Chapter 7

Random Variables

&

Probability Distributions

Sections 7.1 – 7.3

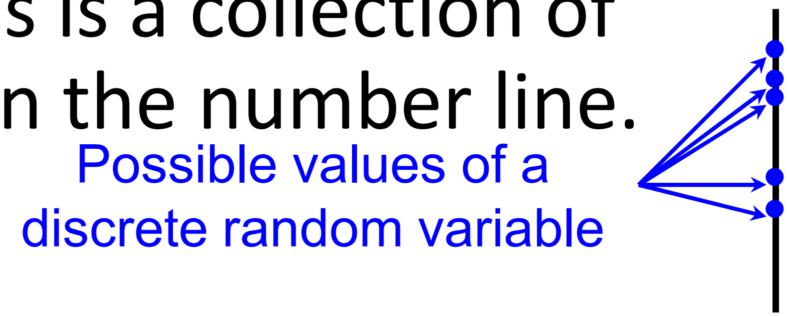
- 7.1 Random Variables
- Probability Distributions: Discrete
- Probability Distributions: Continuous

7.1: Random Variables

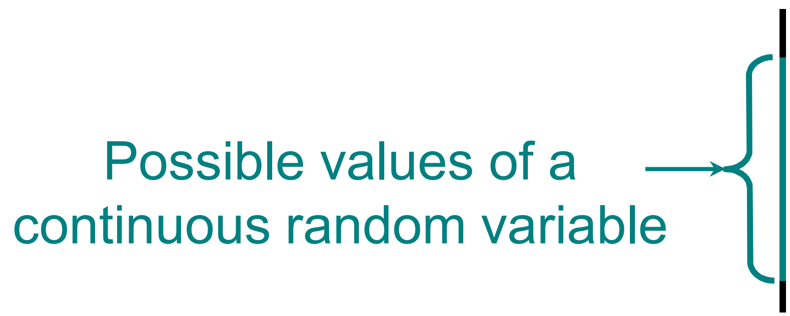
- A numerical variable whose value depends on the outcome of a chance experiment is called a **random variable**. A random variable associates a numerical value with each outcome of a chance experiment.

Discrete and Continuous Random Variables

- A random variable is **discrete** if its set of possible values is a collection of isolated points on the number line.



A random variable is **continuous** if its set of possible values includes an entire interval on the number line.



We will use lowercase letters, such as x and y, to represent random variables.

Examples

- 1. Experiment: A fair die is rolled
- Random Variable: The number on the up face
- Type: **Discrete**

2. Experiment: A pair of fair dice are rolled
Random Variable: The sum of the up faces
Type: **Discrete**

Examples

- 3. Experiment: A coin is tossed until the 1st head turns up
- Random Variable: The number of the toss that the 1st head turns up
- Type: **Discrete**

- 4. Experiment: Choose and inspect a specified size for a manufactured part
- Random Variable: The difference in length (mm) of the part compared to its prescribed optimum.
- Type: **Continuous**

Examples

- 5. Experiment: Inspect the precision of Primary mirror (Hubble Telescope)
- Random Variable: The number of defects on the surface of the mirror
- Type: **Discrete (strictly a count)**

- 6. Experiment: Inspect the precision of Primary mirror (Hubble Telescope)
- Random Variable: Percentage Variation in amount of curvature compared to optimum
- Type: **Continuous (limit of measurement is strictly dependent on precision of tools)**

Examples

- 7. Experiment: Measure the voltage in a outlet in your room
- Random Variable: The voltage
- Type: Continuous

- 8. Experiment: Observe the amount of time it takes a bank teller to serve a customer
- Random Variable: time in minutes
- Type: Continuous

