CHAPTER

Duxbury Thomson Learning

Making Hard Decision

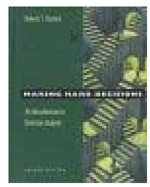
Third Edition

Probability Basics

A. J. Clark School of Engineering •Department of Civil and Environmental Engineering

7b

FALL 2003



By
Dr . Ibrahim. Assakkaf

ENCE 627 – Decision Analysis for Engineering

Department of Civil and Environmental Engineering

University of Maryland, College Park

Introduction

- The central principles of decision analysis is that we can represent uncertainty of any kind through the appropriate use of probability.
- Be able to create and analyze a model that represents the uncertainty faced in a decision.

Introduction

- The nature of the model created naturally depends on the nature of the uncertainty faced and the analysis required depends on the exigencies of the decision situation.
- The term <u>chance event</u> to refer to something about which a decision maker is uncertain. A chance event has more than one possible <u>outcome</u>. When we talk about probabilities, we are concerned with the chances associated with the different possible outcomes.

Introduction

- Objectives of Studying Probability:
  1. To become reasonably comfortable with probability concepts,
  2. To become comfortable in the use of probability to model simple uncertain situations,
  3. To be able to interpret probability statements in terms of the uncertainty that they represent, and
  4. To be able to manipulate and analyze the models you create.

Introduction

- Thus, the fundamental mathematical tools of probability theory can be used to
  - identify all possible outcomes for a specific problem, and
  - define events in the context of all these possibilities
- The basic mathematical tools of probability theory is the *set theory.*

Introduction

- Sets constitute a fundamental concept in probabilistic analysis of engineering systems.
- Establishment of a proper model and obtaining realistic results require the definition of the underlying sets.
- The objective herein is to provide the needed set foundation for probabilistic analysis

Sample Spaces, Sets, and Events

- Sets
  - Definition:

    "A set can be defined as a collection of objects, called elements or components"
  - Capital letters are usually used to denote sets.

    e.g., *A*, *B*, *X*, and *Y*
  - Small letters are used to denote their elements

    e.g., *a*, *b*, *x*, and *y*

Sample Spaces, Sets, and Events

- Sets

$a \in C$ means *a* belongs to *C*

$a \notin C$ means *a* does not belong to *C*

$a, b \in C$ means both *a* and *b* belong to *C*

Sample Spaces, Sets, and Events

- <u>Examples:</u> Sets

  A = {2, 4, 6, 8, 10}

  B = {b:b>0}; where ':' means "such that"

  C = {Maryland, Virginia, Washington}

  D = {P, M, 2, 7, U, E}

  F = {1, 3, 5, 7, 9, 11,…}; the set of odd numbers

Sample Spaces, Sets, and Events

– In set A, 2 belongs to A, but 14 does not belong to A

  • Mathematically

  $2 \in A$ means 2 belongs to A

  $14 \notin A$ means 14 does not belong to A

– Sets can be classified as *<u>finite</u>* and *<u>infinite</u>* sets

  A, C, and D are <u>finite sets</u>

  B and F are <u>infinite sets</u>

Sample Spaces, Sets, and Events

– The element of a set can be either *discrete* or *continuous*.

Elements in sets *A*, *C*, *D*, and *F* are discrete

Elements in set *B* are continuous

– A set without any element is called a null (or empty) set and is denoted as ϕ.

Sample Spaces, Sets, and Events

- Subsets

$A \subset B$	means *A* is a subset of *B*
$A = B$	means *A* and *B* have exactly the same elements
$A \not\subset B$	means *A* is not a subset of *B*
$A \neq B$	means *A* and *B* do not have exactly the same elements

NOTE: the null set ϕ is considered a subset of every set

Sample Spaces, Sets, and Events

- <u>Examples:</u> Subsets

  A_1 = {2, 4} is a subset of A = {2, 4, 6, 8, 10}

  B_1 = {b:7<b≤200} is a subset of B = {b:b>0}

  F = {1,2,3,4,5} is a subset of F = {1,2,3,4,5}

Sample Spaces, Sets, and Events

- Sample Spaces and Events
  - The set of all possible outcomes of random experiment (system) is called a <u>*sample space*</u> and is presented by the symbol *S.*
  - A subset of the sample space *S* is called an <u>*event*</u>.
  - An event without sample points is an empty set, and is called the impossible event ϕ.
  - A set that contains all the sample points is called the certain event (or sample space) *S*.

Sample Spaces, Sets, and Events

- Examples: Sample Spaces

  A = {number of cars waiting (queuing) for a left turn at specified traffic light}

  B = {number of units produced by an assembly line}

  C = {the strength of concrete delivered at a construction site}

  D = {the deformation of a structure under extreme load conditions}

Sample Spaces, Sets, and Events

- Examples: Events

  A_1 = {number of cars waiting (queuing) for a left turn at specified traffic light between 3:30 p.m. and 6:30 p.m. on a working day}

  D_1 = {failure of structure}

Sample Spaces, Sets, and Events

- Example: Sample Spaces and Events

  Roll of a Pair of Dice:

  Consider an experiment of rolling two dice. A convenient sample space that will enable us to answer many questions about events in the following figure.

  What is the event (subset of sample space *S*) that correspond to each of the following outcomes?

  (a) A sum of 7 turns up (b) A sum of 11 turns up

  (c) A sum less than 4 turns up

  (d) A sum of 12 turns up

Sample Spaces, Sets, and Events

Roll of a Pair of Dice

SECOND DIE

FIRST DIE

(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

Sample Spaces, Sets, and Events

- Example: (cont'd)
  - Roll of a Pair of Dice

(a) E_1 = {(6,1), (5,2), (4,3), (3,4), (2,5), (1,6)}

(b) E_2 = {(6,5), (5,6)}

(c) E_3 = {(1,1), (2,1), (1,2)}

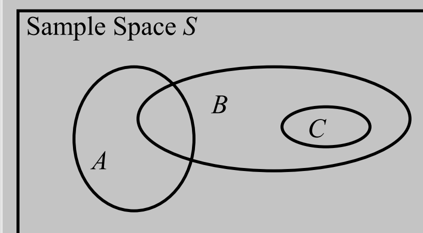
(d) E_4 = {(6,6)}

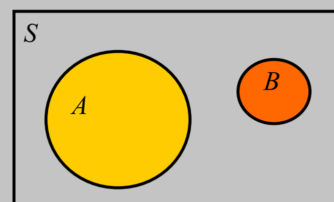
Venn Diagrams

- Events and sets can be presented using spaces that are bounded by closed shapes, such as circles.
- These shapes are called Venn-Euler (or simply Venn) diagrams.
- Belonging, non-belonging, and overlaps between events and sets can be presented by these diagrams.

Venn Diagrams

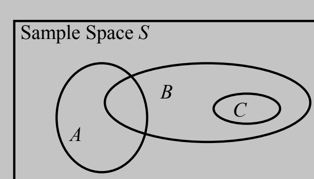
- Events in Venn Diagram





Venn Diagrams

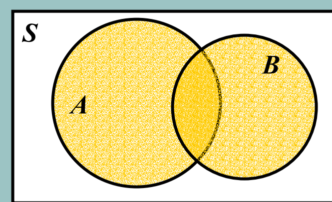
- A sample space *S*, and events *A*, *B*, and *C* are shown in the figure.
- The event *C* is contained in *B* (i.e., $C \subset B$)
- *A* is not equal to *B* (i.e., $A \neq B$).
- The events *A* and *B* have an overlap in the sample space *S*.



