

Statistical Methods and Data Analysis (MA 231)

Chanchal Kundu

Department of Mathematical Sciences
Rajiv Gandhi Institute of Petroleum Technology
Jais, U.P., INDIA

Even Semester, 2024-25



Text Books

- *Probability and Statistics in Engineering* by **Hines, Montgomery, Goldsman and Borror**, John Wiley and Sons (Student Ed.).
- *An Introduction to Mathematical Statistics and its Applications* by **R. Larsen and M. Marx**, Pearson Education.



Text Books

- *Probability and Statistics in Engineering* by **Hines, Montgomery, Goldsman and Borrer**, John Wiley and Sons (Student Ed.).
- *An Introduction to Mathematical Statistics and its Applications* by **R. Larsen and M. Marx**, Pearson Education.

Supplementary Books

- *Introduction to Probability and Statistics* by **J. Susan Milton and J.C. Arnold**, Tata McGraw-Hill Pub. Co. Ltd.
- *Miller and Freund's Probability and Statistics for Engineers* by **R.A. Johnson**, Pearson Education.
- *Probability and Statistics* by **Spiegel, Schiller and Srinivasan**. Tata McGraw-Hill Pub. Co. Ltd.
- *An Introduction to Probability and Statistics* by **Rohatgi, V.K. and Saleh, A.K.**, John Wiley and Sons (Student Ed.).



Part I

Lecture: Introduction to Probability



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:

Probability

- 1 What are the 'chances' that sales will decrease if we increase prices?



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:

Probability

- 1 What are the 'chances' that sales will decrease if we increase prices?
- 2 What is the 'likelihood' that a new assembly method will increase productivity?



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:

Probability

- 1 What are the 'chances' that sales will decrease if we increase prices?
- 2 What is the 'likelihood' that a new assembly method will increase productivity?
- 3 How 'likely' is it that the project will be completed on time?



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:

Probability

- 1 What are the 'chances' that sales will decrease if we increase prices?
- 2 What is the 'likelihood' that a new assembly method will increase productivity?
- 3 How 'likely' is it that the project will be completed on time?
- 4 What are the 'odds' in favor of a new investment being profitable?



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:

Probability

- 1 What are the 'chances' that sales will decrease if we increase prices?
- 2 What is the 'likelihood' that a new assembly method will increase productivity?
- 3 How 'likely' is it that the project will be completed on time?
- 4 What are the 'odds' in favor of a new investment being profitable?



Historical Background. . .

Business decisions are often based on analysis of uncertainties such as the following:

Probability

- ① What are the 'chances' that sales will decrease if we increase prices?
- ② What is the 'likelihood' that a new assembly method will increase productivity?
- ③ How 'likely' is it that the project will be completed on time?
- ④ What are the 'odds' in favor of a new investment being profitable?

- Probability is a numerical measure of the likelihood that an event will occur.
- Measure of the degree of **uncertainty** associated with the above four events.



Introduction of Probability Concepts through Case

The Problem of a Medical Representative



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.
- 2 He needs to meet 100 more doctors in the upcoming 27 days.



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.
- 2 He needs to meet 100 more doctors in the upcoming 27 days.
- 3 Invitation from Mr. Singh, regional manager of northern region. (Lucknow).



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.
- 2 He needs to meet 100 more doctors in the upcoming 27 days.
- 3 Invitation from Mr. Singh, regional manager of northern region. (Lucknow).
- 4 Mr. Chatterjee requires 25 days to complete his target.



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.
- 2 He needs to meet 100 more doctors in the upcoming 27 days.
- 3 Invitation from Mr. Singh, regional manager of northern region. (Lucknow).
- 4 Mr. Chatterjee requires 25 days to complete his target.
- 5 He can't travel to meet the doctors due to (i) Bandh/Hartal and (ii) Heavy rains.



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.
- 2 He needs to meet 100 more doctors in the upcoming 27 days.
- 3 Invitation from Mr. Singh, regional manager of northern region. (Lucknow).
- 4 Mr. Chatterjee requires 25 days to complete his target.
- 5 He can't travel to meet the doctors due to (i) Bandh/Hartal and (ii) Heavy rains.
- 6 There were bandh/hartal for 14 days in the last two years.



Introduction of Probability Concepts through Case

The Problem of a Medical Representative

Introduction to the Case

- 1 Mr. Chatterjee, a sales representative of Ranboxy in Kolkata region.
- 2 He needs to meet 100 more doctors in the upcoming 27 days.
- 3 Invitation from Mr. Singh, regional manager of northern region. (Lucknow).
- 4 Mr. Chatterjee requires 25 days to complete his target.
- 5 He can't travel to meet the doctors due to (i) Bandh/Hartal and (ii) Heavy rains.
- 6 There were bandh/hartal for 14 days in the last two years.
- 7 There is a one in 30 chance that, on any day in this season, the roads are blocked due to flood in the city.



The Problem of Medical Representative

Question 1

What conclusion did Mr. Chatterjee arrive at?



The Problem of Medical Representative

Question 1

What conclusion did Mr. Chatterjee arrive at?

Question 2

What are the methods Mr. Chatterjee used to arrive at this conclusion?



Statement of the Problem

Mr. Chatterjee would undertake going to Lucknow, only if he is confident that he should be working for at least 25 days. This means that the number of days that could be lost due to bandh/hartal or flood should not exceed 1. In order to meet the target, Mr. Chatterjee needs to compute the probability that not more than one day is lost due to bandh/flood. If this probability is sufficiently large, he may decide to accept the invitation.



Basic Terminology in 'Probability'

1 Random Experiment.



Basic Terminology in 'Probability'

- 1 Random Experiment.
- 2 Sample Space.