Examples

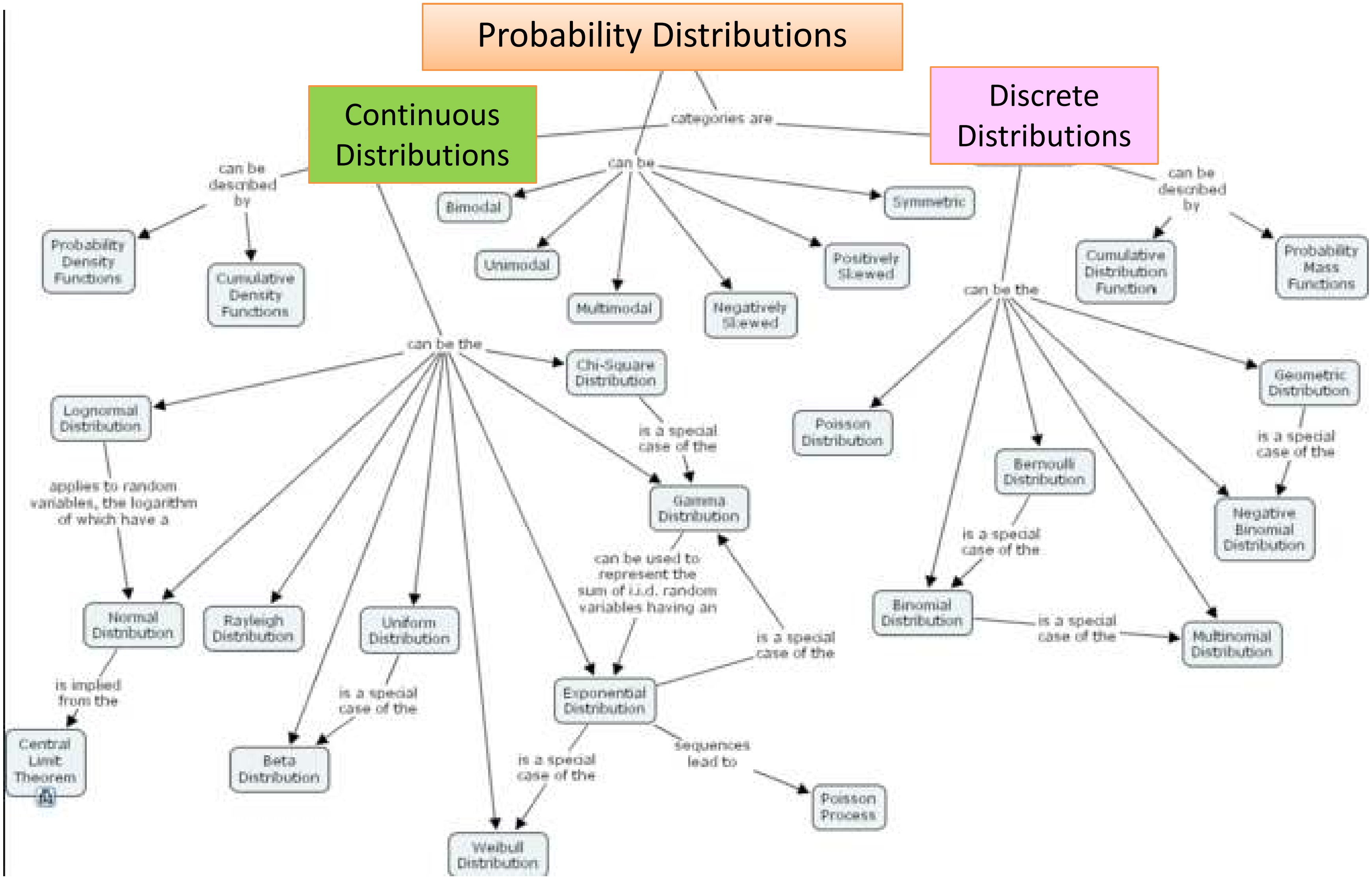
- 9. Experiment: Measure the time until the next customer arrives at a customer service window
- Random Variable: The time
- Type: Continuous

- 10. Experiment: Inspect a randomly chosen circuit board from a production line
- Random Variable:

  1, if the circuit board is defective

  0, if the circuit board is not defective
- Type: Discrete

Schematic Typology



Notation for Random Variables

- For a probability $P(\,X \leq x\,)$, what do x and X mean here?

A chosen constant

-

- Consider X to be the random variable which represents the outcome of a single roll of a die, so that X can take on values of $\{1,2,3,4,5,6\}$
- $P(\,X \leq 2\,)$ means what is the probability that the outcome will be 1 or 2.
- $P(\,X \leq 5\,)$ means what is the probability that the outcome will be 1,2, 3, 4, or 5.

Notation: In general $P(X \leq x)$

- …means the probability that the random variable X is less than or equal to the realization x .
- Our textbook might show the following: Given two common dice, the random variable is the sum of the two dice $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
- What is the probability that the sum is six or less?

$$\boldsymbol{P}(\boldsymbol{X} \leq \mathbf{6}) = \boldsymbol{p} (\boldsymbol{x} \leq \mathbf{6})$$

- What is the probability that the sum is 9?

$$\boldsymbol{P}(\boldsymbol{X} = \mathbf{9}) = \boldsymbol{p} (\boldsymbol{x} = \mathbf{9}) = \boldsymbol{p}(\mathbf{9})$$

7.2: Probability Distributions for Discrete Random Variables

•The **probability distribution of a discrete random variable x** gives the probability associated with each possible x value.

