Chapter 7

Random Variables

&

Probability Distributions

Warm-UP

1) What, if anything, is the difference in these variables from the following examples?

A) $3x - 4 = 16$

B) $3x - 4 \leq 16$

C) $3x - 4y = 16$

2) Use the information to create normal distributions distributions based on the summary statistics:

Set A: $n = 24; \bar{x} = 69, \; s_x = 12.2,$

Set B: $n = 22; \bar{x} = 62.8, \; s_x = 15.7$

Warm-UP

Variables are symbols that represent some unknown quantity, but the unknown can take on different values depending upon the situation:

A) $3x - 4 = 16$; Here the unknown *x*, is simply a single value

B) $3x - 4 \leq 16$; Here the unknown *x*, actually varies, taking on a range of values

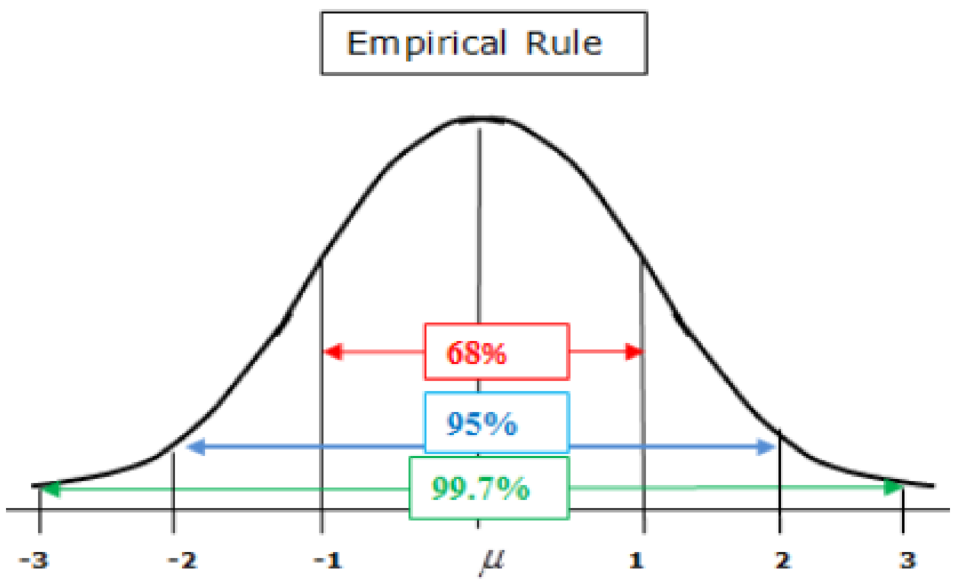
C) $3x - 4y = 16$; Here there are 2 unknowns, $x \text{ and } y$, that have an infinite number of values, related to each other based upon a defined rule

Warm-Up

Use the information to create normal distributions distributions based on the summary statistics:

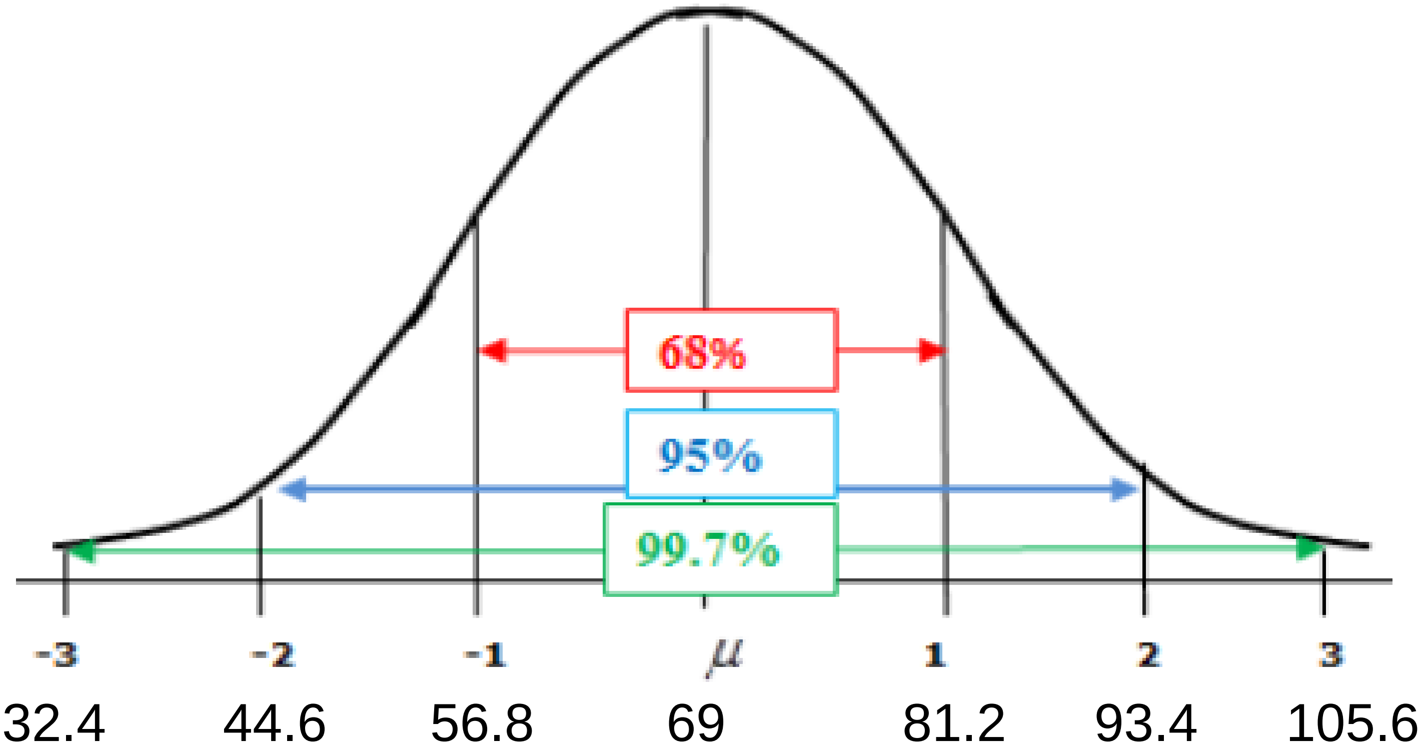
Set A: $n = 24; \bar{x} = 69, \; s_x = 12.2,$