Basic Terminology in 'Probability'

- 1 Random Experiment.
- 2 Sample Space.
- 3 Events and their Classifications.



Random Experiment

A random (or statistical) experiment is an experiment in which:



Random Experiment

A random (or statistical) experiment is an experiment in which:

- 1 All outcomes of the experiment is known in advance;



Random Experiment

A random (or statistical) experiment is an experiment in which:

- 1 All outcomes of the experiment is known in advance;
- 2 It is not possible to predict which outcome will occur in any particular performance of the experiment;



Random Experiment

A random (or statistical) experiment is an experiment in which:

- 1 All outcomes of the experiment is known in advance;
- 2 It is not possible to predict which outcome will occur in any particular performance of the experiment;
- 3 **The experiment can be repeated under identical conditions.**



Random Experiment

A random (or statistical) experiment is an experiment in which:

- 1 All outcomes of the experiment is known in advance;
- 2 It is not possible to predict which outcome will occur in any particular performance of the experiment;
- 3 The experiment can be repeated under identical conditions.



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .
 - Recall that a σ -field is a nonempty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set \emptyset .



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .
 - Recall that a σ -field is a nonempty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set \emptyset .

Sample Space

The sample space of a random experiment is a pair (Ω, S) , where



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .
 - Recall that a σ -field is a nonempty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set \emptyset .

Sample Space

The sample space of a random experiment is a pair (Ω, S) , where

- ① Ω is the set of all possible outcomes of the experiment.



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .
 - Recall that a σ -field is a nonempty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set \emptyset .

Sample Space

The sample space of a random experiment is a pair (Ω, S) , where

- ① Ω is the set of all possible outcomes of the experiment.
- ② S is a σ -field of subsets of Ω .



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .
 - Recall that a σ -field is a nonempty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set \emptyset .

Sample Space

The sample space of a random experiment is a pair (Ω, S) , where

- ① Ω is the set of all possible outcomes of the experiment.
- ② S is a σ -field of subsets of Ω .



Random Experiment

A random (or statistical) experiment is an experiment in which:

- ① All outcomes of the experiment is known in advance;
 - ② It is not possible to predict which outcome will occur in any particular performance of the experiment;
 - ③ The experiment can be repeated under identical conditions.
- To each such experiment we associate a set Ω , the set of all possible outcomes of the experiment. And with Ω we associate a σ -field S of subsets of Ω .
 - Recall that a σ -field is a nonempty class of subsets of Ω that is closed under the formation of countable unions and complements and contains the null set \emptyset .

Sample Space

The sample space of a random experiment is a pair (Ω, S) , where

- ① Ω is the set of all possible outcomes of the experiment.
 - ② S is a σ -field of subsets of Ω .
- The elements of Ω are called sample points and any set $A \in S$ is known as an event. Clearly, A is a collection of sample points.



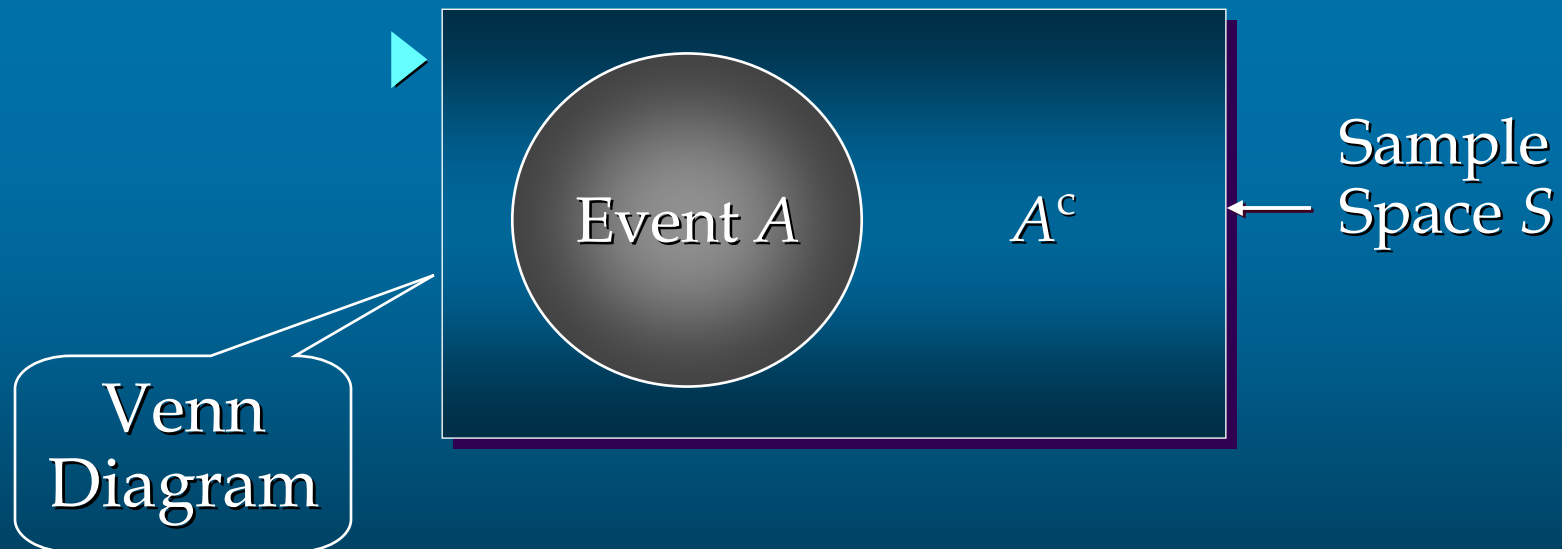
Some Basic Relationships of Probability

There are some basic probability relationships that can be used to compute the probability of an event without knowledge of all the sample point probabilities.