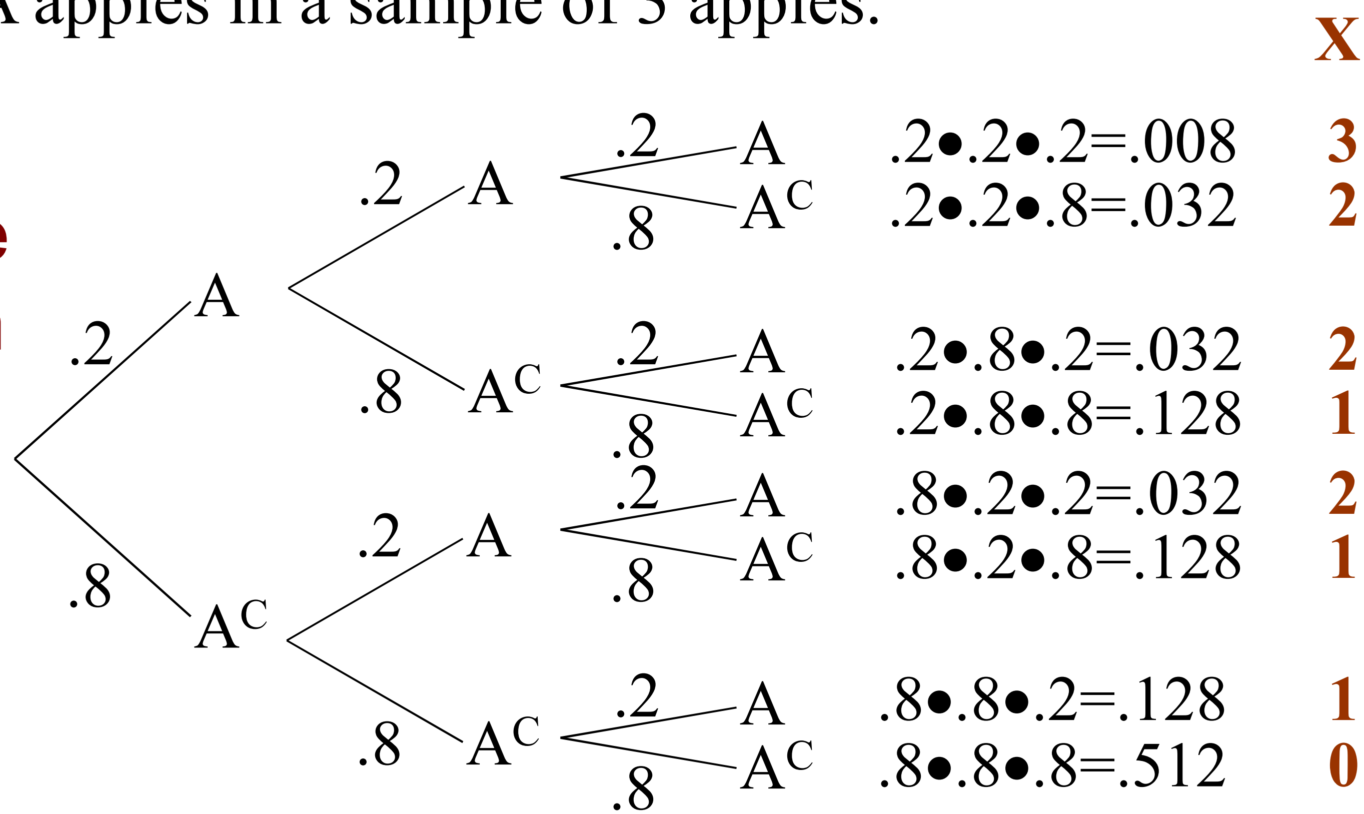
•Each probability is the limiting relative frequency of occurrence of the corresponding x va … dly per …

Roll	1	2	3	4	5	6
p =	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
p =	**0.167**	**0.167**	**0.167**	**0.167**	**0.167**	**0.167**

Example

Suppose that 20% of the apples sent to a sorting line are Grade A. If 3 of the apples sent to this plant are chosen randomly, determine the probability distribution of the number of Grade A apples in a sample of 3 apples.

