac{3}{9} = \frac{24}{90}$

Add these:

$$P(B) = \frac{12}{90} + \frac{10}{90} + \frac{24}{90} = \frac{46}{90}$$

Simplify:

$$P(B) = \frac{23}{45}$$

Part (ii): Probability that X was appointed as manager given the bonus scheme was introduced ($P(X \mid B)$)

Step 1: Use Bayes' Theorem

Bayes' Theorem states:

$$P(X \mid B) = \frac{P(B \mid X) \cdot P(X)}{P(B)}$$

Step 2: Substitute the Values

From Part (i), $P(B) = \frac{23}{45}$, $P(B \mid X) = \frac{3}{10}$, and $P(X) = \frac{4}{9}$. Substitute these:

$$P(X \mid B) = \frac{\left(\frac{3}{10} \cdot \frac{4}{9}\right)}{\frac{23}{45}}$$

Step 3: Simplify the Numerator

$$\frac{3}{10} \cdot \frac{4}{9} = \frac{12}{90}$$

Step 4: Divide by $P(B)$

$$P(X \mid B) = \frac{\frac{12}{90}}{\frac{23}{45}} = \frac{12}{90} \cdot \frac{45}{23} = \frac{540}{2070} = \frac{6}{23}$$

Example: 2

In a neighborhood, 90% children were falling sick due to flu and 10% due to measles and no other disease. The probability of observing rashes for measles is 0.95 and for flu is 0.08. If a child develops rashes, Find the child's probability of having flu.

Step 1: Key Formula (Bayes' Theorem)

Bayes' Theorem states:

$$P(\text{Flu} \mid \text{Rashes}) = \frac{P(\text{Rashes} \mid \text{Flu}) \cdot P(\text{Flu})}{P(\text{Rashes})}$$

Where:

- $P(\text{Flu} \mid \text{Rashes})$: Probability the child has flu given they have rashes (what we want to calculate).
- $P(\text{Rashes} \mid \text{Flu})$: Probability of rashes given the child has flu (0.08).
- $P(\text{Flu})$: Probability the child has flu ($90\% = 0.9$).
- $P(\text{Rashes})$: Total probability of rashes.

Step 2: Total Probability of Rashes

To calculate $P(\text{Rashes})$, we use the **law of total probability:**

$$P(\text{Rashes}) = P(\text{Rashes} \mid \text{Flu}) \cdot P(\text{Flu}) + P(\text{Rashes} \mid \text{Measles}) \cdot P(\text{Measles})$$

From the problem:

- $P(\text{Flu}) = 0.9$
- $P(\text{Measles}) = 0.1$ (since 90% is due to flu, 10% must be due to measles).
- $P(\text{Rashes} \mid \text{Flu}) = 0.08$
- $P(\text{Rashes} \mid \text{Measles}) = 0.95$

Substitute the values:

$$P(\text{Rashes}) = (0.08 \cdot 0.9) + (0.95 \cdot 0.1)$$

Calculate:

$$P(\text{Rashes}) = 0.072 + 0.095 = 0.167$$

Step 3: Apply Bayes' Theorem

Now substitute the values into Bayes' Theorem:

$$P(\text{Flu} \mid \text{Rashes}) = \frac{P(\text{Rashes} \mid \text{Flu}) \cdot P(\text{Flu})}{P(\text{Rashes})}$$

Substitute the values:

$$P(\text{Flu} \mid \text{Rashes}) = \frac{0.08 \cdot 0.9}{0.167}$$

Calculate:

$$P(\text{Flu} \mid \text{Rashes}) = \frac{0.072}{0.167} \approx 0.431$$

Final Answer

The probability that the child has flu, given that they have developed rashes, is approximately:

$$\boxed{0.431 \text{ or } 43.1\%}$$

Example: 3

It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam e-mails, and the probability for a false positive (a non-spam email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

Step 1: Key Formula (Bayes' Theorem)

Bayes' Theorem is given by:

$$P(\text{Non-Spam} \mid \text{Detected as Spam}) = \frac{P(\text{Detected as Spam} \mid \text{Non-Spam}) \cdot P(\text{Non-Spam})}{P(\text{Detected as Spam})}$$

Where:

- $P(\text{Non-Spam} \mid \text{Detected as Spam})$: Probability the email is non-spam given it is detected as spam (what we want to calculate).
- $P(\text{Detected as Spam} \mid \text{Non-Spam})$: Probability that a non-spam email is detected as spam (false positive rate = $5\% = 0.05$).
- $P(\text{Non-Spam})$: Probability that an email is non-spam ($50\% = 0.5$).
- $P(\text{Detected as Spam})$: Total probability that an email is detected as spam.

Step 2: Total Probability of Being Detected as Spam

To calculate $P(\text{Detected as Spam})$, we use the **law of total probability**:

$$P(\text{Detected as Spam}) = P(\text{Detected as Spam} \mid \text{Spam}) \cdot P(\text{Spam}) + P(\text{Detected as Spam} \mid \text{Non-Spam}) \cdot P(\text{Non-Spam})$$

From the problem:

- $P(\text{Spam}) = 0.5$ (50% of emails are spam).
- $P(\text{Non-Spam}) = 0.5$ (remaining 50% are non-spam).
- $P(\text{Detected as Spam} \mid \text{Spam}) = 99\% = 0.99$ (true positive rate).
- $P(\text{Detected as Spam} \mid \text{Non-Spam}) = 5\% = 0.05$ (false positive rate).