Basic Operations

- The ***union*** of *A* and *B* which is denoted as

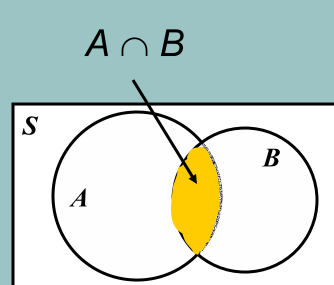
  $A \cup B$ is the set of all elements that belong to *A* or *B* or both.



- Two or more events are called collectively ***exhaustive*** events if the union of these events results in the sample space.

Basic Operations

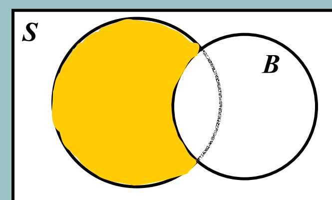
- The ***intersection*** of *A* and *B*, which is denoted as $A \cap B$, is the set of all elements that belong to both *A* and *B*.



- Events are termed ***mutually exclusive*** if the occurrence of one event precludes the occurrence of the other events

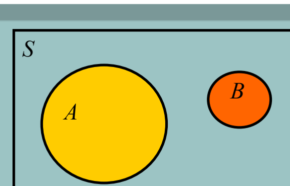
Basic Operations

- The ***difference*** of events *A* and *B*, which is designated *A* – *B*, is the set of all elements that belong to *A* but not to *B.*

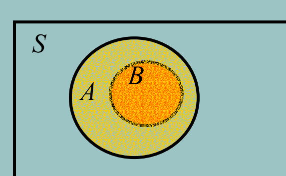


Basic Operations

- If $A \cap B = \varnothing$, then the sets *A* and *B* are said to be disjoint (mutually exclusive).

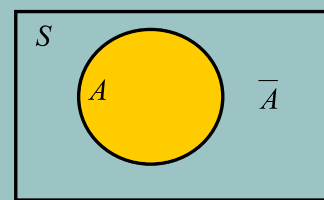


- If $B \subset A$, then $A - B$ is called the complement of *B* relative to *A* and is denoted by B_A .



Basic Operations

- The event that contains all of the elements that do not belong to an event A is called the **complement** of A, and is denoted by $\overline{A}$



Basic Operations

- Example: Operations on Sets and Events
  - The following are example sets:

    A = {2, 4, 6, 8, 10}

    B = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}

    C = {1, 3, 7, 9, 11, …}; the set of odd numbers

    F_1 = {failure of a structure due to earthquake}

    F_2 = {failure of a structure due to strong winds}

    F_3 = {failure of a structure due to an extreme overload }

Basic Operations

- Example: Operations on Sets and Events
  - The following operations can be executed for the previous example sets:

    $A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

    $A \cap B = \{2, 4, 6, 8, 10\}$

    $\overline{A} = \{2, 4, 8, 10, 12, \ldots\}$; the set of even numbers

    $F_1 \cup F_2$ = {failure of the structure due to an earthquake or strong wind}

    $\overline{F_3}$ ={non-failure of the structure due to an extreme overload}

Venn Diagram and Basic Operations

- Example: Venn Diagram and Basic Operations
  - A city has two daily newspapers, the Wildcat and the Journal. The following information was obtained from a survey of 100 residents of the city. 35 people subscribe to Wildcat, 60 subscribe to the Journal, 20 subscribe to both papers

Venn Diagram and Basic Operations

1. How many people in the survey subscribe to the Wildcat but not to the journal?
2. How many subscribe to the Journal but not to the Wildcat?
3. How many do not subscribe to either paper?
4. Organize this information in a table.

Venn Diagram and Basic Operations

- Solution
  - Let *S* be the group of people surveyed.
  - Let *W* be the set of people who subscribe to the Wildcat, and
  - Let *J* be the set of people who subscribe to the Journal

Venn Diagram and Basic Operations

- Solution (cont'd)
  - Hence,
    - $\overline{W}$ the set of people in the survey group *S* who do not subscribe to the Wildcat.
    - $\overline{J}$ the set of people who do not subscribe to Journal.

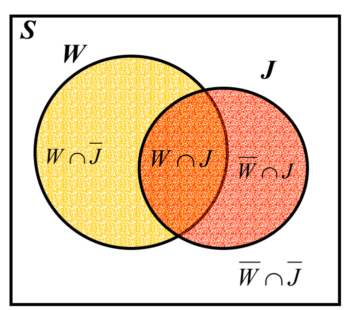
Venn Diagram and Basic Operations

$W \cap J$ = Set of people who subscribe to both paper

$W \cap \overline{J}$ = Set of people who subscribe to Wildcat but not to the Journal

$\overline{W} \cap J$ = Set of people who subscribe to the Journal but not the Wildcat

$\overline{W} \cap \overline{J}$ = Set of people who do not subscribe to either paper



Venn Diagram and Basic Operations

– Solution (cont'd)

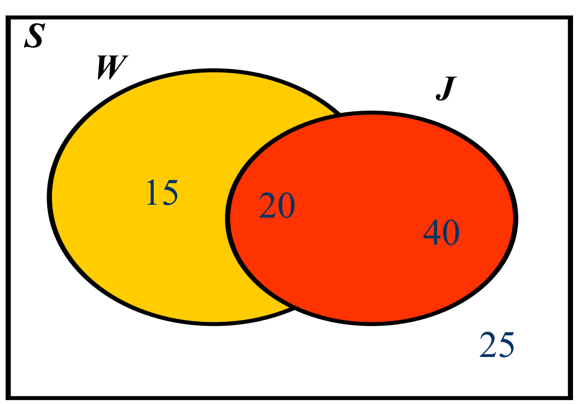
The given information can expressed in the terms of set notation as

$n(S) = 100$, $n(W) = 35$, $n(J) = 60$

$n(W \cap J) = 20$

This information with a Venn diagram can be used to answer the questions. To begin, we place 20 in $W \cap J$ in the diagram

Venn Diagram and Basic Operations



Venn Diagram and Basic Operations

- Solution (cont'd)

1. The number of people subscribe to the Wildcat but not to the Journal is

$$n(W \cap \overline{J}) = 35 - 20 = 15$$

2. The number of people who subscribe to Journal but not to the Wildcat is

$$n(\overline{W} \cap J) = 60 - 20 = 40$$

Venn Diagram and Basic Operations

- Solution (cont'd)

3. The number of people who do not subscribe to either paper is

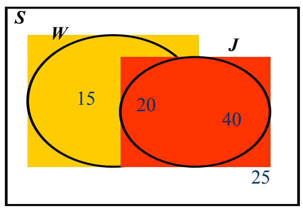
$$n(\overline{W} \cap \overline{J}) = 100 - 15 - 20 - 40 = 25$$

4. The following table contains the same information as in the Venn diagram figure, but organized in a different format:

Venn Diagram and Basic Operations

- Solution (cont'd)

		Journal		
		Subscriber, J	Subscriber, $\overline{J}$	Totals
Wildcat	Subscriber, W	20	15	35
	Nonsubscriber, $\overline{W}$	40	25	65
	Totals	60	40	100



Additional Operational Rules

Rule Type	Operations
Identity Laws	$A \cup \varnothing = A,\ A \cap \varnothing = \varnothing,\ A \cup S = S,\ A \cap S = A$
Idem potent Laws	$A \cup A = A,\ A \cap A = A$
Complement Laws	$A \cup \overline{A} = S,\ A \cap \overline{A} = \varnothing,\ \overline{\overline{A}} = A,\ \overline{S} = \varnothing,\ \overline{\varnothing} = S$
Commutative Laws	$A \cup B = B \cup A,\ A \cap B = B \cap A$
Associative Laws	$(A \cup B) \cup C = A \cup (B \cup C),\ (A \cap B) \cap C = A \cap (B \cap C)$
Distributive Laws	$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
De Morgan's Law	$\overline{(A \cup B)} = \overline{A} \cap \overline{B},\ \overline{(E_1 \cup E_2 ... \cup E_n)} = \overline{E_1} \cap \overline{E_2} ... \cap \overline{E_n}$ $\overline{(A \cap B)} = \overline{A} \cup \overline{B},\ \overline{(E_1 \cap E_2 \cap ... \cap E_n)} = \overline{E_1} \cup \overline{E_2} \cup ... \cup \overline{E_n}$
Combinations of Laws	$\overline{(A \cup (B \cap C))} = \overline{A} \cap \overline{(B \cap C)} = (\overline{A} \cap \overline{B}) \cup (\overline{A} \cap \overline{C})$