Consider the tree diagram



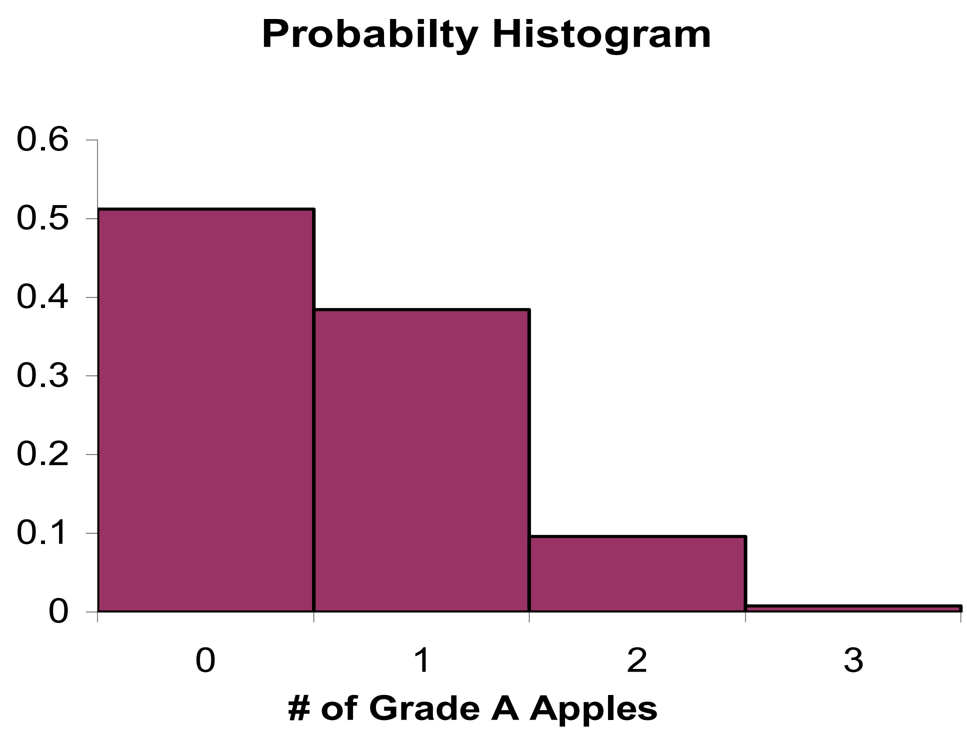
The Results in Table Form

x	p(x)
0	$1(.8)^3$
1	$3(.8)^2(.2)^1$
2	$3(.8)^1(.2)^2$
3	$1(.2)^3$

or

x	p(x)
0	0.512
1	0.384
2	0.096
3	0.008

Results in Graphical Form (Probability Histogram)



For a probability histogram, the area of a bar is the probability of obtaining that value associated with that bar.

Properties of Discrete Probability Distributions

- The probabilities p_i must satisfy

  1. $0 \le p_i \le 1$ for each i

  2. $p_1 + p_2 + ... + p_k = 1$

- The probability P(X in A) of any event is found by summing the p_i for the outcomes x_i making up A.

Example

The number of items a given salesman sells per customer is a random variable. The table below is for a specific salesman (Wilbur) in a clothing store in the mall. The probability distribution of X is given below:

x	0	1	2	3	4	5	6
p(x)	0.20	0.35	0.15	0.12	0.10	0.05	0.03

Note: $0 \leq p(x) \leq 1$ for each x

$\Sigma p(x) = 1$ (the sum is over all values of x)

Example - continued

x	0	1	2	3	4	5	6
p(x)	0.20	0.35	0.15	0.12	0.10	0.05	0.03

The probability that he sells at least three items to a randomly selected customer is

$P(X \geq 3) = 0.12 + 0.10 + 0.05 + 0.03 = \mathbf{0.30}$

The probability that he sells *at most three items* to a randomly selected customer is