- ▶ Complement of an Event
- ▶ Union of Two Events
- ▶ Intersection of Two Events
- ▶ Mutually Exclusive Events

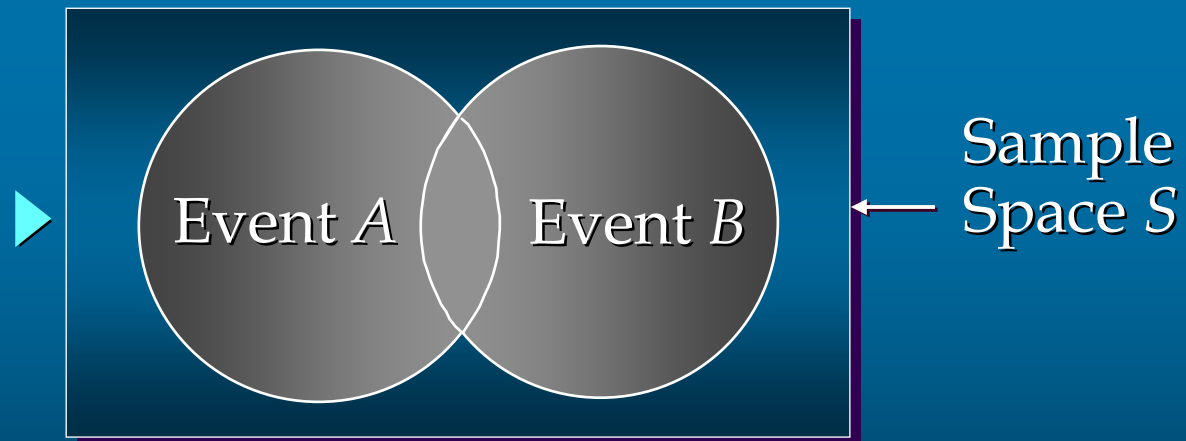
Complement of an Event

- ▶ The complement of event A is defined to be the event consisting of all sample points that are not in A .
- ▶ The complement of A is denoted by A^c .



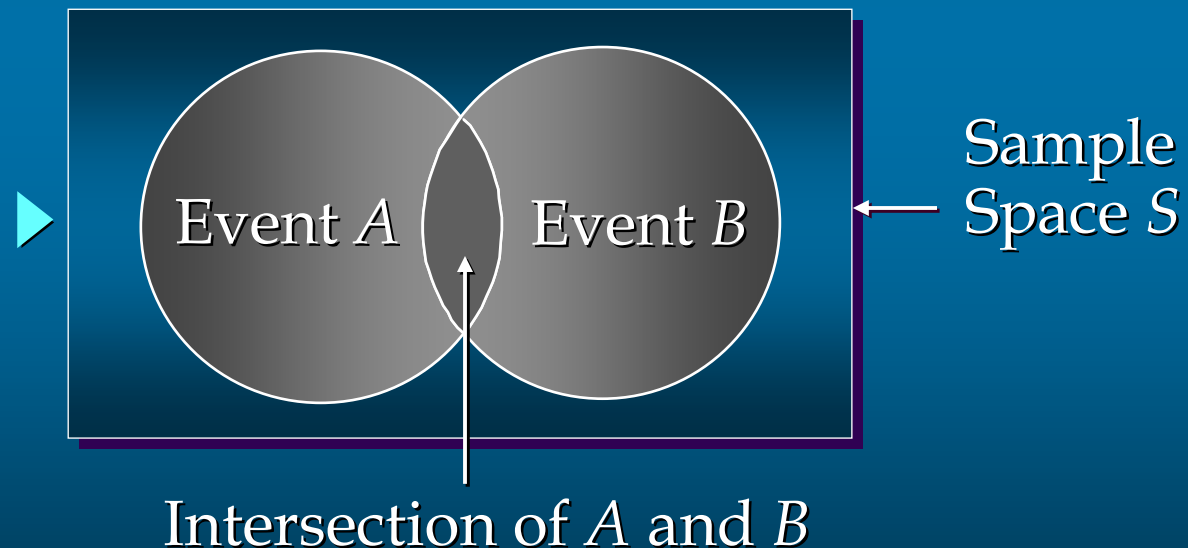
Union of Two Events

- ▶ The union of events A and B is the event containing all sample points that are in A or B or both.
- ▶ The union of events A and B is denoted by $A \cup B$.



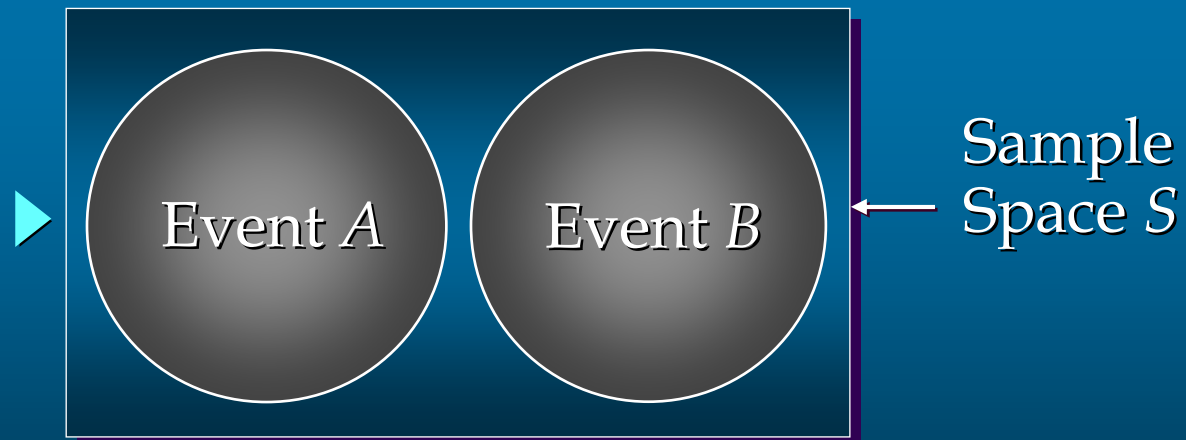
Intersection of Two Events

- ▶ The intersection of events A and B is the set of all sample points that are in both A and B .
- ▶ The intersection of events A and B is denoted by $A \cap B$.