Definition of Probability

- Relative Frequency
  - In an experiment (or system) that can be repeated *N* times with *n* occurrences of an event of interest, the relative frequency of occurrence can be considered as the *probability of occurrence*.
  - In this case, the probability of occurrence is

$$\mathrm{P}(X = x_0) = \frac{n}{N}$$

Definition of Probability

- Example: Product Reliability
  - A factory produces a product. A sample of size *N* was taken from a production line. The number of non-defective products was determined to be *n*.

$$\text{Probability (non-defective)} = \frac{n}{N}$$

A Little Probability Theory

- Probabilities must satisfy the following three main requirements:
  1. Probabilities Must Lie Between 0 and 1.
  2. Probabilities Must Add Up.
  3. Total Probability Must Equal 1.

Probability Requirements — Rule 1

1 - Probabilities Must Lie Between 0 and 1

Every probability (p) must be positive, and between 0 and 1, inclusive ($0 < p < 1$). This is a sensible requirement. In informal terms it simply means nothing can have more than 100% chance of occurring or less than a 0% chance.

Rule 1: Every probability must be between 0 and 1 (inclusive).

1 ~ 100% chance of occurring

.

.

.

0 ~ 0% chance of occurring

Probability Requirements — Rule 2

2 - Probabilities Must Add Up

Suppose several outcomes are mutually exclusive (only one can happen, not both). The probability that one or the other occurs is then the sum of the individual probabilities.

Rule 2: Suppose two outcomes A_1, A_2 are mutually exclusive. Then the probability of either A_1 or A_2 occurring is:

$$P\,(A_1 \text{ or } A_2) = P\,(A_1) + P\,(A_2)$$

Probability Requirements-Rule 2(cont'd)

- Example 1: Stock Market
  - Consider the stock market. Suppose there is a 30% chance that the market will go up and a 45% chance that it will stay the same (as measured by the Dow Jones average).
  - It cannot do both at once, and so the probability that it will either go up or stay the same must be 75%.

Probability Requirements-Rule 2(cont'd)

- Example 2: Coin flipping, mutually exclusive, collectively exhaustive

A_1 = "heads" $P(A_1)$=0.5 for a fair coin

A_2 = "tails" $P(A_2)$=0.5 for a fair coin

Then,

union

$P(A_1 \text{ or } A_2) = 1$ i.e $P(A_1 \cup A_2) = 1$

$P(A_1) + P(A_2) = 0.5 + 0.5$, so

$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2) = 1.0$

Probability Requirements —Rule 3

3 - Total Probability Must Equal 1

- Suppose a set of outcomes is <u>mutually exclusive and collectively exhaustive</u>. This means that one (and only one) of the possible outcomes must occur. The probabilities for this set must sum to 1.

  Informally, if we have a set of outcomes such that one of them has to occur, then there is a 100% chance that one of them will indeed come to pass.

Rule 3: Total Probability Must Equal 1

If a set of outcome $A_1, A_2, \ldots, A_n$ is mutually exclusive and collectively exhaustive, then

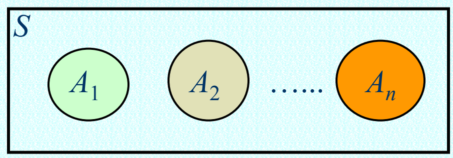
$P(A_1 \text{ or } A_2 \text{ or } \ldots \text{ or } A_n)$

$= P(A_1) + P(A_2) + \ldots + P(A_n) = 1$

Useful Formulas

- Formula I
  - If $A1$, $A2$,…….An are *mutually exclusive* events on the sample space S, then

$$\boxed{\mathrm{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \ldots + \mathrm{P}(A_n)}$$

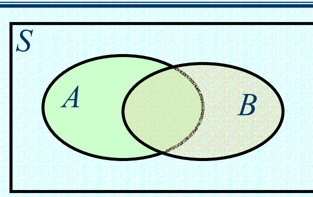


Useful Formulas

- Formula II

  If events A and B are *not mutually exclusive* events on the sample space S, then

$$\boxed{\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)}$$



Useful Formulas

- Computational Rules

Additional computational rules can be developed based on the previous axioms. The following are examples rules:

$$\mathrm{P}(A \cup B \cup C) = \mathrm{P}(A) + \mathrm{P}(B) + \mathrm{P}(C) - \mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C)$$

$$\mathrm{P}(\overline{\mathrm{A}}) = 1 - \mathrm{P}(A)$$

If $A \subseteq B$, then $\mathrm{P}(A) \leq \mathrm{P}(B)$

Useful Formulas

- <u>Example:</u> Union and Intersection

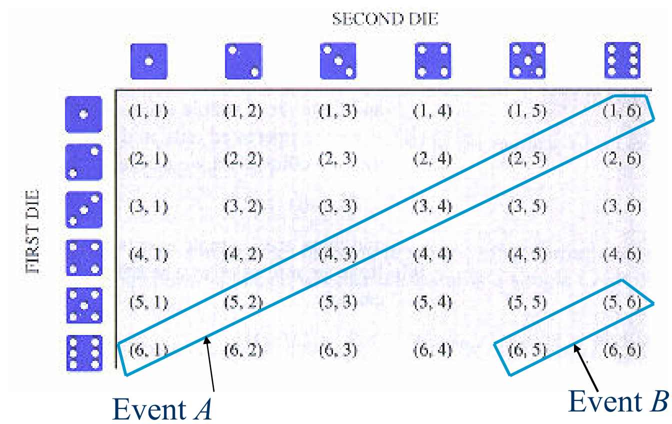
Suppose a pair of dice are rolled:

a. What is the probability that a sum of 7 or 11 turns up?

b. What is the probability that both dice turn up the same or that a sum less than 5 turns up?

Useful Formulas

Example (cont'd): Part a Solution



Useful Formulas

- Example (cont'd): Part a solution

  Let

  A = event that sum of 7 turns up

  B = event that sum of 11turns up

  Then

  $A \cup B$ = the event that sum of 7 or 11 turns up

Useful Formulas

- <u>Example(cont'd):</u> Part a solution

  A = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

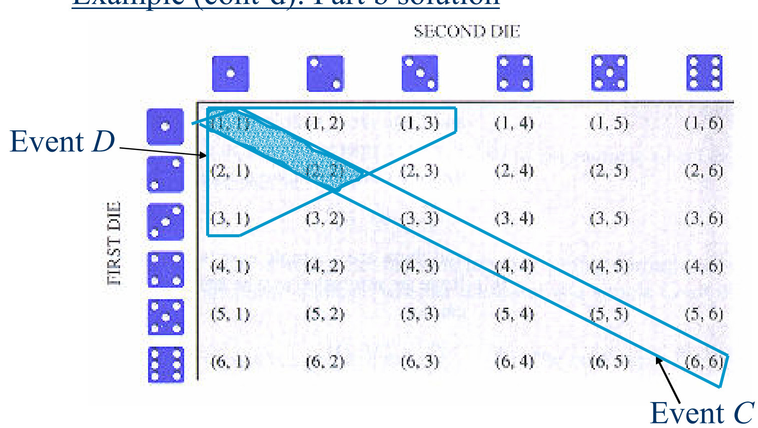
  B = {(5,6), (6,5)}

  Since events A and B are *<u>mutually exclusive</u>*, then the event that a sum of 7 or 11 turns up is