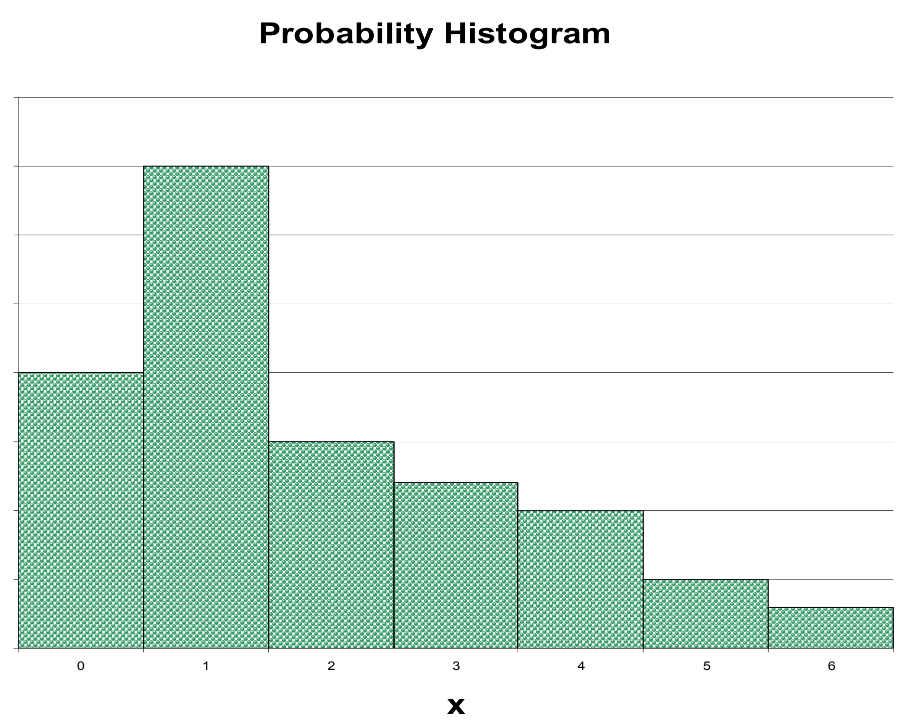
$P(X \leq 3) = 0.20 + 0.35 + 0.15 + 0.12 = \mathbf{0.82}$

The probability that he sells between (inclusive) 2 and 4 items to a randomly selected customer is

$P(2 \leq X \leq 4) = 0.15 + 0.12 + 0.10 = \mathbf{0.37}$

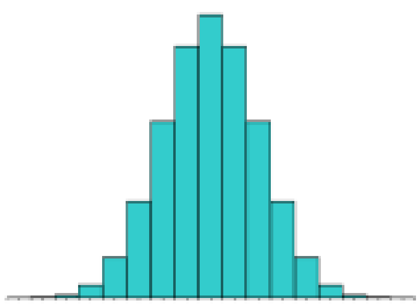
Probability Histogram

- A probability histogram has its vertical scale adjusted in a manner that makes the area associated with each bar equal to the probability of the event that the random variable takes on the value describing the bar.

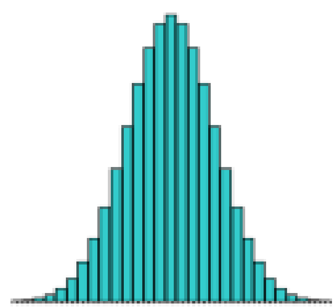


Continuous Probability Distributions

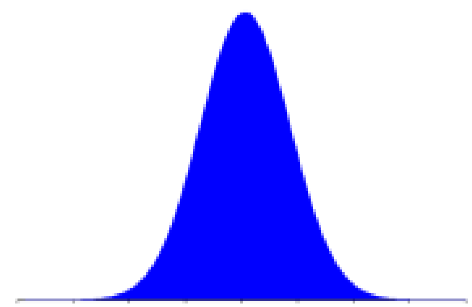
If one looks at the distribution of the actual amount of water (in ounces) in "one gallon" bottles of spring water they might see something such as



Amount measured to nearest hundredths of an ounce.



Amount measured to nearest ten thousands of an ounce.