Substitute these values:

$$P(\text{Detected as Spam}) = (0.99 \cdot 0.5) + (0.05 \cdot 0.5)$$

Calculate:

$$P(\text{Detected as Spam}) = 0.495 + 0.025 = 0.52$$

Step 3: Apply Bayes' Theorem

Now substitute the values into Bayes' Theorem:

$$P(\text{Non-Spam} \mid \text{Detected as Spam}) = \frac{P(\text{Detected as Spam} \mid \text{Non-Spam}) \cdot P(\text{Non-Spam})}{P(\text{Detected as Spam})}$$

Substitute the values:

$$P(\text{Non-Spam} \mid \text{Detected as Spam}) = \frac{0.05 \cdot 0.5}{0.52}$$

Calculate:

$$P(\text{Non-Spam} \mid \text{Detected as Spam}) = \frac{0.025}{0.52} \approx 0.0481$$

Final Answer

The probability that an email detected as spam is actually **non-spam** is approximately:

$$\boxed{0.0481 \text{ or } 4.81\%}$$

Bayes' Theorem for Hypotheses

Bayes' Theorem for multiple hypotheses provides a way to update the probability of each hypothesis based on observed evidence.

Formula

If $H_1, H_2, \ldots, H_n$ are mutually exclusive and exhaustive hypotheses, and E is some observed evidence, Bayes' Theorem states:

$$P(H_k \mid E) = \frac{P(E \mid H_k) \cdot P(H_k)}{\sum_{i=1}^{n} P(E \mid H_i) \cdot P(H_i)}$$

Where:

- $P(H_k \mid E)$: Posterior probability of hypothesis H_k given evidence E.
- $P(E \mid H_k)$: Likelihood of observing E if hypothesis H_k is true.
- $P(H_k)$: Prior probability of hypothesis H_k.
- $\sum_{i=1}^{n} P(E \mid H_i) \cdot P(H_i)$: Total probability of observing E across all hypotheses.

Example Problem

Scenario

A diagnostic lab is testing for a rare disease. The disease can be caused by one of three possible viruses: V_1, V_2, or V_3. The probabilities are:

- $P(V_1) = 0.4$, $P(V_2) = 0.35$, and $P(V_3) = 0.25$.

The test result E (evidence) shows a positive result. The probabilities of getting a positive result given each virus are:

- $P(E \mid V_1) = 0.8$,
- $P(E \mid V_2) = 0.6$,
- $P(E \mid V_3) = 0.7$.

What is the probability that the disease is caused by virus V_1, given the positive test result?

Step 1: Apply Bayes' Theorem

Using the formula:

$$P(V_1 \mid E) = \frac{P(E \mid V_1) \cdot P(V_1)}{\sum_{i=1}^{3} P(E \mid V_i) \cdot P(V_i)}$$

Step 2: Calculate Total Probability of E

The total probability of observing E is:

$$P(E) = P(E \mid V_1) \cdot P(V_1) + P(E \mid V_2) \cdot P(V_2) + P(E \mid V_3) \cdot P(V_3)$$

Substitute the values:

$$P(E) = (0.8 \cdot 0.4) + (0.6 \cdot 0.35) + (0.7 \cdot 0.25)$$

$$P(E) = 0.32 + 0.21 + 0.175 = 0.705$$

Step 3: Calculate $P(V_1 \mid E)$

Substitute into Bayes' Theorem:

$$P(V_1 \mid E) = \frac{P(E \mid V_1) \cdot P(V_1)}{P(E)}$$

$$P(V_1 \mid E) = \frac{0.8 \cdot 0.4}{0.705}$$

$$P(V_1 \mid E) = \frac{0.32}{0.705} \approx 0.454$$

Final Answer

The probability that the disease is caused by V_1, given the positive test result, is approximately:

$$\boxed{0.454 \text{ or } 45.4\%}$$

Example: 4

- Consider a medical diagnosis problem in which there are two alternative hypotheses:

H1:That a patient has a particular form of cancer

H2:That the patient does not

- The available data is from a particular laboratory test with two possible outcomes

  + Positive

  - Negative

- Over the entire population of people only 0.008 have this disease. The test returns a corrective positive result in only 98% of the cases in which the disease is actually present and a correct negative result in only 97% of the cases in which the disease is not present.

- How does P(cancer/+) compare to P(~ cancer/+)?

To solve this problem, we will use **Bayes' Theorem** to compute:

1. $P(\text{cancer} \mid +)$: The probability that the patient has cancer given a positive test result.
2. $P(\sim \text{cancer} \mid +)$: The probability that the patient does not have cancer given a positive test result.

Finally, we will compare $P(\text{cancer} \mid +)$ and $P(\sim \text{cancer} \mid +)$.

Step 1: Given Information

- $P(\text{cancer}) = 0.008$ (prior probability of having cancer).
- $P(\sim \text{cancer}) = 1 - P(\text{cancer}) = 0.992$ (prior probability of not having cancer).
- $P(+ \mid \text{cancer}) = 0.98$ (true positive rate — test correctly identifies cancer).
- $P(- \mid \sim \text{cancer}) = 0.97$ (true negative rate — test correctly identifies no cancer).
- $P(+ \mid \sim \text{cancer}) = 1 - P(- \mid \sim \text{cancer}) = 0.03$ (false positive rate — test incorrectly identifies no cancer as positive).

Step 2: Formula for Bayes' Theorem

For any hypothesis H:

$$P(H \mid +) = \frac{P(+ \mid H) \cdot P(H)}{P(+)}$$

Here:

- H: Hypothesis (e.g., cancer or no cancer).
- $+$: Evidence (positive test result).
- $P(+ \mid H)$: Likelihood of a positive test result given the hypothesis.
- $P(H)$: Prior probability of the hypothesis.
- $P(+)$: Total probability of a positive test result (calculated using the **law of total probability**).