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) = \frac{6}{36} + \frac{2}{36} = \boxed{\frac{2}{9}}$$

Useful Formulas

<u>Example (cont'd): Part b solution</u>



Useful Formulas

- Example (cont'd): Part b solution

  Let

  C = event that the both dice turn up the same

  D = event that the sum is less than 5

  Then

  $C \cup D$ = the event that both dice turn up the same or the sum is less than 5

Useful Formulas

- Example (cont'd): Part b solution

  $C = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$

  $D = \{(1,1), (1,2), (1,3), (2,1), (2,2), (3,1)\}$

  Since $C \cap D = \{(1,1), (2,2)\}$, C and D are not mutually exclusive. And the event that both dice turn up the same or the sum is less than 5 is

$$\mathrm{P}(C \cup D) = \mathrm{P}(C) + \mathrm{P}(D) - \mathrm{P}(C \cap B) = \frac{6}{36} + \frac{6}{36} - \frac{2}{36} = \boxed{\frac{5}{18}}$$

Conditional Probability

- The probabilities previously discussed are based on and relate to sample space *S*. However, it is common in many engineering problems to have interest of occurrence of events that are conditioned on occurrence of a subset of the sample space. This is introduces the concept of conditional probability.

Conditional Probability

- The probability of an event may change if we are told of the occurrence of another event.
  - Example 1:
    - If an adult is selected at random from all adults in the United States, the probability of that person having lung cancer would not be too high.

Conditional Probability

– Example 1 (cont'd):

- However, if we are told that the person is also a heavy smoker, then we would certainly want to revise the probability upward. In other words, the probability would be much higher because smoking (specially heavy) causes cancer.
- In general, the probability of the occurrence of an event *A*, given the occurrence of event *B*, is called conditional probability and is denoted by P(*A*|*B*).

Conditional Probability

– Example 1 (cont'd):

- In this example, events *A* and *B* would be

  A = Adult has lung cancer

  B = Adult is a heavy smoker
- And P(*A*|*B*) would represent the probability of an adult having lung cancer, given that he or she is a heavy smoker.

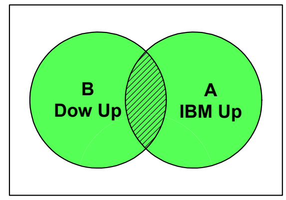
Conditional Probability

- Example 2

Tracking Dow Jones stock prices

A = your stock's price (IBM) up

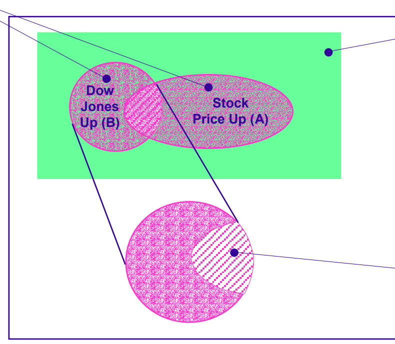
B = Dow Jones up



4 Cases:	**Dow Jones price**	**IBM price**
1	Up	Up
2	Up	Down
3	Down	Up
4	Down	Down

Conditional Probability

Portions representing possibility of DJ going up and SP going down or vise-a-versa.



Neither the Dow Jones index nor the stock price goes up

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

Stock Price Up and Dow Jones Up (*joint outcome or intersection*)

Conditional Probability

- Conditional Probability

  For events *A* and *B* in an arbitrary sample space *S*, the conditional probability of *A* given *B* can be computed as follows:

$$\boxed{\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \qquad \mathrm{P}(B) \neq 0}$$

Conditional Probability

- Example 3:

  The objective herein is to try to formulate a precise definition of P(*A*|*B*) through a simple example.

  What is the probability of rolling a prime numbers (2, 3, or 5) in a single roll of a fair die?

  Let *S* = {1, 2, 3, 4, 5, 6}

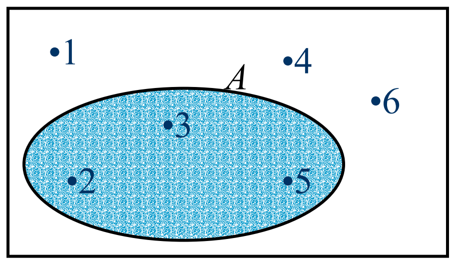
Conditional Probability

- <u>Example 3 (cont'd):</u>

Then the event of rolling a prime number is

A = {2, 3, 5}

$$\mathrm{P}(A) = \frac{n(A)}{n(S)} = \frac{3}{6} = \frac{1}{2}$$



Conditional Probability

- <u>Example 3 (cont'd):</u>

Now suppose we are asked **"In a single roll of a fair die, what is the probability that a prime number has turned up if we are given the <u>additional information</u> that an odd number has turned up?"**

The additional information that another event has occurred, namely,

B = {odd number turns up}

put the problem in a new light.

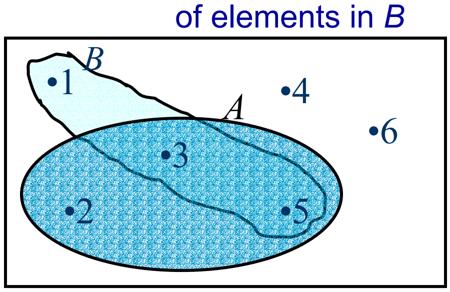
Conditional Probability

- Example: (cont'd):
  - We are now interested only in the part of event *A* (rolling a prime number) that is in event *B* (rolling an odd number).
  - Event *B*, since we know it has occurred, becomes the new sample space.
  - The following Venn diagrams illustrate the various relationships:

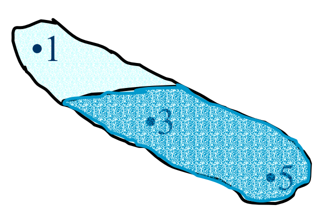
Conditional Probability

- Example: (cont'd):

Thus, the probability of *A* given *B* is the number of *A* elements in *B* divided by the total number of elements in *B*



$$\mathrm{P}(A \mid B) = \frac{n(A \cap B)}{n(B)} = \frac{2}{3}$$



Conditional Probability

- Example: (cont'd):

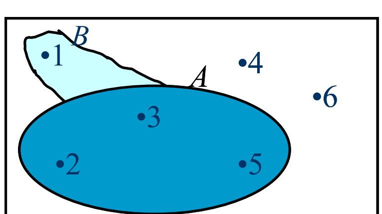
Dividing the numerator and denominator of $n(A \cap B)/n(B)$ by $n(S)$, the number of elements in the original sample space, the expression for the conditional probability can be verified as follows:|

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

Conditional Probability

- Example: (cont'd):

Using the above expression to compute $P(A|B)$ for this example, the same result (as it should be) is obtained as follows:



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

Conditional Probability

- Properties of Conditional Probability
  1. The complement of an event:

$$\boxed{P(\overline{A} \mid B) = 1 - P(A \mid B)}$$

  2. The multiplication rule for two events *A* and *B*:

$$P(A \cap B) = P(A \mid B)P(B) \qquad \text{if } P(B) \neq 0$$
$$P(A \cap B) = P(B \mid A)P(A) \qquad \text{if } P(A) \neq 0$$