Limiting curve as the accuracy increases

7.3: Probability Distribution for a Continuous Random Variable

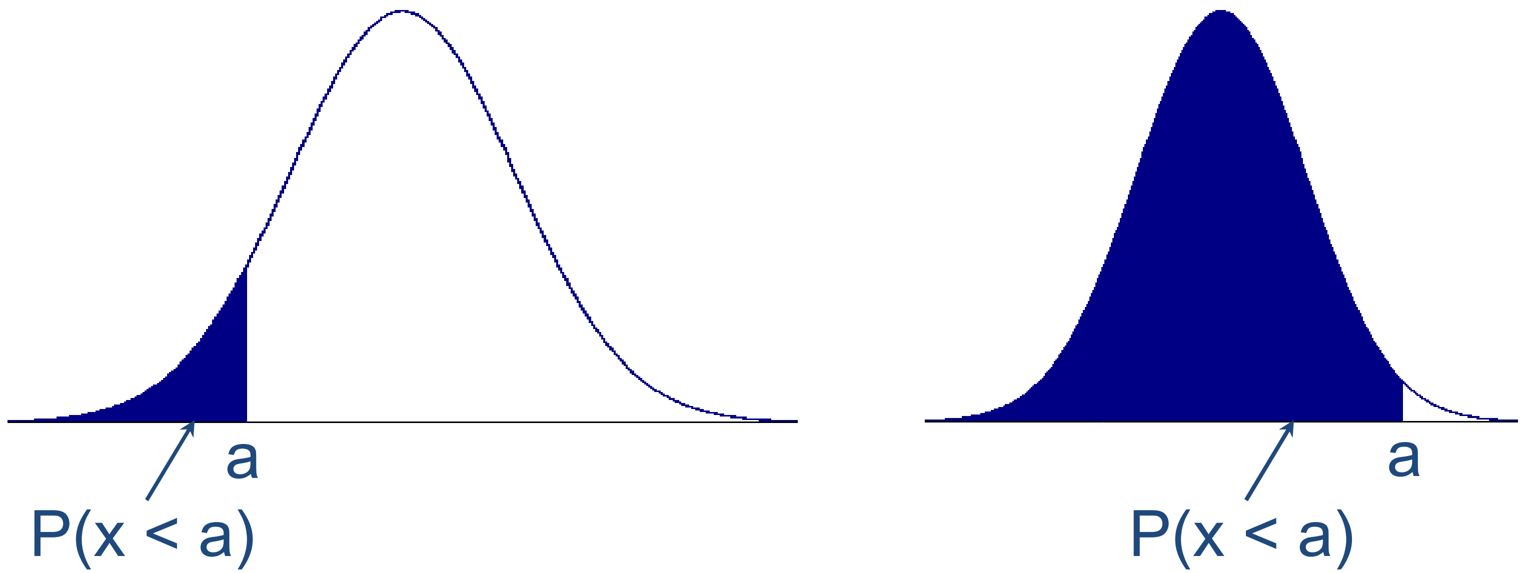
- A **probability distribution for a continuous random variable x** is specified by a mathematical function denoted by f(x) which is called the **density function**. The graph of a density function is a smooth curve (the **density curve**).

- The following requirements must be met:

    1. $f(x) \geq 0$
    2. The total area under the density curve is equal to 1.

- The probability that x falls in any particular interval is the area under the density curve that lies above the interval.