Step 3: Calculate Total Probability of a Positive Test ($P(+)$)

$$P(+) = P(+ \mid \text{cancer}) \cdot P(\text{cancer}) + P(+ \mid \sim \text{cancer}) \cdot P(\sim \text{cancer})$$

Substitute the values:

$$P(+) = (0.98 \cdot 0.008) + (0.03 \cdot 0.992)$$

Calculate:

$$P(+) = 0.00784 + 0.02976 = 0.0376$$

Step 4: Calculate $P(\text{cancer} \mid +)$

Using Bayes' Theorem:

$$P(\text{cancer} \mid +) = \frac{P(+ \mid \text{cancer}) \cdot P(\text{cancer})}{P(+)}$$

Substitute the values:

$$P(\text{cancer} \mid +) = \frac{0.98 \cdot 0.008}{0.0376}$$

Calculate:

$$P(\text{cancer} \mid +) = \frac{0.00784}{0.0376} \approx 0.2085$$

Step 5: Calculate $P(\sim \text{cancer} \mid +)$

Using Bayes' Theorem:

$$P(\sim \text{cancer} \mid +) = \frac{P(+ \mid \sim \text{cancer}) \cdot P(\sim \text{cancer})}{P(+)}$$

Substitute the values:

$$P(\sim \text{cancer} \mid +) = \frac{0.03 \cdot 0.992}{0.0376}$$

Calculate:

$$P(\sim \text{cancer} \mid +) = \frac{0.02976}{0.0376} \approx 0.7915$$

Step 6: Compare $P(\text{cancer} \mid +)$ and $P(\sim \text{cancer} \mid +)$

From the calculations:

- $P(\text{cancer} \mid +) \approx 0.2085$ (20.85%).
- $P(\sim \text{cancer} \mid +) \approx 0.7915$ (79.15%).

The probability of not having cancer given a positive test result ($P(\sim \text{cancer} \mid +)$) is significantly higher than the probability of having cancer given a positive test result ($P(\text{cancer} \mid +)$).

Final Answer

- $P(\text{cancer} \mid +) \approx 0.2085$ or 20.85%
- $P(\sim \text{cancer} \mid +) \approx 0.7915$ or 79.15%

Example: 5

If the weather is sunny, Then the player will play or not?

i.e., Play /sunny = Yes or No.

Note: If we know P(Yes/Sunny) and P(No/Sunny)

then we can answer the question asked

Weather	Play
Sunny	No
Overcast	Yes
Rainy	Yes
Sunny	Yes
Sunny	Yes
Overcast	Yes
Rainy	No
Rainy	No
Sunny	Yes
Rainy	Yes
Sunny	No
Overcast	Yes
Overcast	Yes
Rainy	No

We will calculate $P(\text{Yes} \mid \text{Sunny})$ and $P(\text{No} \mid \text{Sunny})$ using **Bayes' Theorem**. This will help us determine whether the player will play or not if the weather is sunny.

Step 1: Information from the Table

From the data in the table:

- Total number of instances: 14
- Weather is **Sunny** in 5 instances (as seen in the table).
- Out of these:
  - Player says **Yes** to play in 2 cases.
  - Player says **No** to play in 3 cases.

Step 2: Bayes' Theorem

Bayes' Theorem formula for $P(\text{Yes} \mid \text{Sunny})$ and $P(\text{No} \mid \text{Sunny})$ is:

$$P(\text{Yes} \mid \text{Sunny}) = \frac{P(\text{Sunny} \mid \text{Yes}) \cdot P(\text{Yes})}{P(\text{Sunny})}$$

$$P(\text{No} \mid \text{Sunny}) = \frac{P(\text{Sunny} \mid \text{No}) \cdot P(\text{No})}{P(\text{Sunny})}$$

Step 3: Calculate Individual Probabilities

1. $P(\text{Sunny})$:

$$P(\text{Sunny}) = \frac{\text{Number of Sunny cases}}{\text{Total cases}} = \frac{5}{14}$$

2. $P(\text{Yes})$:

$$P(\text{Yes}) = \frac{\text{Number of Yes cases}}{\text{Total cases}} = \frac{9}{14}$$

3. $P(\text{No})$:

$$P(\text{No}) = \frac{\text{Number of No cases}}{\text{Total cases}} = \frac{5}{14}$$

4. $P(\text{Sunny} \mid \text{Yes})$:

Out of 9 "Yes" cases, 2 are "Sunny":

$$P(\text{Sunny} \mid \text{Yes}) = \frac{2}{9}$$

5. $P(\text{Sunny} \mid \text{No})$:

Out of 5 "No" cases, 3 are "Sunny":

$$P(\text{Sunny} \mid \text{No}) = \frac{3}{5}$$

Step 4: Calculate $P(\text{Yes} \mid \text{Sunny})$

Substitute into Bayes' Theorem:

$$P(\text{Yes} \mid \text{Sunny}) = \frac{P(\text{Sunny} \mid \text{Yes}) \cdot P(\text{Yes})}{P(\text{Sunny})}$$

Substitute the values:

$$P(\text{Yes} \mid \text{Sunny}) = \frac{\frac{2}{9} \cdot \frac{9}{14}}{\frac{5}{14}}$$

Simplify:

$$P(\text{Yes} \mid \text{Sunny}) = \frac{\frac{2}{14}}{\frac{5}{14}} = \frac{2}{5}$$

Step 5: Calculate $P(\text{No} \mid \text{Sunny})$

Substitute into Bayes' Theorem:

$$P(\text{No} \mid \text{Sunny}) = \frac{P(\text{Sunny} \mid \text{No}) \cdot P(\text{No})}{P(\text{Sunny})}$$

Substitute the values:

$$P(\text{No} \mid \text{Sunny}) = \frac{\frac{3}{5} \cdot \frac{5}{14}}{\frac{5}{14}}$$

Simplify:

$$P(\text{No} \mid \text{Sunny}) = \frac{\frac{3}{14}}{\frac{5}{14}} = \frac{3}{5}$$

Step 5: Calculate $P(\text{No} \mid \text{Sunny})$

Substitute into Bayes' Theorem:

$$P(\text{No} \mid \text{Sunny}) = \frac{P(\text{Sunny} \mid \text{No}) \cdot P(\text{No})}{P(\text{Sunny})}$$

Substitute the values:

$$P(\text{No} \mid \text{Sunny}) = \frac{\frac{3}{5} \cdot \frac{5}{14}}{\frac{5}{14}}$$

Simplify:

$$P(\text{No} \mid \text{Sunny}) = \frac{\frac{3}{14}}{\frac{5}{14}} = \frac{3}{5}$$

Step 6: Conclusion

From the calculations:

- $P(\text{Yes} \mid \text{Sunny}) = \frac{2}{5} = 0.4$ (40%)
- $P(\text{No} \mid \text{Sunny}) = \frac{3}{5} = 0.6$ (60%)

Thus, if the weather is sunny, the player is **more likely NOT to play.**

Naive Bayes Classifier

Naive Bayes is a probabilistic classifier based on **Bayes' Theorem** and assumes that the features (variables) are **conditionally independent** given the class. This "naive" assumption simplifies the computation significantly, making it scalable and efficient.