Conditional Probability

- Properties of Conditional Probability
  3. The multiplications rule for three events *A*, *B*, and *C*:

$$P(A \cap B \cap C) = P(A \mid (B \cap C))P(B \mid C)P(C)$$
$$= P((A \cap B) \mid C)P(C)$$
$$\text{if } P(C) \neq 0 \text{ and } P(B \cap C) \neq 0$$

Conditional Probability

- Properties of Conditional Probability
  4. For mutually independent events *A* and *B*:

$$\boxed{\begin{aligned} P(A \mid B) = 0 \\ P(B \mid A) = 0 \end{aligned}}$$

  5. For statistically independent events *A* and *B*:

$$\boxed{\begin{aligned} &P(A \mid B) = P(A) \\ &P(B \mid A) = P(B) \\ &P(A \cap B) = P(A)P(B) \end{aligned}}$$

Theorem of Total Probability

- If structural damage (*D*) to a building can only be caused by three events: *fire* (*F*), strong *wind* (*W*), or *earthquake* (*E*), then *D* will depend on whether *F*, *W*, or *E* has occurred, and the likelihood of occurrence of *F*, *W*, and *E.*
- If we assume further that *F*, *W*, and *E* are collectively exhaustive and mutually exclusive events, then

Theorem of Total Probability

the probability of damage to the building can be computed as

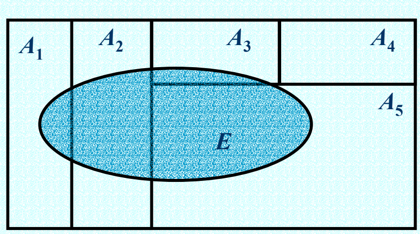
$$P(D) = P(D \mid F)P(F) + P(D \mid W)P(W) + P(D \mid E)P(E)$$

Each term in the right-hand side of the above equation calculates the probability of damage given that *fire*, *wind*, or *earthquake* has occurred. The concept of above equation is called the theorem of total probability.

Theorem of Total Probability

If A_1, A_2, A_3,.. A_n,represents a partition of a sample space S, and $E \subset S$ represents an arbitrary event, the theorem states that

$$\boxed{P(E) = P(A_1)P(E \mid A_1) + P(A_2)P(E \mid A_2) + \ldots + P(A_n)P(E \mid A_n)}$$



Theorem of Total Probability

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

from which we can derive:

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

Now expanding P(A) with the formula for total probability, we obtain:

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

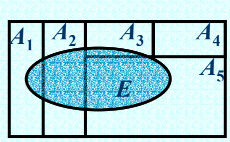
Note: Bayes' theorem is extremely useful in decision analysis, especially when using information.

Theorem of Total Probability

– **Bayes' Theorem**

It is useful in computing the reverse probability of the type $P(A_i | E)$, for $I = 1, 2, \ldots, n$. The reverse probability can be computed as

$$P(A_i | E) = \frac{P(A_i)P(E | A_i)}{P(A_1)P(E | A_1) + P(A_2)P(E | A_2) + \ldots + P(A_n)P(E | A_n)} = \frac{P(A_i)P(E | A_i)}{P(E)}$$



Theorem of Total Probability

- Example: Defective Products in Production Lines

  Consider a factory with three production lines, L_1, L_2, and L_3. Products are either defective (D) or non-defective (ND). The following probabilities are given:

$$\mathrm{P}(D \mid L_1) = 0.1$$
$$\mathrm{P}(D \mid L_2) = 0.1$$
$$\mathrm{P}(D \mid L_3) = 0.2$$

Theorem of Total Probability

- Example (cont'd): Defective Products in Production Lines

  Assuming 20, 30, and 50% of the components are manufactures by lines 1, 2, and 3, the probability of defective components is

$$\mathrm{P}(D) = \mathrm{P}(L_1)\mathrm{P}(D \mid L_1) + \mathrm{P}(L_2)\mathrm{P}(D \mid L_2) + \mathrm{P}(L_3)\mathrm{P}(D \mid L_3)$$
$$= 0.1 \times 0.2 + 0.1 \times 0.3 + 0.2 \times 0.5 = 0.15$$

Uncertain Quantities

- Many uncertain events have quantitative outcomes.
- If an event is not quantitative in the first place, we might define a variable that has a quantitative outcome based on the original event.
- The set of probabilities associated with all possible outcomes of an uncertain quantity is called its *probability distribution*.
- The probabilities in a probability distribution must add to 1 because the events - numerical outcomes - are mutually exclusive.

Note:

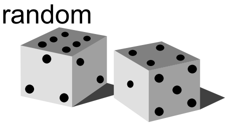
1. Uncertain quantities (often called *random variables*) and their probability distributions play a central role in decision analysis.
2. It is helpful to distinguish between *discrete* and *continuous* uncertain quantities.

Random Variables

- A *random variable* is defined as a function that assigns a real value to every outcome (event) for an engineering system.
- Random variables are commonly classified into two types: *discrete* and *continuous* random variables.

Example: Random Variables

- ***Discrete Random Variables***
- The outcome of a roll of a die may only take on the integer values from 1 to 6.
- The number of floods per year at a point on a river can only take on integer values, so it is also a discrete random variable.



- ***Continuous Random Variables***
- The average of all scores on a test having a maximum possible score of 100 may take on any value including non-integers, between 0 and 100.
- The yield strength of steel can take any non-negative value.

Discrete Probability Distributions

- **The discrete probability distribution case is characterized by an uncertain quantity that can assume a finite or countable number of possible values.**
- **When we specify a probability distribution for a discrete uncertain quantity, we can express the distribution in several ways.**
- **The two approaches that are particularly useful are:**
  - **The probability mass function, and**
  - **A cumulative distribution function (CDF).**

Discrete Probability Distributions

- Probability Mass Function
  - The set of ordered pairs (x_i, $P(x_i)$) is a probability mass function or probability distribution of the discrete random variable X, if for each possible outcome x_i:

$$P_X(x_i) = \mathrm{P}(X = x_i)$$
$$0 \le P_X(x_i) \le 1$$
$$\sum_{i=1}^{N} P_X(x_i) = 1$$

Discrete Probability Distributions

- Cumulative Mass Function
  - The cumulative mass function $F_X(x_i)$of a discrete random variable X with probability mass function $P_X(x_i)$ is given by

$$F_X(x_i) = \mathrm{P}(X \le x_i) = \sum_{j=1}^{i} P_X(x_j)$$

Discrete Probability Distributions

- Example
  - Suppose that you think that no cookie in a batch of oatmeal cookies could have more than five raisons. A possible probability mass function would be

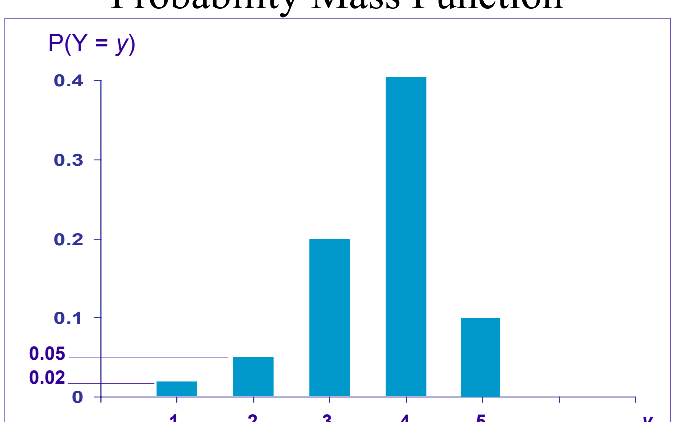
P (Y = 0 raisins) = 0.02
P (Y = 1 raisins) = 0.05
P(Y = 2 raisins) = 0.20
P (Y = 3 raisins) = 0.40
P (Y = 4 raisins) = 0.22
P (Y = 5 raisins) = 0.11

$$P_X(x_i) = \mathrm{P}(X = x_i)$$
$$0 \le P_X(x_i) \le 1$$
$$\sum_{i=1}^{N} P_X(x_i) = 1$$