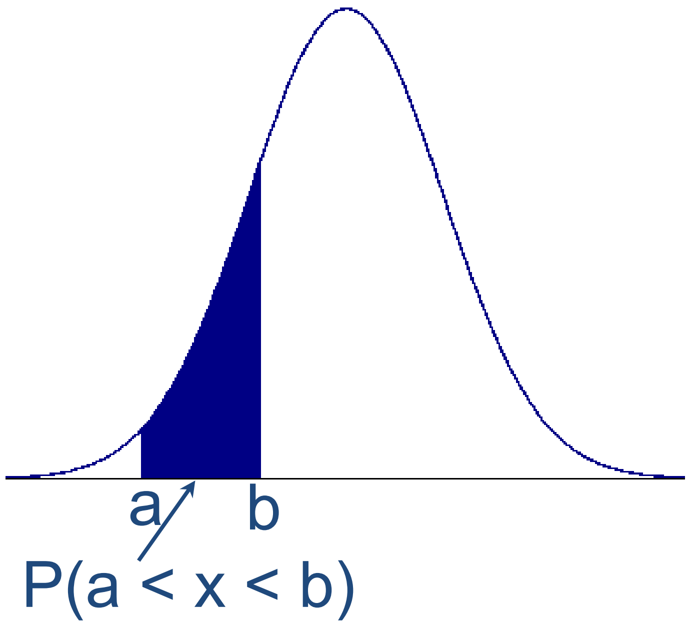
Some Illustrations

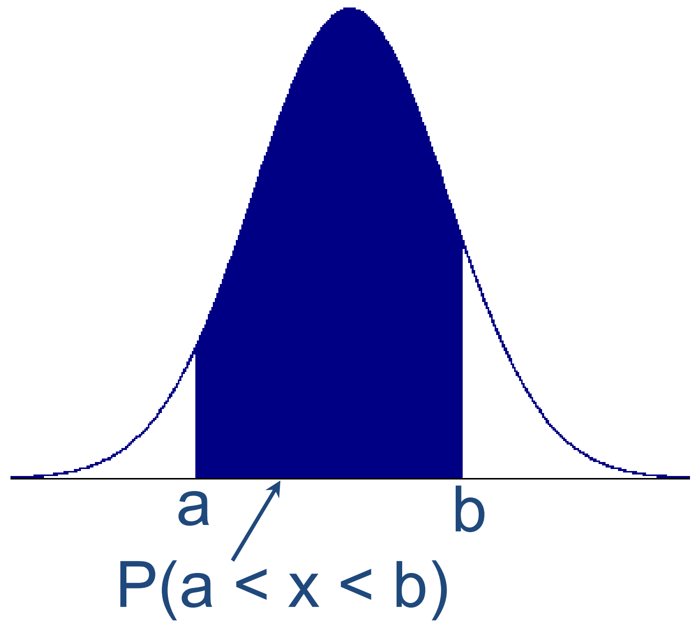


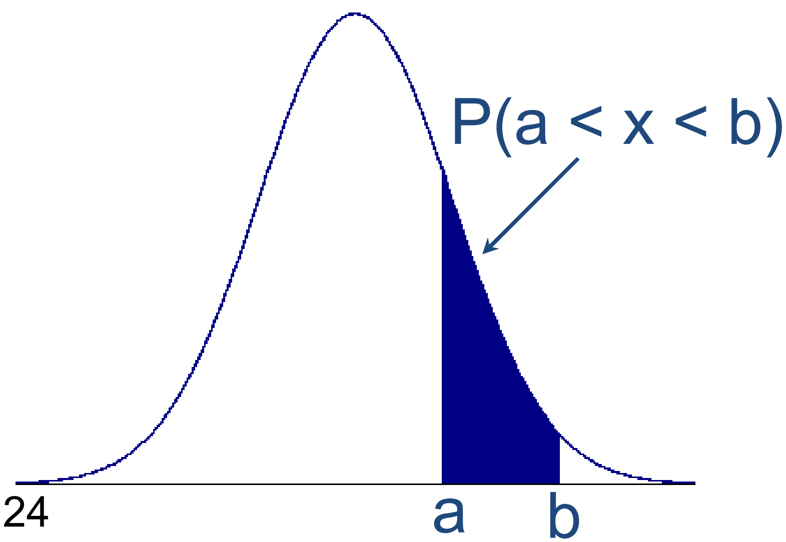
Notice that for a continuous random variable x,
P(x = a) = 0 for any specific value a because the "area above a point" under the curve is a line segment and hence has o area.

Specifically this means P(x < a) = P(x ≤ a).

Some Illustrations







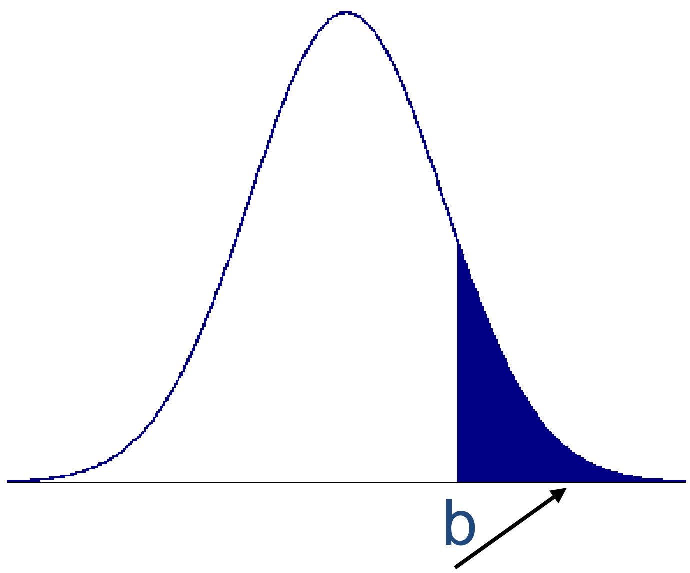
Note:

$P(a < x < b) = P(a \le x < b)$

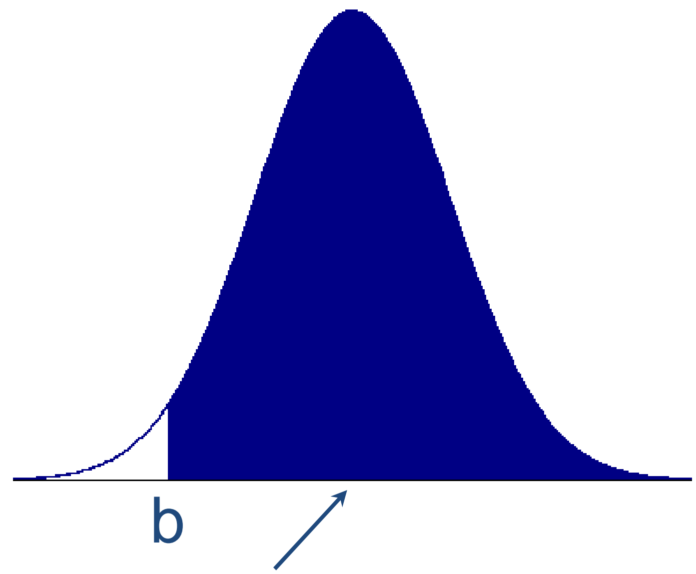
$= P(a < x \le b)$

$= P(a \le x \le b)$

Some Illustrations



P(x > b) = P(b < x)