Bayes' Theorem

The general formula of Bayes' Theorem is:

$$P(C \mid X) = \frac{P(X \mid C) \cdot P(C)}{P(X)}$$

Where:

- $P(C \mid X)$: Posterior probability of class C given the feature set X.
- $P(X \mid C)$: Likelihood of the feature set X given class C.
- $P(C)$: Prior probability of class C.
- $P(X)$: Total probability of feature set X (acts as a normalization factor).

Naive Bayes Formula

For multiple features $X = \{x_1, x_2, \ldots, x_n\}$, Naive Bayes assumes conditional independence:

$$P(C \mid X) \propto P(C) \cdot P(x_1 \mid C) \cdot P(x_2 \mid C) \cdot \cdots \cdot P(x_n \mid C)$$

Here:

- $P(C)$: Prior probability of the class C.
- $P(x_i \mid C)$: Likelihood of the feature x_i given class C.

The class with the **highest posterior probability** $P(C \mid X)$ is chosen as the predicted class.

Example

Problem: Classify Emails as Spam or Not Spam

Given:

1. Two classes: Spam (S) and Not Spam (NS).
2. Features: Words in the email (e.g., "win," "offer").
3. Training Data (for simplicity, counts):
    - Spam: Appears 70% of the time ($P(S) = 0.7$).
    - Not Spam: Appears 30% of the time ($P(NS) = 0.3$).
    - Word "win" appears in:
        - 50% of spam emails ($P(\text{win} \mid S) = 0.5$).
        - 10% of non-spam emails ($P(\text{win} \mid NS) = 0.1$).
    - Word "offer" appears in:
        - 40% of spam emails ($P(\text{offer} \mid S) = 0.4$).
        - 20% of non-spam emails ($P(\text{offer} \mid NS) = 0.2$).

Given an email containing "win" and "offer," classify it as Spam or Not Spam.

Step 1: Calculate $P(S \mid \text{win, offer})$

Using Naive Bayes:

$$P(S \mid \text{win, offer}) \propto P(S) \cdot P(\text{win} \mid S) \cdot P(\text{offer} \mid S)$$

Substitute the values:

$$P(S \mid \text{win, offer}) \propto 0.7 \cdot 0.5 \cdot 0.4 = 0.14$$

Step 2: Calculate $P(NS \mid \text{win, offer})$

Using Naive Bayes:

$$P(NS \mid \text{win, offer}) \propto P(NS) \cdot P(\text{win} \mid NS) \cdot P(\text{offer} \mid NS)$$

Substitute the values:

$$P(NS \mid \text{win, offer}) \propto 0.3 \cdot 0.1 \cdot 0.2 = 0.006$$

Step 3: Normalize Probabilities

To compare $P(S \mid \text{win, offer})$ and $P(NS \mid \text{win, offer})$, normalize by dividing each by the total probability:

$$P(S \mid \text{win, offer}) = \frac{0.14}{0.14 + 0.006} \approx 0.96$$

$$P(NS \mid \text{win, offer}) = \frac{0.006}{0.14 + 0.006} \approx 0.04$$

Step 4: Conclusion

Since $P(S \mid \text{win, offer}) > P(NS \mid \text{win, offer})$, the email is classified as **Spam**.

Key Points

1. **Conditional Independence:** Naive Bayes assumes features are independent given the class.
2. **Scalability:** Efficient for high-dimensional data (like text classification).
3. **Applications:** Used in spam detection, sentiment analysis, medical diagnosis, etc.

Final Formula

The general formula for Naive Bayes Classification is:

$$P(C \mid X) \propto P(C) \cdot \prod_{i=1}^{n} P(x_i \mid C)$$

Example 6:

Suppose we got the new message with the words ‘**Dear Friend**’, Decide whether this new message is a normal or spam message?

i.e. Normal/ Dear, Friend = Yes or No

Note if we know P (Normal/ Dear, Friend) and P(Spam/ Dear, Friend) then we can answer the question asked

Email word	Spam	Email word	Spam
Dear	Yes	Friend	No
Friend	No	Friend	Yes
Dear	No	Dear	No
Dear	No	Lunch	No
Dear	No	Friend	No
Friend	No	Dear	No
Lunch	No	Dear	No
Friend	No	Dear	No
Lunch	No	Dear	No
Dear	Yes	Money	Yes
Money	Yes	Money	No
Money	Yes	Money	Yes

We will use **Naive Bayes Classifier** to decide whether the player will play or not given the conditions:

Today: (Outlook = Sunny, Temp = Hot, Humidity = Normal, Windy = False)

Step 1: Extract Information from the Table

The table shows 14 instances of data.

- Total instances: 14
- Play Tennis = Yes: 9

$$P(\text{Play} = \text{Yes}) = \frac{9}{14}$$

- Play Tennis = No: 5

$$P(\text{Play} = \text{No}) = \frac{5}{14}$$

Step 2: Calculate Conditional Probabilities

We calculate $P(\text{Sunny} \mid \text{Yes})$, $P(\text{Hot} \mid \text{Yes})$, $P(\text{Normal} \mid \text{Yes})$, $P(\text{False} \mid \text{Yes})$, and similarly for No.