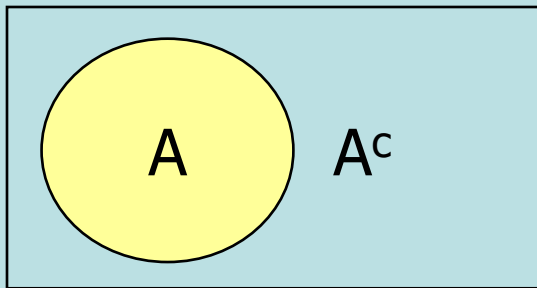
Mutually Exclusive Events

- ▶ Two events are said to be mutually exclusive if the events have no sample points in common.
- ▶ Two events are mutually exclusive if, when one event occurs, the other cannot occur.

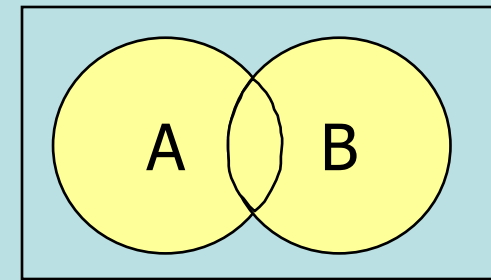


Basic Relationships of Probability...

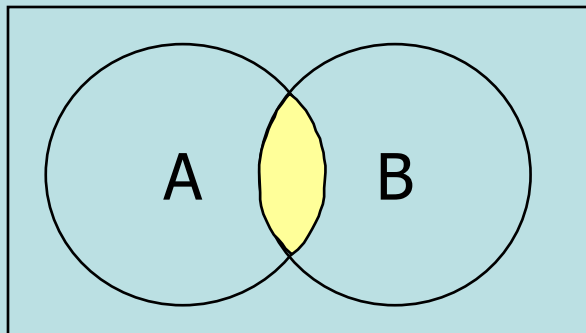
Complement of Event



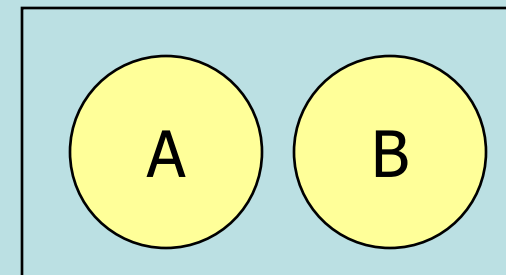
Union of Events



Intersection of Events



Mutually Exclusive Events



Assigning Probabilities

▶ Classical Method

Assigning probabilities based on the assumption of equally likely outcomes

▶ Relative Frequency Method

Assigning probabilities based on experimentation or historical data

▶ Subjective Method

Assigning probabilities based on judgment

Classical Definition

Definition

Let E be a random experiment, (Ω, S) be the sample space where Ω contains n event points all of which are known to be equally likely. If any event $A \in S$ connected with E contains $1 \leq m(A) \leq n$ of these events points, then the probability of A is defined by

$$\begin{aligned}
 P(A) &= \frac{m(A)}{n} \\
 &= \frac{\text{Number of favorable events}}{\text{Total number of events in } \Omega}
 \end{aligned}$$

Example:

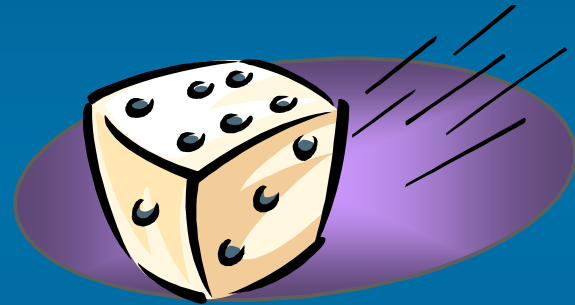


Classical Method

If an experiment has n possible outcomes, this method would assign a probability of $1/n$ to each outcome.

Example

- ▶ Experiment: Rolling a die
- ▶ Sample Space: $S = \{1, 2, 3, 4, 5, 6\}$
- ▶ Probabilities: Each sample point has a $1/6$ chance of occurring



Relative Frequency Method



- Example: Lucas Tool Rental
- ▶ Lucas Tool Rental would like to assign probabilities to the number of car polishers it rents each day. Office records show the following frequencies of daily rentals for the last 40 days.

<u>Number of Polishers Rented</u>	<u>Number of Days</u>
0	4
1	6
2	18
3	10
4	2

Relative Frequency Method

Each probability assignment is given by dividing the frequency (number of days) by the total frequency (total number of days).



<u>Number of Polishers Rented</u>	<u>Number of Days</u>	<u>Probability</u>
0	4	.10
1	6	.15
2	18	.45
3	10	.25
4	<u>2</u>	<u>.05</u>
	40	1.00

4/40

Rules for Probability

Let (Ω, \mathcal{S}) be the sample space associated with a random experiment. A set function P defined on \mathcal{S} is called a probability measure (or simply, probability) if it satisfies the following conditions:

❶ $0 \leq P(A) \leq 1$ and $P(\bar{A}) = 1 - P(A)$.