Discrete Probability Distributions

- Example (cont'd)

Probability Mass Function



Discrete Probability Distributions

- Example (cont'd)
  - A cumulative distribution gives the probability that an uncertain quantity is less than or equal to a specific value P($X \le x$). For this example the CMF is

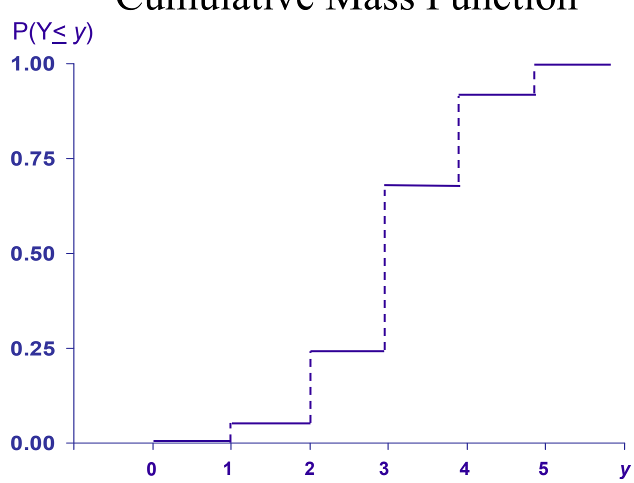
P($Y \le$ 0 raisins) = 0.02
P($Y \le$ 1 raisins) = 0.07
P($Y \le$ 2 raisins) = 0.27
P($Y \le$ 3 raisins) = 0.67
P($Y \le$ 4 raisins) = 0.89
P($Y \le$ 5 raisins) = 1.00

$$\boxed{F_X(x_i) = \mathrm{P}(X \le x_i) = \sum_{j=1}^{i} P_X(x_j)}$$

Discrete Probability Distributions

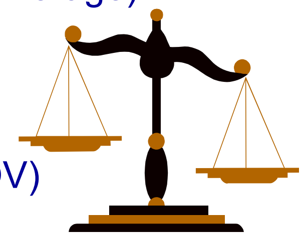
- Example (cont'd)

Cumulative Mass Function



Discrete Probability Distributions

- Measures of Probability
  - Expected Value (Mean or Average)
  - Variance
  - Standard Deviation, σ
  - Coefficient of variation (COV)
  - Skewness



Discrete Probability Distributions

Definition:

- **The variation of uncertain quantity *X* is denoted by Var(*X*) or σ_X^2 (Greek sigma) and is calculated mathematically by:**

$$\begin{aligned}
\mathrm{Var}(X) &= [x_1 - \mathrm{E}(X)]^2 \mathrm{P}(X = x_1) + [x_2 - \mathrm{E}(X)]^2 \mathrm{P}(X = x_2) \\
&\quad + \ldots + [x_n - \mathrm{E}(X)]^2 \mathrm{P}(X = x_n) \\
&= \sum_{i=1}^{n} [x_i - \mathrm{E}(X)]^2 \mathrm{P}(X = x_i) \\
&= \mathrm{E}[X - \mathrm{E}(X)]^2
\end{aligned}$$

Discrete Probability Distributions

Definition:

- **The *standard deviation* of *X*, denoted by σ_X is just the square root of the variance.**
- **Because the variance is the expected value of the squared differences, the standard deviation can be thought of as a "best guess" as to how far the outcome of the *X* might lie from E(*X*).**

Note:

1. **A large standard deviation and variance means that the probability distribution is quite spread out; a large difference between the outcome and the expected value is anticipated. For this reason, the variance and the standard deviation of a probability distribution are used as measures of variability or Risk.**
2. **A large variance or standard deviation would indicate a situation in which the outcomes are highly variance.**

Discrete Probability Distributions

- Measures of Discrete Random Variables
  - If *X* is a discrete random variable with PMF $P_X(x)$, the following expressions can be used to compute the mean, variance, and skewness:

$$E(X) = \mu = \sum_{i=1}^{n} x_i P_X(x_i)$$

$$\text{Variance} = \text{Var}(X) = \sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 P_X(x_i)$$

$$\text{Skewness} = \lambda = \sum_{i=1}^{n} (x_i - \mu)^3 P_X(x_i)$$

Discrete Probability Distributions

- Special Cases
  - If the function $Y = g(X) = a + b\,X$, then

$$\mathrm{E}(Y) = a + b\mathrm{E}(X)$$
$$\mathrm{Var}(Y) = b^2\mathrm{Var}(X)$$

Where a and b are real numbers.

Discrete Probability Distributions

- Special Cases
  - If the function $Y = g(X)$ is given by

$$Y = g(\boldsymbol{X}) = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

Then

$$\mathrm{E}(Y) = a_0 + a_1\mathrm{E}(X_1) + a_2\mathrm{E}(X_2) + \ldots + a_n\mathrm{E}(X_n)$$

and

$$\mathrm{Var}(Y) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \mathrm{Cov}(X_i, X_j) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j}$$

If the random variables of $\boldsymbol{X}$ are uncorrelated, then

$$\mathrm{Var}(Y) = \sum_{i=1}^{n} a_i^2 \mathrm{Var}(X_i)$$

Discrete Probability Distributions

- Example: Three Cars
  - If there are three cars, the following situations are possible:
    - All three cars in good condition.
    - Two cars are good and one is bad.
    - One car is good and two cars are bad.
    - All three cars are in bad condition

S

G	G	G	B	B	B	G	B
G	G	B	G	B	G	B	B
G	B	G	G	G	B	B	B

Venn Diagram

Discrete Probability Distributions

- Example: Three Cars
  - Assume that a car will be in good condition 90% of the time and in bad condition 10% of the time.
  - Thus, P(G) = 0.90 and P(B) = 0.10
  - If X is a random variable representing the number of good cars at a given time, for this problem, X = 0, 1, 2, or 3.
  - The PMF's for these values of X can be computed as shown in the next viewgraph.

	G	G	G	B	B	B	G	B
	G	G	B	G	B	G	B	B
	G	B	G	G	G	B	B	B
$X =$	3		2			1		0

Discrete Probability Distributions

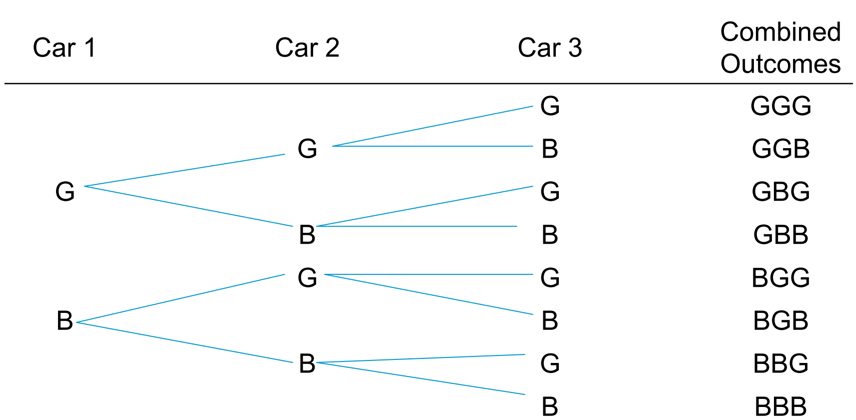
- Example (cont'd): Three Cars
  - PMF for three cars

$$P_X(0) = \mathrm{P}(X = 0) = 0.1 \times 0.1 \times 0.1 \quad = 0.001$$
$$P_X(1) = \mathrm{P}(X = 1) = 3 \times 0.9 \times 0.1 \times 0.1 = 0.027$$
$$P_X(2) = \mathrm{P}(X = 2) = 3 \times 0.9 \times 0.9 \times 0.1 = 0.243$$
$$P_X(3) = \mathrm{P}(X = 3) = 0.9 \times 0.9 \times 0.9 \quad = 0.729$$
$$\sum 1.000$$