1. For Play = Yes:

- $P(\text{Sunny} \mid \text{Yes})$:
  - Total Sunny days = 5
  - Sunny days where Play = Yes = 2:

$$P(\text{Sunny} \mid \text{Yes}) = \frac{2}{9}$$

- $P(\text{Hot} \mid \text{Yes})$:
  - Total Hot days = 4
  - Hot days where Play = Yes = 2:

$$P(\text{Hot} \mid \text{Yes}) = \frac{2}{9}$$

- $P(\text{Normal Humidity} \mid \text{Yes})$:
  - Total Normal Humidity days = 7
  - Normal Humidity days where Play = Yes = 6:

$$P(\text{Normal} \mid \text{Yes}) = \frac{6}{9}$$

- $P(\text{Windy} = \text{False} \mid \text{Yes})$:
  - Total days where Windy = False = 8
  - Windy = False days where Play = Yes = 6:

$$P(\text{False} \mid \text{Yes}) = \frac{6}{9}$$

2. For Play = No:

- $P(\text{Sunny} \mid \text{No})$:
  - Total Sunny days = 5
  - Sunny days where Play = No = 3:

$$P(\text{Sunny} \mid \text{No}) = \frac{3}{5}$$

- $P(\text{Hot} \mid \text{No})$:
  - Total Hot days = 4
  - Hot days where Play = No = 2:

$$P(\text{Hot} \mid \text{No}) = \frac{2}{5}$$

- $P(\text{Normal Humidity} \mid \text{No})$:
  - Total Normal Humidity days = 7
  - Normal Humidity days where Play = No = 1:

$$P(\text{Normal} \mid \text{No}) = \frac{1}{5}$$

- $P(\text{Windy} = \text{False} \mid \text{No})$:
  - Total days where Windy = False = 8
  - Windy = False days where Play = No = 2:

$$P(\text{False} \mid \text{No}) = \frac{2}{5}$$

Step 3: Apply Naive Bayes Formula

For Play = Yes:

$P(\text{Yes} \mid \text{Sunny, Hot, Normal, False}) \propto P(\text{Yes}) \cdot P(\text{Sunny} \mid \text{Yes}) \cdot P(\text{Hot} \mid \text{Yes}) \cdot P(\text{Normal} \mid \text{Yes}) \cdot P(\text{False} \mid \text{Yes})$

Substitute values:

$$P(\text{Yes} \mid \text{Sunny, Hot, Normal, False}) \propto \frac{9}{14} \cdot \frac{2}{9} \cdot \frac{2}{9} \cdot \frac{6}{9} \cdot \frac{6}{9}$$

Simplify:

$$P(\text{Yes} \mid \text{Sunny, Hot, Normal, False}) \propto \frac{9}{14} \cdot \frac{144}{6561} = \frac{1296}{91854}$$

For Play = No:

$P(\text{No} \mid \text{Sunny, Hot, Normal, False}) \propto P(\text{No}) \cdot P(\text{Sunny} \mid \text{No}) \cdot P(\text{Hot} \mid \text{No}) \cdot P(\text{Normal} \mid \text{No}) \cdot P(\text{False} \mid \text{No})$

Substitute values:

$$P(\text{No} \mid \text{Sunny, Hot, Normal, False}) \propto \frac{5}{14} \cdot \frac{3}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} \cdot \frac{2}{5}$$

Simplify:

$$P(\text{No} \mid \text{Sunny, Hot, Normal, False}) \propto \frac{5}{14} \cdot \frac{12}{625} = \frac{60}{8750}$$

Step 4: Compare Probabilities

- Calculate normalized probabilities for $P(\text{Yes})$ and $P(\text{No})$.
- $P(\text{Yes})$ is **much higher**, so the player will **play tennis.**

Final Answer:

The player will:

$$\boxed{\text{Play Tennis}}$$

Practice problems

Problem 1

Two set of candidates are competing for the positions on the board of directors of a company. The probabilities that the first and the second sets will win are 0.6 and 0.4 respectively. If the first set wins, the probability of introducing a new product is 0.8, and the corresponding probability if the second set wins is 0.3. What is the probability that the product will be introduced?

Step 1: Define the Events

Let:

- A_1: Event that the **first set wins**.
- A_2: Event that the **second set wins**.
- $P(A_1) = 0.6$: Probability the first set wins.
- $P(A_2) = 0.4$: Probability the second set wins.
- $P(B \mid A_1) = 0.8$: Probability of introducing the product if the first set wins.
- $P(B \mid A_2) = 0.3$: Probability of introducing the product if the second set wins.
- B: Event that the product will be introduced.

We aim to find $P(B)$, the probability that the product will be introduced.

Step 2: Law of Total Probability

The total probability of B is given by:

$$P(B) = P(B \mid A_1) \cdot P(A_1) + P(B \mid A_2) \cdot P(A_2)$$

Step 3: Substitute the Values

Substitute the given probabilities:

$$P(B) = (0.8 \cdot 0.6) + (0.3 \cdot 0.4)$$

Step 4: Calculate

1. Calculate the first term:

$$0.8 \cdot 0.6 = 0.48$$

2. Calculate the second term:

$$0.3 \cdot 0.4 = 0.12$$

3. Add the terms:

$$P(B) = 0.48 + 0.12 = 0.6$$

Final Answer

The probability that the product will be introduced is:

$$\boxed{0.6 \text{ or } 60\%}$$

Problem 2

In a bolt factory, Machines A, B, C manufacture respectively 25%,35% and 40 % of the total. Of their output 5,4,2 percent are known to be defective bolts.

A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by

(i) Machine A

(ii) Machine B or C

Step 1: Given Information

- $P(A) = 0.25$: Probability the bolt is manufactured by Machine A.
- $P(B) = 0.35$: Probability the bolt is manufactured by Machine B.
- $P(C) = 0.40$: Probability the bolt is manufactured by Machine C.

Defective probabilities:

- $P(D \mid A) = 0.05$: Probability the bolt is defective if produced by Machine A.
- $P(D \mid B) = 0.04$: Probability the bolt is defective if produced by Machine B.
- $P(D \mid C) = 0.02$: Probability the bolt is defective if produced by Machine C.

Total probability of defectiveness, $P(D)$, is needed to normalize the probabilities.