Rules for Probability

Let (Ω, \mathcal{S}) be the sample space associated with a random experiment. A set function P defined on \mathcal{S} is called a probability measure (or simply, probability) if it satisfies the following conditions:

- 1 $0 \leq P(A) \leq 1$ and $P(\bar{A}) = 1 - P(A)$.
- 2 $P(\emptyset) = 0$ and $P(\Omega) = 1$.



Rules for Probability

Let (Ω, S) be the sample space associated with a random experiment. A set function P defined on S is called a probability measure (or simply, probability) if it satisfies the following conditions:

- ① $0 \leq P(A) \leq 1$ and $P(\bar{A}) = 1 - P(A)$.
- ② $P(\emptyset) = 0$ and $P(\Omega) = 1$.
- ③ If A_1, A_2, A_3, \dots be a finite or infinite sequence of pairwise mutually exclusive events i.e., $A_i \cap A_j = \emptyset, i \neq j$ then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

$$\Rightarrow P(A_1 + A_2 + A_3 + \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$



Rules for Probability

Let (Ω, S) be the sample space associated with a random experiment. A set function P defined on S is called a probability measure (or simply, probability) if it satisfies the following conditions:

- ① $0 \leq P(A) \leq 1$ and $P(\bar{A}) = 1 - P(A)$.
- ② $P(\emptyset) = 0$ and $P(\Omega) = 1$.
- ③ If A_1, A_2, A_3, \dots be a finite or infinite sequence of pairwise mutually exclusive events i.e., $A_i \cap A_j = \emptyset, i \neq j$ then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

$$\Rightarrow P(A_1 + A_2 + A_3 + \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$



Rules for Probability

Let (Ω, S) be the sample space associated with a random experiment. A set function P defined on S is called a probability measure (or simply, probability) if it satisfies the following conditions:

- ① $0 \leq P(A) \leq 1$ and $P(\bar{A}) = 1 - P(A)$.
- ② $P(\emptyset) = 0$ and $P(\Omega) = 1$.
- ③ If A_1, A_2, A_3, \dots be a finite or infinite sequence of pairwise mutually exclusive events i.e., $A_i \cap A_j = \emptyset, i \neq j$ then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

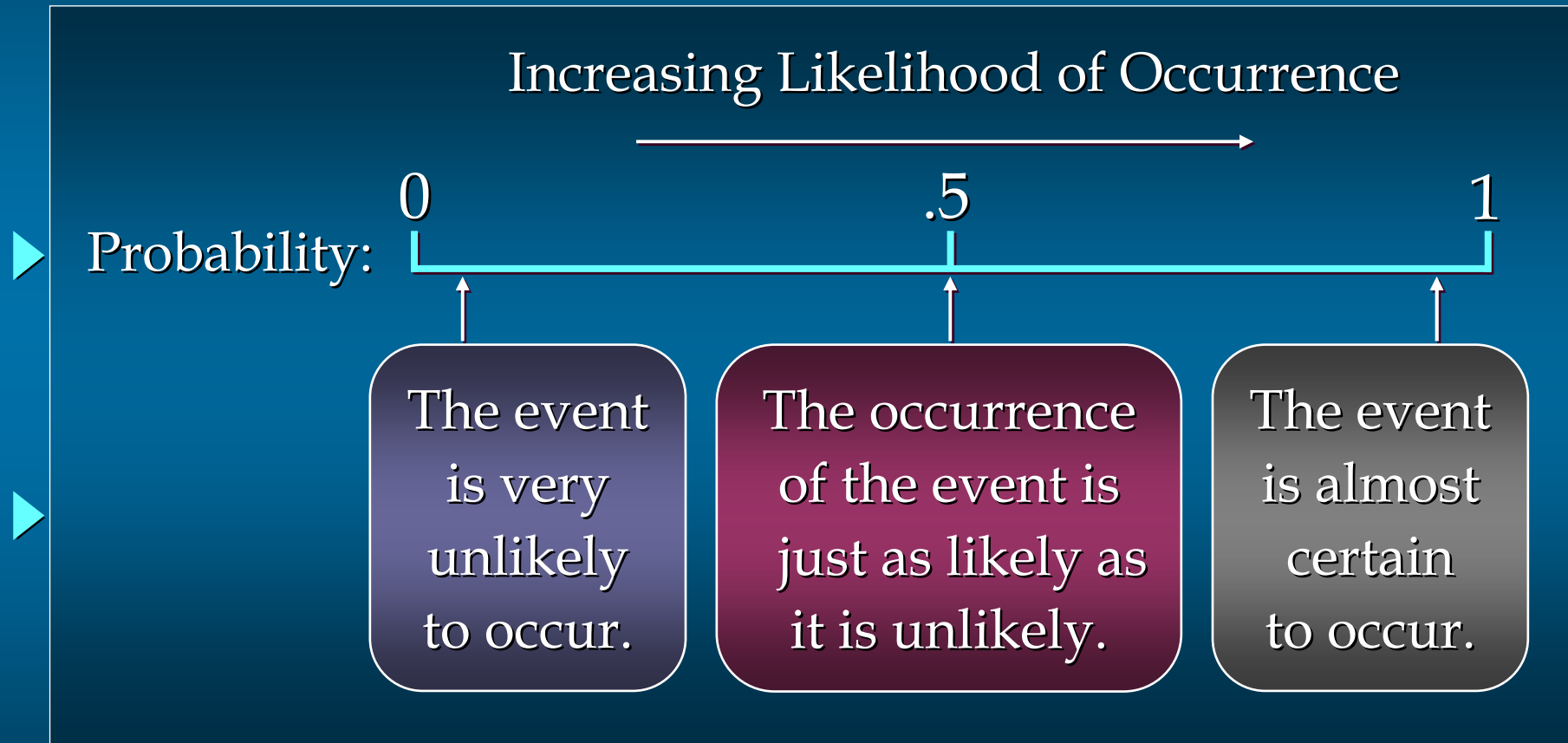
$$\Rightarrow P(A_1 + A_2 + A_3 + \dots) = P(A_1) + P(A_2) + P(A_3) + \dots$$

Definition

The triple (Ω, S, P) is called a probability space.