G G G	G G B	G B G	B G G	B B G	B G B	G B B	B B B
3	2			1			0

Discrete Probability Distributions

- Example (cont'd): Three Cars



Discrete Probability Distributions

- <u>Example (cont'd):</u> Three Cars

$$E(X)=\mu=\sum_{i=1}^{n}x_iP_X(x_i)=0\times0.001+1\times0.027+2\times0.243+3\times0.729=2.7$$

$$\text{Var}(X)=\sigma^2=\sum_{i=1}^{n}(x_i-\mu)^2P_X(x_i)=(0-2.7)^2\times0.001+(1-2.7)^2\times0.027$$
$$+(2-2.7)^2\times0.243+(3\text{-}2.7)^2\times0.729=0.27$$

$$\text{Skewness}=\lambda=\sum_{i=1}^{n}(x_i-\mu)^3P_X(x_i)=(0-2.7)^3\times0.001+(1-2.7)^3\times0.027$$
$$+(2-2.7)^3\times0.243+(3\text{-}2.7)^3\times0.729=-0.216$$

$$\sigma=\sqrt{0.27}=52,\qquad \text{COV}=\frac{0.52}{2.7}=0.19$$

Continuous Probability Distributions

- **Continuous uncertain quantities. The uncertain quantity can take any value within some range.**

<u>Example:</u>

- **The temperature tomorrow at O'Hara Airport in Chicago at noon is an uncertain quantity that can be anywhere between, say, 50ºF and 120º F.**

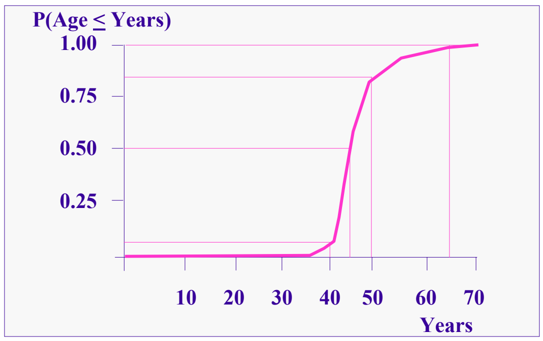
<u>Note:</u>

1. **With continuous uncertain quantities, it is not reasonable to speak of the probability that that a specific value occurs.**
2. **The probability of a particular value occurring is equal to zero: P(Y = y) = 0.**
3. **The probability of any particular value must be infinitely small.**
4. **We typically speak of interval probabilities: P (a $\leq Y \leq b$). The CDF for a continuous uncertain quantity can be constructed on the basis of such intervals.**

Continuous Probability Distributions

Example:

■ Let us suppose we are interested in a movie star's age. Table of cumulative probabilities for a move star's age.



P(Age ≤ 29) = 0.00
P(Age ≤ 40) = 0.05
P(Age ≤ 44) = 0.50
P(Age ≤ 50) = 0.85
P(Age ≤ 65) = 1.00

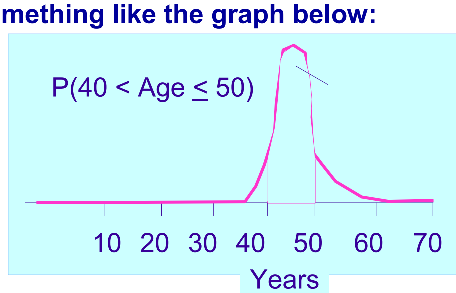
The CDF allows us to calculate the probability for any interval.

Continuous Probability Distributions

■ The CDF for a continuous uncertain quantity corresponds closely to the CDF for the discrete case. The density function $f(x)$ can be built up from the CDF. It is a function in which the area under the curve within a specific interval represents the probability that the uncertain quantity will fall in that interval.

Example:

■ The density function f(Age) for the movie star's age might look something like the graph below:

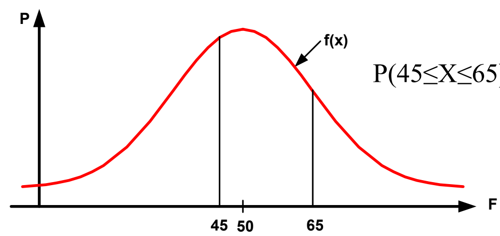


Probability density function for movie star's age.

Continuous Probability Distributions

- **Note**
  - **Unlike discrete r.v.'s, continuous ones can taken as a specific value within probability of zero.**

    P(X=x)=0 for continuous r.v. X
  - **Can have P(a≤X≤b) > 0 (ranges)**
  - **Instead of probability mass function (discrete r.v.), we have probability density function. (continuous r.v.)**
- **Example: Temperature X at College Park at noon throughout the year:**



P(45≤X≤65) = ?

Continuous Probability Distributions

- The ***probability density function*** **(PDF)** defines the probability of occurrence for a continuous random variable.
- The probability that the random variable *X* lies within the interval from x_1 to x_2 is given by:

$$P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} f_X(x)\,dx$$

where $f_X(x)$ = probability density function

Continuous Probability Distributions

- If the interval is made infinitesimally small, x_1 approaches x_2 and $P(x_1 \leq X \leq x_2)$ approaches zero.
- This illustrates a property that distinguishes discrete random variables from continuous variables.
- Therefore, the probability that a continuous random variable takes on a specific value equals zero

Continuous Probability Distributions

- Some Useful Properties:

$$P(-\infty \leq X \leq +\infty) = \int_{-\infty}^{+\infty} f_X(x)dx = 1$$

$$P(X \geq x_0) = \int_{x_0}^{+\infty} f_X(x)dx = 1 - P(X < x_0)$$

Continuous Probability Distributions

– The ***cumulative distribution function*** **(CDF)** of a continuous random variable is defined by

$$F_X(x_0) = \mathrm{P}(X \le x_0) = \int_{-\infty}^{x_0} f_X(x)\,dx$$

where $f_X(x)$ = probability density function

Continuous Probability Distributions

- Example A:

The continuous random variable X has the following probability density function:

$$f_X(x) = \begin{cases} kx & \text{for } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant. Find the value of k that is necessary for $f_X(x)$ to be a legitimate probability density function.

Continuous Probability Distributions

- Example A (cont'd):

Plot both the density and cumulative functions. What is the probability that *X* equals 1? What is the probability that *X* takes on a value less than 0.5? What is the probability that *X* is greater than 1.0 and less than 1.5?