Step 2: Total Probability of a Defective Bolt

Using the **Law of Total Probability**:

$$P(D) = P(D \mid A) \cdot P(A) + P(D \mid B) \cdot P(B) + P(D \mid C) \cdot P(C)$$

Substitute the given values:

$$P(D) = (0.05 \cdot 0.25) + (0.04 \cdot 0.35) + (0.02 \cdot 0.40)$$

Calculate:

$$P(D) = 0.0125 + 0.014 + 0.008 = 0.0345$$

Step 3: Probability it was Manufactured by Machine A

Using **Bayes' Theorem**:

$$P(A \mid D) = \frac{P(D \mid A) \cdot P(A)}{P(D)}$$

Substitute the values:

$$P(A \mid D) = \frac{(0.05 \cdot 0.25)}{0.0345}$$

Calculate:

$$P(A \mid D) = \frac{0.0125}{0.0345} \approx 0.3623$$

Step 4: Probability it was Manufactured by Machine B or C

Step 4.1: For Machine B

Using Bayes' Theorem:

$$P(B \mid D) = \frac{P(D \mid B) \cdot P(B)}{P(D)}$$

Substitute the values:

$$P(B \mid D) = \frac{(0.04 \cdot 0.35)}{0.0345}$$

Calculate:

$$P(B \mid D) = \frac{0.014}{0.0345} \approx 0.4058$$

Step 4.2: For Machine C

Using Bayes' Theorem:

$$P(C \mid D) = \frac{P(D \mid C) \cdot P(C)}{P(D)}$$

Substitute the values:

$$P(C \mid D) = \frac{(0.02 \cdot 0.40)}{0.0345}$$

Calculate:

$$P(C \mid D) = \frac{0.008}{0.0345} \approx 0.2319$$

Step 4.3: Combined Probability for Machine B or C

$$P(B \text{ or } C \mid D) = P(B \mid D) + P(C \mid D)$$

Substitute the values:

$$P(B \text{ or } C \mid D) = 0.4058 + 0.2319 = 0.6377$$

Final Answers

1. Probability that the defective bolt was manufactured by **Machine A:**

$$\boxed{0.3623 \text{ or } 36.23\%}$$

2. Probability that the defective bolt was manufactured by **Machine B or C:**

$$\boxed{0.6377 \text{ or } 63.77\%}$$

Problem 3

The results of an investigation by an expert on a fire accident in a skyscraper are summarized below:

(i) P(There should have been short circuit)=0.8

(ii) P(LPG cylinder explosion)

(iii) Chance of fire accident is 30% given a short circuit and 95% given an LPG explosion.

Based on these, What do you think is the most probable causes of fire? Statistically justify your answer.

To determine the most probable cause of the fire, we will compare the probabilities of the two possible causes:

1. **Short Circuit**
2. **LPG Cylinder Explosion**

We are given:

- $P(\text{Short Circuit}) = 0.8$
- $P(\text{LPG Explosion}) = 1 - P(\text{Short Circuit}) = 0.2$ (since these are the only two possible causes).
- $P(\text{Fire} \mid \text{Short Circuit}) = 0.3$
- $P(\text{Fire} \mid \text{LPG Explosion}) = 0.95$

Step 1: Bayes' Theorem for Cause of Fire

We aim to calculate:

1. $P(\text{Short Circuit} \mid \text{Fire})$: Probability that the fire was caused by a short circuit.
2. $P(\text{LPG Explosion} \mid \text{Fire})$: Probability that the fire was caused by an LPG explosion.

Bayes' Theorem states:

$$P(\text{Cause} \mid \text{Fire}) = \frac{P(\text{Fire} \mid \text{Cause}) \cdot P(\text{Cause})}{P(\text{Fire})}$$

Where $P(\text{Fire})$ is the total probability of fire, calculated using the **Law of Total Probability:**

$$P(\text{Fire}) = P(\text{Fire} \mid \text{Short Circuit}) \cdot P(\text{Short Circuit}) + P(\text{Fire} \mid \text{LPG Explosion}) \cdot P(\text{LPG Explosion})$$

Step 2: Calculate $P(\text{Fire})$

Substitute the given values:

$$P(\text{Fire}) = (0.3 \cdot 0.8) + (0.95 \cdot 0.2)$$

Calculate each term:

$$P(\text{Fire}) = 0.24 + 0.19 = 0.43$$

Step 3: Calculate $P(\text{Short Circuit} \mid \text{Fire})$

Substitute into Bayes' Theorem:

$$P(\text{Short Circuit} \mid \text{Fire}) = \frac{P(\text{Fire} \mid \text{Short Circuit}) \cdot P(\text{Short Circuit})}{P(\text{Fire})}$$

Substitute the values:

$$P(\text{Short Circuit} \mid \text{Fire}) = \frac{0.3 \cdot 0.8}{0.43}$$

Calculate:

$$P(\text{Short Circuit} \mid \text{Fire}) = \frac{0.24}{0.43} \approx 0.5581$$

Step 4: Calculate $P(\text{LPG Explosion} \mid \text{Fire})$

Substitute into Bayes' Theorem:

$$P(\text{LPG Explosion} \mid \text{Fire}) = \frac{P(\text{Fire} \mid \text{LPG Explosion}) \cdot P(\text{LPG Explosion})}{P(\text{Fire})}$$

Substitute the values:

$$P(\text{LPG Explosion} \mid \text{Fire}) = \frac{0.95 \cdot 0.2}{0.43}$$

Calculate:

$$P(\text{LPG Explosion} \mid \text{Fire}) = \frac{0.19}{0.43} \approx 0.4419$$

Step 5: Compare the Probabilities

- $P(\text{Short Circuit} \mid \text{Fire}) \approx 0.5581$ (55.81%)
- $P(\text{LPG Explosion} \mid \text{Fire}) \approx 0.4419$ (44.19%)

The most probable cause of the fire is a **Short Circuit**, as $P(\text{Short Circuit} \mid \text{Fire}) > P(\text{LPG Explosion} \mid \text{Fire})$.