Probability as a Numerical Measure of the Likelihood of Occurrence



Addition Law

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$



Addition Law

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Exercise

Consider the previous year situation of IITK in Quantitative Methods for Decision Making. Of 200 students taking the course, 160 passed the mid term, 140 end term; 124 both the exams. After reviewing the grades, the instructor decided to give a passing grade to any student who passed at least one of the two exams. What is the probability of a student receiving a passing grade in this course?



Example

Why are some mutual fund managers more successful than others? One possible factor is where the manager earned his or her MBA. The following table compares mutual fund performance against the ranking of the school where the fund manager earned their MBA:

	Mutual fund outperforms the market	Mutual fund doesn't outperform the market
Top 20 MBA program	.11	.29
Not top 20 MBA program	.06	.54

E.g. This is the probability that a mutual fund outperforms **AND** the manager was in a top-20 MBA program; it's a **joint probability**.

Continued...

Alternatively, we could introduce shorthand notation to represent the events:

A_1 = Fund manager graduated from a top-20 MBA program

A_2 = Fund manager did not graduate from a top-20 MBA program

B_1 = Fund outperforms the market

B_2 = Fund does not outperform the market

	B_1	B_2
A_1	.11	.29
A_2	.06	.54

E.g. $P(A_2 \text{ and } B_1) = .06$

= the probability a fund outperforms the market

and the manager isn't from a top-20 school.

Marginal Probabilities...

Marginal probabilities are computed by adding across rows and down columns; that is they are calculated in the *margins* of the table:

$$P(A_2) = .06 + .54$$

“what’s the probability a fund manager isn’t from a top school?”

	B ₁	B ₂	P(A _i)
A ₁	.11	.29	.40
A ₂	.06	.54	.60
P(B _j)	.17	.83	1.00

$$P(B_1) = .11 + .06$$

“what’s the probability a fund outperforms the market?”

BOTH margins must add to 1
(useful error check)

Conditional Probability...

Conditional probability is used to determine how two events are related; that is, we can determine the probability of one event *given* the occurrence of another related event.

Conditional probabilities are written as $P(A | B)$ and read as “the probability of *A given B*” and is calculated as:

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

Conditional Probability...

Again, the probability of an event *given* that another event has occurred is called a conditional probability...

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

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

Note how “A given B” and “B given A” are related...

Conditional Probability...

In previous Example: What's the probability that a fund will outperform the market *given* that the manager graduated from a top-20 MBA program?

Recall:

A_1 = Fund manager graduated from a top-20 MBA program

A_2 = Fund manager did not graduate from a top-20 MBA program

B_1 = Fund outperforms the market

B_2 = Fund does not outperform the market

Thus, we want to know “what is $P(B_1 | A_1)$?”

Conditional Probability...

We want to calculate $P(B_1 | A_1)$

	B ₁	B ₂	P(A _i)
A ₁	.11	.29	.40
A ₂	.06	.54	.60
P(B _j)	.17	.83	1.00

$$P(B_1 | A_1) = \frac{P(A_1 \text{ and } B_1)}{P(A_1)} = \frac{.11}{.40} = .275$$

Thus, there is a 27.5% chance that that a fund will outperform the market given that the manager graduated from a top-20 MBA program.

Independence...

One of the objectives of calculating conditional probability is to determine whether two events are related.

In particular, we would like to know whether they are *independent*, that is, if the probability of one event is *not affected* by the occurrence of the other event.

Two events A and B are said to be *independent* if

$$P(A|B) = P(A)$$

or

$$P(B|A) = P(B)$$

Independence...

For example, we saw that

$$P(B_1 | A_1) = .275$$

The marginal probability for B_1 is: $P(B_1) = 0.17$

Since $P(B_1|A_1) \neq P(B_1)$, B_1 and A_1 are *not independent* events.

Stated another way, they are *dependent*. That is, the probability of one event (B_1) *is affected* by the occurrence of the other event (A_1).

Conditional Probability

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

provided $P(A) \neq 0$.



Conditional Probability

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

provided $P(A) \neq 0$.

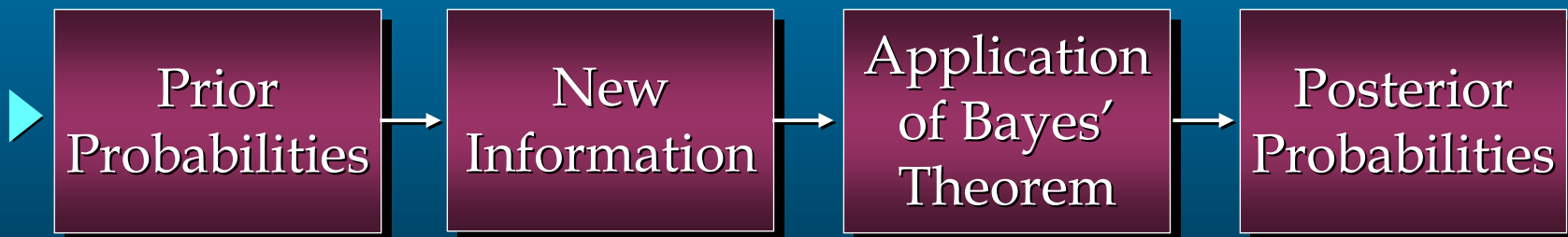
Exercise

For two events A and B ,
 $P(A) = 0.5$, $P(B) = 0.6$, $P(A \cap B) = 0.4$. Find
 $P(A|B)$, $P(B|A)$. Are A and B independent?