For $f_X(x)$ to be a legitimate PDF, it must satisfy the following equation:

$$\mathrm{P}(-\infty < X < +\infty) = \int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x = 1$$

Continuous Probability Distributions

- Example A (cont'd):

Therefore,

$$1 = \int_{-\infty}^{+\infty} f_X(x)\,\mathrm{d}x = \int_0^2 kx\,\mathrm{d}x = k\int_0^2 x\,\mathrm{d}x = \left.\tfrac{kx^2}{2}\right|_0^2$$

$$= k\frac{4}{2} = 2k \quad \Rightarrow \quad k = 0.5$$

The cumulative distribution function is given by

$$F_X(x_0) = \int_0^{x_0} 0.5x\,\mathrm{d}x = \left.\frac{x^2}{4}\right|_0^{x_0} = \frac{x_0^2}{4}$$

Continuous Probability Distributions

- Example A (cont'd):

The density and cumulative functions are provided in the following table:

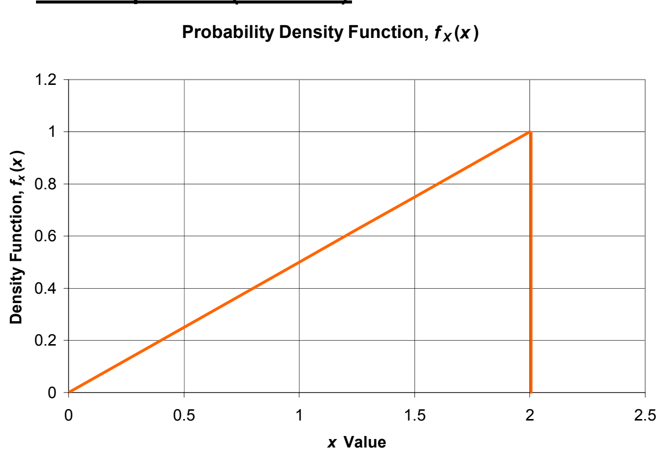
x	Density Function, $f_X(x)$	Cumulative Distribution Function, $F_X(x)$
0	0	0
0.5	0.25	0.0625
1	0.5	0.25
1.5	0.75	0.5625
2	1	1

$$f_X(x) = \begin{cases} \dfrac{x}{2} & \text{for } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(x_0) = \frac{x_0^2}{4}$$

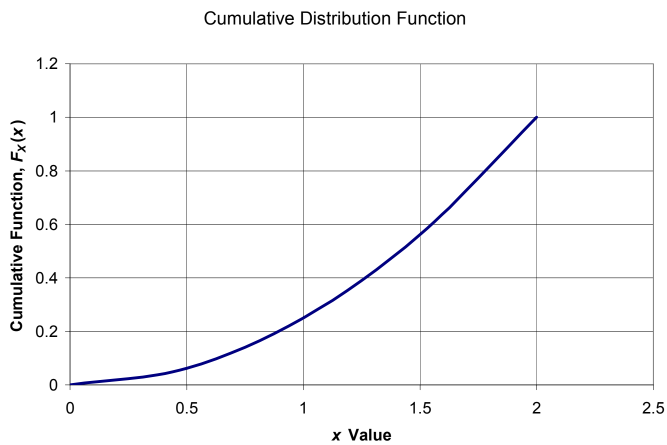
Continuous Probability Distributions

- Example A (cont'd):



Continuous Probability Distributions

- Example A (cont'd):



Continuous Probability Distributions

- Example A (cont'd):
  - Because probabilities of continuous random variables are defined for regions rather than point values,

$$\mathrm{P}(X = 1) = 0$$

  - The P(*X*<0.5) can be determined from the cumulative function as

$$\mathrm{P}(X < 0.5) = \left.\frac{x_0^2}{4}\right|_{x_0=0}^{x_0=0.5} = \frac{1}{16}$$

Continuous Probability Distributions

- Example A (cont'd):
  - Similarly, the cumulative function can be used to find the probability for the following region:

$$\mathrm{P}(1 < X < 1.5) = \mathrm{P}(X < 1.5) - \mathrm{P}(X < 1.0)$$

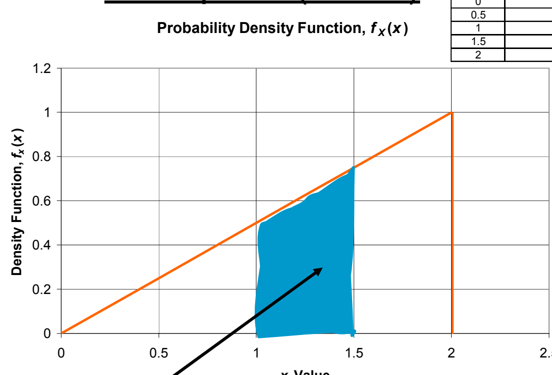
$$= \left.\frac{x_0^2}{4}\right|_{x_0=0}^{x_0=1.5} - \left.\frac{x_0^2}{4}\right|_{x_0=0}^{x_0=1.0}$$

$$= \frac{9}{16} - \frac{1}{4} = \frac{5}{16} = 0.313$$

Continuous Probability Distributions

- Example A (cont'd):

x	Density Function, $f_X(x)$	Cumulative Distribution Function, $F_X(x)$
0	0	0
0.5	0.25	0.0625
1	0.5	0.25
1.5	0.75	0.5625
2	1	1

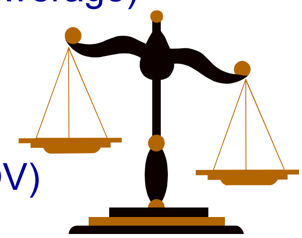


$$\text{area} = (1.5 - 1)\frac{0.75 + 0.5}{2} = 0.313$$

Shaded area under the curve = P(1.0 <X<1.5)

Continuous Probability Distributions

- Measures of Probability
  - Expected Value (Mean or Average)
  - Variance
  - Standard Deviation, σ
  - Coefficient of variation (COV)
  - Skewness



Continuous Probability Distributions

- Measures of Continuous Random Variables
  - If X is a continuous random variable with PDF $f_X(x)$, the following expressions can be used to compute the mean, variance, and skewness:

$$E(X) = \mu = \int_{-\infty}^{+\infty} x f_X(x)\,\mathrm{d}x$$

$$\text{Variance} = \mathrm{Var}(X) = \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)\,\mathrm{d}x$$

$$\text{Skewness} = \lambda = \int_{-\infty}^{+\infty} (x-\mu)^3 f_X(x)\,\mathrm{d}x$$

Continuous Probability Distributions

- Measures Continuous Random Variables
  - Useful expression for the Variance

$$\text{Variance} = \text{Var}(X) = \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)\,\mathrm{d}x$$
$$= E\left(X^2\right) - \mu^2$$

where

$$E\left(X^2\right) = \int_{-\infty}^{+\infty} x^2 f_X(x)\,\mathrm{d}x$$

Continuous Probability Distributions

- Example B

For the continuous random variable *X* of Example A that has the following probability density function:

$$f_X(x) = \begin{cases} kx & \text{for } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Determine the mean, variance, standard deviation, and coefficient of variation (*COV*) of the random variable *X*.

Continuous Probability Distributions

- Example B (cont'd):

  The density and cumulative functions are provided in the following table:

x	Density Function, $f_X(x)$	Cumulative Distribution Function, $F_X(x)$
0	0	0
0.5	0.25	0.0625
1	0.5	0.25
1.5	0.75	0.5625
2	1	1

$$f_X(x) = \begin{cases} \dfrac{x}{2} & \text{for } 0 \le x \le 2 \\ 0 & \text{otherwise} \end{cases} \qquad F_X(x_0) = \frac{x_0^2}{4}$$

Continuous Probability Distributions

- Example B (cont'd):
  - The *mean* value can be computed as follows:

$$\mu = E(X) = \int_{-\infty}^{+\infty} x f_X(x)\,dx + \int_0^2 x(0.5x)\,dx = \frac{4}{3}$$

  - The *variance* can be calculated as follows:

$$\sigma^2 = VAR(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f_X(x)\,dx$$

$$= \int_0^2 \left(x - \frac{4}{3}\right)^2 (0.5x)\,dx = \frac{2}{9}$$

Continuous Probability Distributions

- Example (cont'd):
  - The *standard deviation* and *the coefficient of variation* (*COV*) can be computed as follows:

$$\text{Standard Deviation} = \sqrt{\sigma^2} = \sqrt{\frac{2}{9}} = 0.4714$$

and

$$COV(X) = \frac{0.4714}{4/3} = 0.3771$$