Final Answer:

The most probable cause of the fire is:

$$\boxed{\text{Short Circuit}}$$

Problem 4

The chances that an academician, a business man and a politician becoming Vice Chancellor of a university are 0.5, 0.3 and 0.2 respectively. The probability that research work will be promoted in the university by these 3 gentlemen are respectively are 0.8, 0.6 and 0.4. It is found Research work has been promoted by the university. What is the chance that an academician has become the VC?

This problem involves using **Bayes' Theorem** to calculate the probability that an academician has become the Vice Chancellor (VC) given that research work has been promoted.

Step 1: Define Events

Let:

- A_1: Event that an **Academician** becomes the VC.
- A_2: Event that a **Businessman** becomes the VC.
- A_3: Event that a **Politician** becomes the VC.
- R: Event that research work is promoted.

Given Probabilities:

- $P(A_1) = 0.5$, $P(A_2) = 0.3$, $P(A_3) = 0.2$ (prior probabilities for becoming VC).
- $P(R \mid A_1) = 0.8$, $P(R \mid A_2) = 0.6$, $P(R \mid A_3) = 0.4$ (likelihood of promoting research).

We aim to calculate:

$$P(A_1 \mid R)$$

Using **Bayes' Theorem**:

$$P(A_1 \mid R) = \frac{P(R \mid A_1) \cdot P(A_1)}{P(R)}$$

Step 2: Total Probability of R (Research Work Promoted)

Using the Law of Total Probability:

$$P(R) = P(R \mid A_1) \cdot P(A_1) + P(R \mid A_2) \cdot P(A_2) + P(R \mid A_3) \cdot P(A_3)$$

Substitute the values:

$$P(R) = (0.8 \cdot 0.5) + (0.6 \cdot 0.3) + (0.4 \cdot 0.2)$$

Calculate:

$$P(R) = 0.4 + 0.18 + 0.08 = 0.66$$

Step 3: Apply Bayes' Theorem

Using Bayes' Theorem:

$$P(A_1 \mid R) = \frac{P(R \mid A_1) \cdot P(A_1)}{P(R)}$$

Substitute the values:

$$P(A_1 \mid R) = \frac{0.8 \cdot 0.5}{0.66}$$

Calculate:

$$P(A_1 \mid R) = \frac{0.4}{0.66} \approx 0.606$$

Final Answer

The probability that an academician has become the VC given that research work is promoted is:

$$\boxed{0.606 \text{ or } 60.6\%}$$

Problem 5

Four technicians regularly make repairs when breakdowns occur on an automated production line.

- Janet, who services 20% of the breakdowns, makes an incomplete repair 1 time in 20;
- Tom, who services 60% of the breakdowns, makes an incomplete repair 1 time in 10;
- Georgia, who services 15% of the breakdowns, makes an incomplete repair 1 time in 10;
- Peter, who services 5% of the breakdowns, makes an incomplete repair 1 time in 20.

For the next problem with the production line diagnosed as being due to an initial repair that was incomplete, what is the probability that this initial repair was made by Janet?

Step 1: Define Events

Let:

- J: Janet made the repair.
- T: Tom made the repair.
- G: Georgia made the repair.
- P: Peter made the repair.
- I: The repair was incomplete.

Given Probabilities:

- $P(J) = 0.20$: Janet makes 20% of the repairs.
- $P(T) = 0.60$: Tom makes 60% of the repairs.
- $P(G) = 0.15$: Georgia makes 15% of the repairs.
- $P(P) = 0.05$: Peter makes 5% of the repairs.

Incomplete repair probabilities:

- $P(I \mid J) = \frac{1}{20} = 0.05$: Janet makes an incomplete repair 1 in 20 times.
- $P(I \mid T) = \frac{1}{10} = 0.10$: Tom makes an incomplete repair 1 in 10 times.
- $P(I \mid G) = \frac{1}{10} = 0.10$: Georgia makes an incomplete repair 1 in 10 times.
- $P(I \mid P) = \frac{1}{20} = 0.05$: Peter makes an incomplete repair 1 in 20 times.

We aim to calculate:

$$P(J \mid I)$$

Step 2: Total Probability of I (Incomplete Repair)

Using the **Law of Total Probability**:

$$P(I) = P(I \mid J) \cdot P(J) + P(I \mid T) \cdot P(T) + P(I \mid G) \cdot P(G) + P(I \mid P) \cdot P(P)$$

Substitute the given values:

$$P(I) = (0.05 \cdot 0.20) + (0.10 \cdot 0.60) + (0.10 \cdot 0.15) + (0.05 \cdot 0.05)$$

Calculate each term:

$$P(I) = 0.01 + 0.06 + 0.015 + 0.0025 = 0.0875$$

Step 3: Apply Bayes' Theorem

Using Bayes' Theorem:

$$P(J \mid I) = \frac{P(I \mid J) \cdot P(J)}{P(I)}$$

Substitute the values:

$$P(J \mid I) = \frac{0.05 \cdot 0.20}{0.0875}$$

Calculate:

$$P(J \mid I) = \frac{0.01}{0.0875} \approx 0.1143$$

Final Answer

The probability that the incomplete repair was made by **Janet** is approximately:

$$\boxed{0.1143 \text{ or } 11.43\%}$$

Explanation in Simple Terms

1. **Step 1:** Determine the probability of an incomplete repair for each technician (likelihood) and their respective share of all repairs (prior probabilities).
2. **Step 2:** Calculate the total probability of an incomplete repair considering all technicians (Law of Total Probability).
3. **Step 3:** Use Bayes' Theorem to compute the posterior probability ($P(J \mid I)$), which combines the likelihood of an incomplete repair given Janet made it and her share of all repairs.
4. The result shows that Janet is responsible for about **11.43%** of incomplete repairs.

Problem 6

Seventy percent of the light aircraft that disappear while in flight in a certain country are subsequently discovered. Of the aircraft that are discovered,60% have an emergency locator, whereas 90% of the aircraft not discovered do not have such a locator. Suppose that a light aircraft has disappeared.

(a) If it has an emergency locator, find the probability that it will not be discovered.