Bayes' Theorem

- ▶ ■ Often we begin probability analysis with initial or prior probabilities.
- ▶ ■ Then, from a sample, special report, or a product test we obtain some additional information.
- ▶ ■ Given this information, we calculate revised or posterior probabilities.
- ▶ ■ Bayes' theorem provides the means for revising the prior probabilities.



Bayes' Theorem

■ Example: L. S. Clothiers

▶ A proposed shopping center will provide strong competition for downtown businesses like L. S. Clothiers. If the shopping center is built, the owner of L. S. Clothiers feels it would be best to relocate to the center.

▶ The shopping center cannot be built unless a zoning change is approved by the town council. The planning board must first make a recommendation, for or against the zoning change, to the council.



Bayes' Theorem



■ Prior Probabilities

Let:

A_1 = town council approves the zoning change

A_2 = town council disapproves the change

Using subjective judgment:

$$P(A_1) = .7, \quad P(A_2) = .3$$

Bayes' Theorem



■ New Information

The planning board has recommended against the zoning change. Let B denote the event of a negative recommendation by the planning board.

Given that B has occurred, should L. S. Clothiers revise the probabilities that the town council will approve or disapprove the zoning change?

Bayes' Theorem



■ Conditional Probabilities

Past history with the planning board and the town council indicates the following:

$$P(B | A_1) = .2$$

$$P(B | A_2) = .9$$

► Hence:

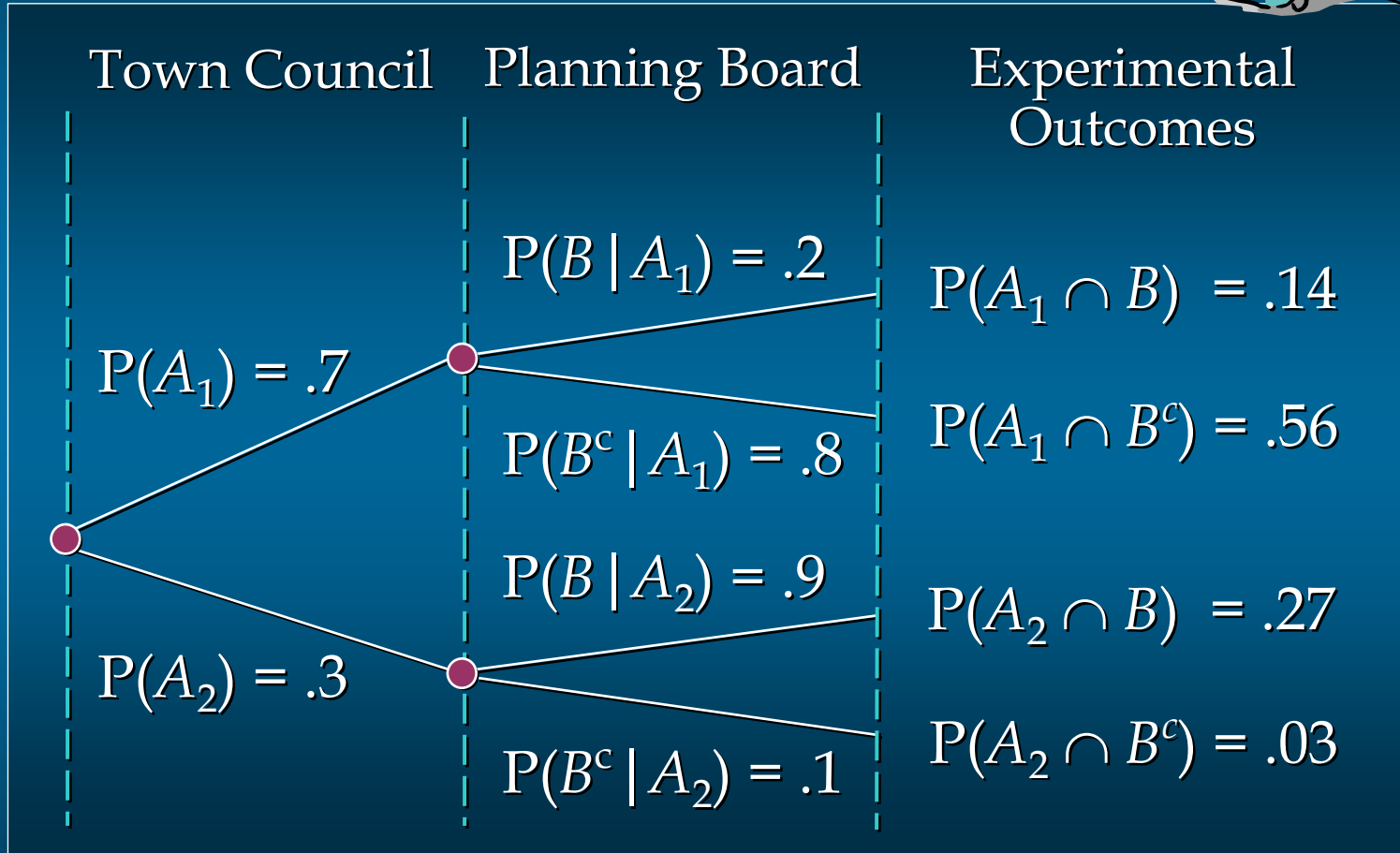
$$P(B^C | A_1) = .8$$

$$P(B^C | A_2) = .1$$

Bayes' Theorem



Tree Diagram



Bayes' Theorem

- ▶ ■ To find the posterior probability that event A_i will occur given that event B has occurred, we apply Bayes' theorem.