P(x > b) = P(b < x)

Example

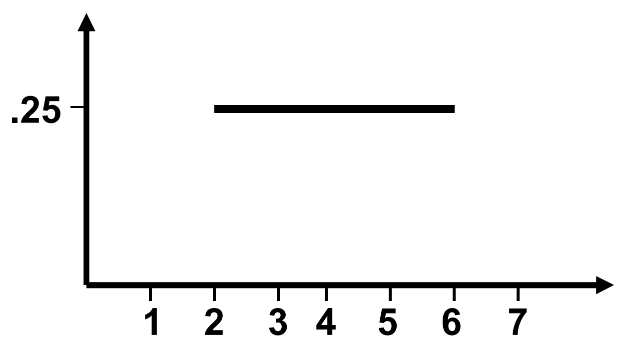
Define a continuous random variable x by

x = the weight of the crumbs in ounces left on the floor of a restaurant during a one hour period.

Suppose that x has a probability distribution with density function

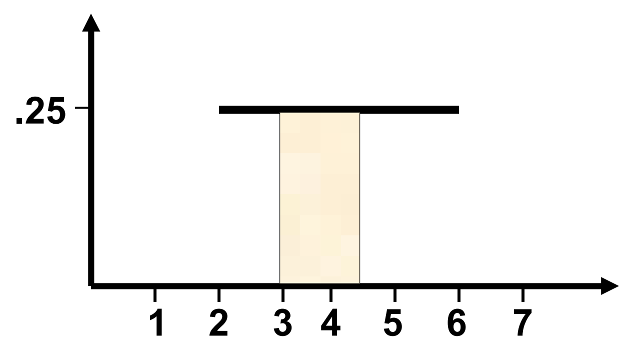
$$f(x) = \begin{cases} .25 & 2 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

The graph looks like



Example

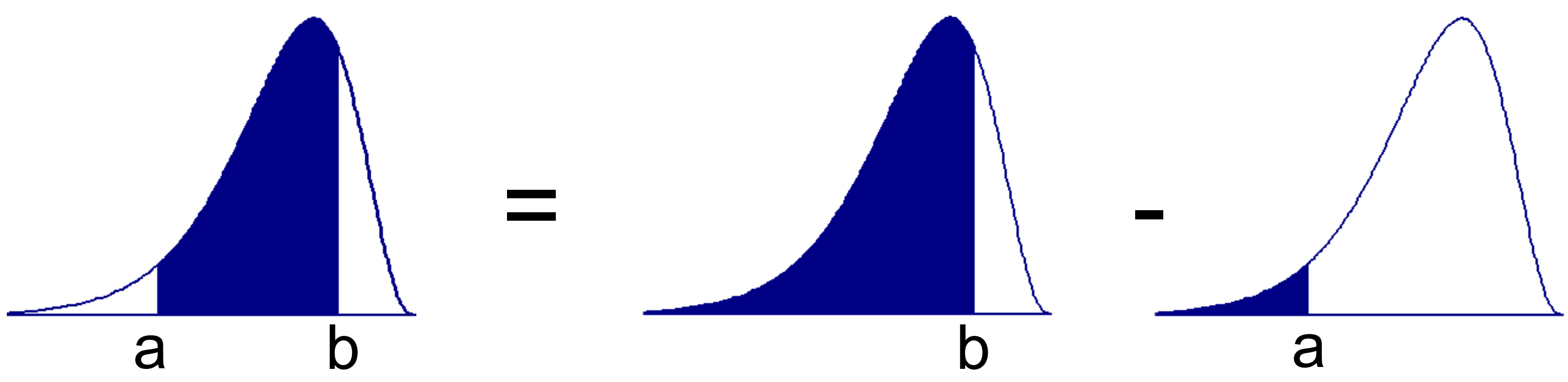
Find the probability that during a given 1 hour period between 3 and 4.5 ounces of crumbs are left on the restaurant floor.



The probability is represented by the shaded area in the graph. Since that shaded area is a rectangle, area = (base)(height)=(1.5)(.25) = .375

Method of Probability Calculation

- The probability that a continuous random variable x lies between a lower limit a and an upper limit b is
- $P(a < x < b)$ = (cumulative area to the left of b) –
- (cumulative area to the left of a)
- $= P(x < b) - P(x < a)$



Next topics

- Section 7.4: Mean & Standard Deviation of a Random Variable
- Section 7.5 : Binomial & Geometric Distributions