(b) If it doesn't have a locator, find the probability that it will be discovered.

Step 1: Define the Events

Let:

- D: Aircraft is discovered.
- $\sim D$: Aircraft is not discovered.
- L: Aircraft has an emergency locator.
- $\sim L$: Aircraft does not have an emergency locator.

Given Probabilities:

- $P(D) = 0.7$: Probability the aircraft is discovered.
- $P(\sim D) = 1 - P(D) = 0.3$: Probability the aircraft is not discovered.
- $P(L \mid D) = 0.6$: Probability the aircraft has a locator given it is discovered.
- $P(\sim L \mid\sim D) = 0.9$: Probability the aircraft does not have a locator given it is not discovered.

Step 2: Using the Complement Rule

From $P(L \mid D) = 0.6$, the probability of not having a locator given the aircraft is discovered is:

$$P(\sim L \mid D) = 1 - P(L \mid D) = 1 - 0.6 = 0.4$$

Similarly, from $P(\sim L \mid\sim D) = 0.9$, the probability of having a locator given the aircraft is not discovered is:

$$P(L \mid\sim D) = 1 - P(\sim L \mid\sim D) = 1 - 0.9 = 0.1$$

Step 3: Total Probability for $P(L)$

We calculate the overall probability that the aircraft has a locator $P(L)$ using the **Law of Total Probability**:

$$P(L) = P(L \mid D) \cdot P(D) + P(L \mid\sim D) \cdot P(\sim D)$$

Substitute the values:

$$P(L) = (0.6 \cdot 0.7) + (0.1 \cdot 0.3)$$

Calculate:

$$P(L) = 0.42 + 0.03 = 0.45$$

The probability that the aircraft has no locator ($P(\sim L)$) is:

$$P(\sim L) = 1 - P(L) = 1 - 0.45 = 0.55$$

Part (a): Probability the aircraft is not discovered given it has a locator ($P(\sim D \mid L)$)

Using **Bayes' Theorem**:

$$P(\sim D \mid L) = \frac{P(L \mid\sim D) \cdot P(\sim D)}{P(L)}$$

Substitute the values:

$$P(\sim D \mid L) = \frac{0.1 \cdot 0.3}{0.45}$$

Calculate:

$$P(\sim D \mid L) = \frac{0.03}{0.45} \approx 0.0667$$

Part (b): Probability the aircraft is discovered given it does not have a locator ($P(D \mid\sim L)$)

Using **Bayes' Theorem**:

$$P(D \mid\sim L) = \frac{P(\sim L \mid D) \cdot P(D)}{P(\sim L)}$$

Substitute the values:

$$P(D \mid\sim L) = \frac{0.4 \cdot 0.7}{0.55}$$

Calculate:

$$P(D \mid\sim L) = \frac{0.28}{0.55} \approx 0.5091$$

Final Answers

1. **Part (a):** If the aircraft has a locator, the probability that it will not be discovered is:

$$\boxed{0.0667 \text{ or } 6.67\%}$$

2. **Part (b):** If the aircraft does not have a locator, the probability that it will be discovered is:

$$\boxed{0.5091 \text{ or } 50.91\%}$$

Problem 7

Two firms V and W consider bidding on a road-building job, which may or may not be awarded depending on the amounts of the bids. Firm V submits a bid and the probability is 3/4 that it will get the job provided firm W does not bid. The probability is 3/4 that W will bid, and if it does, the probability that V will get the job is only 1/3. (a) what is the probability that V will get the job? (b) If V gets the job, what is the probability that W did not bid?

Step 1: Definitions

From the problem, we have the following probabilities:

- $P(B) = \frac{3}{4}$: Probability that W bids.
- $P(\sim B) = 1 - P(B) = \frac{1}{4}$: Probability that W does not bid.
- $P(A \mid B) = \frac{1}{3}$: Probability that V gets the job if W bids.
- $P(A \mid \sim B) = \frac{3}{4}$: Probability that V gets the job if W does not bid.

We aim to find:

1. $P(A)$: Probability that V gets the job.
2. $P(\sim B \mid A)$: Probability that W did not bid, given V gets the job.

Step 2: Total Probability for $P(A)$

Using the **Law of Total Probability**:

$$P(A) = P(A \mid B) \cdot P(B) + P(A \mid \sim B) \cdot P(\sim B)$$

Substitute the given probabilities:

$$P(A) = \left(\frac{1}{3} \cdot \frac{3}{4}\right) + \left(\frac{3}{4} \cdot \frac{1}{4}\right)$$

Calculate each term:

1. First term: $\frac{1}{3} \cdot \frac{3}{4} = \frac{3}{12} = \frac{1}{4}$
2. Second term: $\frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$

Combine:

$$P(A) = \frac{1}{4} + \frac{3}{16} = \frac{4}{16} + \frac{3}{16} = \frac{7}{16}$$

So, the probability that V gets the job is:

$$\boxed{P(A) = \frac{7}{16}}$$

Step 3: Bayes' Theorem for $P(\sim B \mid A)$

Using Bayes' Theorem:

$$P(\sim B \mid A) = \frac{P(A \mid \sim B) \cdot P(\sim B)}{P(A)}$$

Substitute the values:

$$P(\sim B \mid A) = \frac{\left(\frac{3}{4} \cdot \frac{1}{4}\right)}{\frac{7}{16}}$$

Calculate:

1. Numerator: $\frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$
2. Denominator: $P(A) = \frac{7}{16}$

Simplify:

$$P(\sim B \mid A) = \frac{\frac{3}{16}}{\frac{7}{16}} = \frac{3}{7}$$

So, the probability that W did not bid, given V got the job, is:

$$\boxed{P(\sim B \mid A) = \frac{3}{7}}$$

Final Validation

1. **Probability that V gets the job:**

$$\boxed{\frac{7}{16}}$$

   This matches the textbook answer.

2. **Probability that W did not bid, given V gets the job:**

$$\boxed{\frac{3}{7}}$$

   This also matches the textbook answer.