- ▶
$$P(A_i | B) = \frac{P(A_i)P(B | A_i)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2) + \dots + P(A_n)P(B | A_n)}$$

- ▶ ■ Bayes' theorem is applicable when the events for which we want to compute posterior probabilities are mutually exclusive and their union is the entire sample space.

Bayes' Theorem



■ Posterior Probabilities

Given the planning board's recommendation not to approve the zoning change, we revise the prior probabilities as follows:

$$\begin{aligned} \text{▶ } P(A_1 | B) &= \frac{P(A_1)P(B | A_1)}{P(A_1)P(B | A_1) + P(A_2)P(B | A_2)} \\ \text{▶ } &= \frac{(.7)(.2)}{(.7)(.2) + (.3)(.9)} \\ &= .34 \end{aligned}$$

Bayes' Theorem



■ Conclusion

The planning board's recommendation is good news for L. S. Clothiers. The posterior probability of the town council approving the zoning change is .34 compared to a prior probability of .70.

Bayes' Theorem

If $A_1, A_2, A_3, \dots, A_n$ be a given set of n pairwise mutually exclusive events, one of which certainly occurs, then for any arbitrary event X

$$P(X) = P(A_1)P(X|A_1) + P(A_2)P(X|A_2) + \dots + P(A_n)P(X|A_n)$$

and if $P(X) \neq 0$

$$P(A_i|X) = \frac{P(A_i)P(X|A_i)}{P(A_1)P(X|A_1) + P(A_2)P(X|A_2) + \dots + P(A_n)P(X|A_n)}$$



Tabular Approach

■ Step 1

Prepare the following three columns:

Column 1 – The mutually exclusive events for which posterior probabilities are desired.

Column 2 – The prior probabilities for the events.

Column 3 – The conditional probabilities of the new information *given* each event.

Tabular Approach



(1) Events A_i	(2) Prior Probabilities $P(A_i)$	(3) Conditional Probabilities $P(B A_i)$	(4)	(5)
A_1	.7	.2		
A_2	<u>.3</u>	.9		
	1.0			

Tabular Approach

■ Step 2

Column 4

Compute the joint probabilities for each event and the new information B by using the multiplication law.

Multiply the prior probabilities in column 2 by the corresponding conditional probabilities in column 3. That is, $P(A_i \cap B) = P(A_i) P(B | A_i)$.

Tabular Approach



(1)	(2)	(3)	(4)	(5)
Events	Prior Probabilities	Conditional Probabilities	Joint Probabilities	
A_i	$P(A_i)$	$P(B A_i)$	$P(A_i \cap B)$	
A_1	.7	.2	.14	
A_2	<u>.3</u>	.9	<u>.27</u>	
	1.0			

.7 x .2

Tabular Approach



■ Step 2 (continued)

We see that there is a .14 probability of the town council approving the zoning change and a negative recommendation by the planning board.

There is a .27 probability of the town council disapproving the zoning change and a negative recommendation by the planning board.

Tabular Approach



■ Step 3

Column 4

Sum the joint probabilities. The sum is the probability of the new information, $P(B)$. The sum $.14 + .27$ shows an overall probability of $.41$ of a negative recommendation by the planning board.

Tabular Approach



(1)	(2)	(3)	(4)	(5)
Events	Prior Probabilities	Conditional Probabilities	Joint Probabilities	
A_i	$P(A_i)$	$P(B A_i)$	$P(A_i \cap B)$	
A_1	.7	.2	.14	
A_2	<u>.3</u>	.9	<u>.27</u>	
	1.0		$P(B) = .41$	



Tabular Approach

- Step 4

- Column 5

Compute the posterior probabilities using the basic relationship of conditional probability.

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

The joint probabilities $P(A_i \cap B)$ are in column 4 and the probability $P(B)$ is the sum of column 4.

Tabular Approach



(1) Events A_i	(2) Prior Probabilities $P(A_i)$	(3) Conditional Probabilities $P(B A_i)$	(4) Joint Probabilities $P(A_i \cap B)$	(5) Posterior Probabilities $P(A_i B)$
A_1	.7	.2	.14	.3415
A_2	<u>.3</u>	.9	<u>.27</u>	<u>.6585</u>
	1.0		$P(B) = .41$	1.0000

.14 / .41



Exercise

The Wayne Manufacturing Company purchases a certain part from suppliers A_1 , A_2 and A_3 . They supply 20%, 50% and 30% respectively. The quality of the purchased parts varies with the sources having 5%, 4% and 3% bad, respectively. The parts are used in one of the company's major products.

- What is the probability that the company's major product is assembled with a bad part?
- If a bad part is found, which supplier is the likely source?



Solution

The prior probabilities of events A_1 , A_2 and A_3 are $P(A_1) = 0.20$, $P(A_2) = 0.50$, $P(A_3) = 0.30$. The conditional probability of event B given A_1 , A_2 , A_3 are $P(B|A_1) = 0.05$, $P(B|A_2) = 0.04$, $P(B|A_3) = 0.03$. Compute $P(B \cap A_1)$, $P(B \cap A_2)$, $P(B \cap A_3)$. Apply Bayes' theorem to compute the posterior probabilities $P(A_1|B)$, $P(A_2|B)$ and $P(A_3|B)$.