Set B: $n = 22; \bar{x} = 62.8, \; s_x = 15.7$



Normal Distribution

Set A: $n = 24; \bar{x} = 69,\ s_x = 12.2$



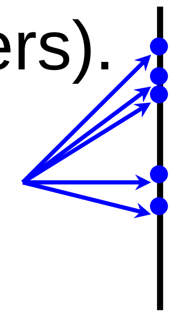
7.1: Random Variables

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

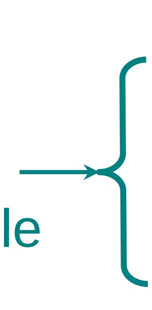
A random variable is **discrete** if its set of possible values is a collection of isolated points on the number line (usually integers).

Possible values of a discrete random variable



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

Possible values of a continuous random variable



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 1^{st} head turns up

   Random Variable: The number of the toss that the 1^{st} 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

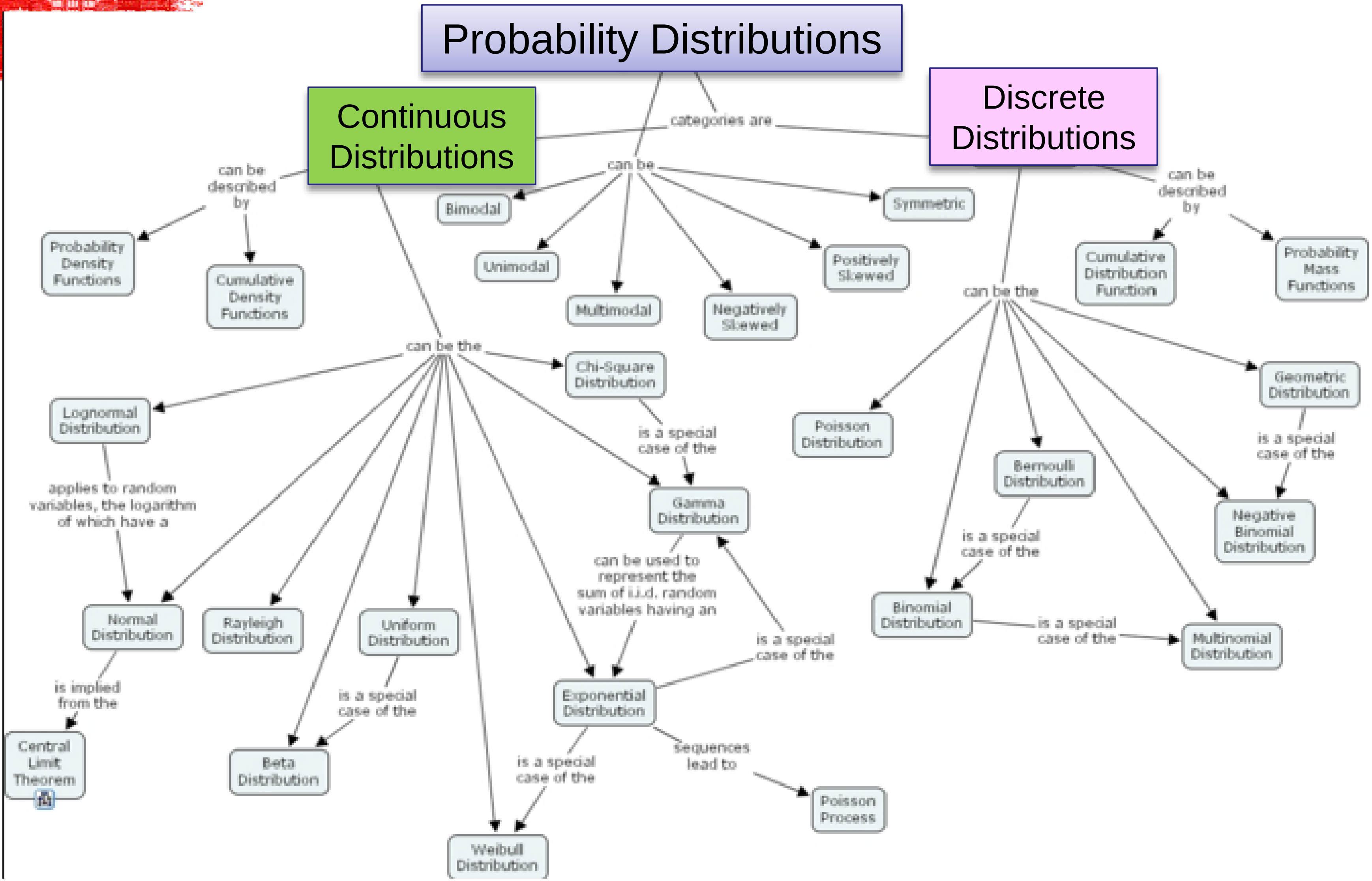
    Random Variable:

    1, if the circuit board is defective

    0, if the circuit board is not defective

    Type: Discrete

Schematic Typology



Common Distributions for Statistics

<u>Discrete Distributions</u>

- Binomial Distributions
- Geometric Distributions
- Poission Distributions (*future stats classes*)

<u>Continuous Distributions</u>

- Normal Distributions
- Chi-Square Distributions
- Student’s *t* Distributions (*t* Distributions)

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?

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

What is the probability that the sum is 9?

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

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.

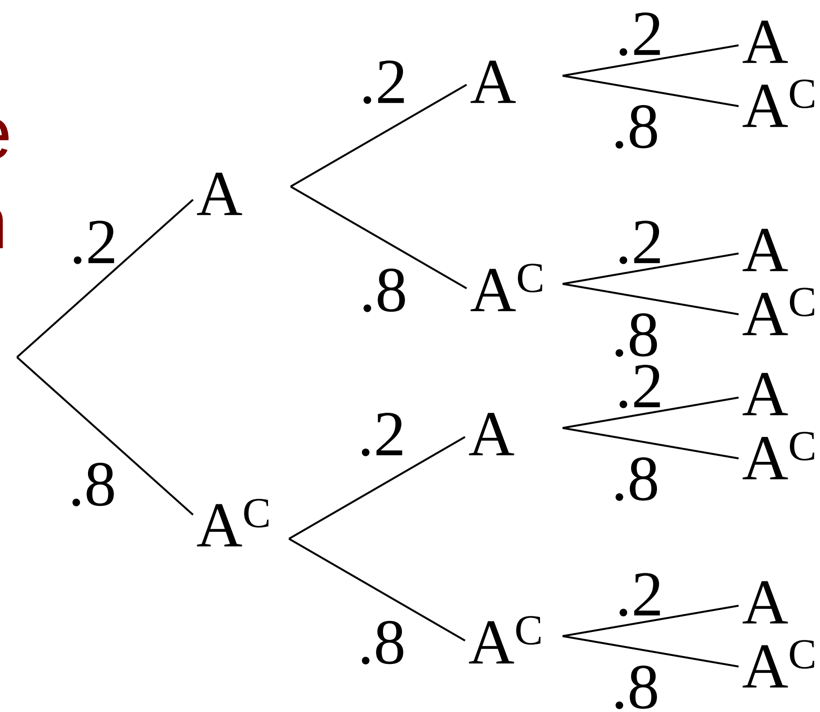
Each probability is the limiting relative frequency of occurrence of the corresponding x value when the experiment is repeatedly performed (LOLn).

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



	X
.2•.2•.2=.008	3
.2•.2•.8=.032	2
.2•.8•.2=.032	2
.2•.8•.8=.128	1
.8•.2•.2=.032	2
.8•.2•.8=.128	1
.8•.8•.2=.128	1
.8•.8•.8=.512	0

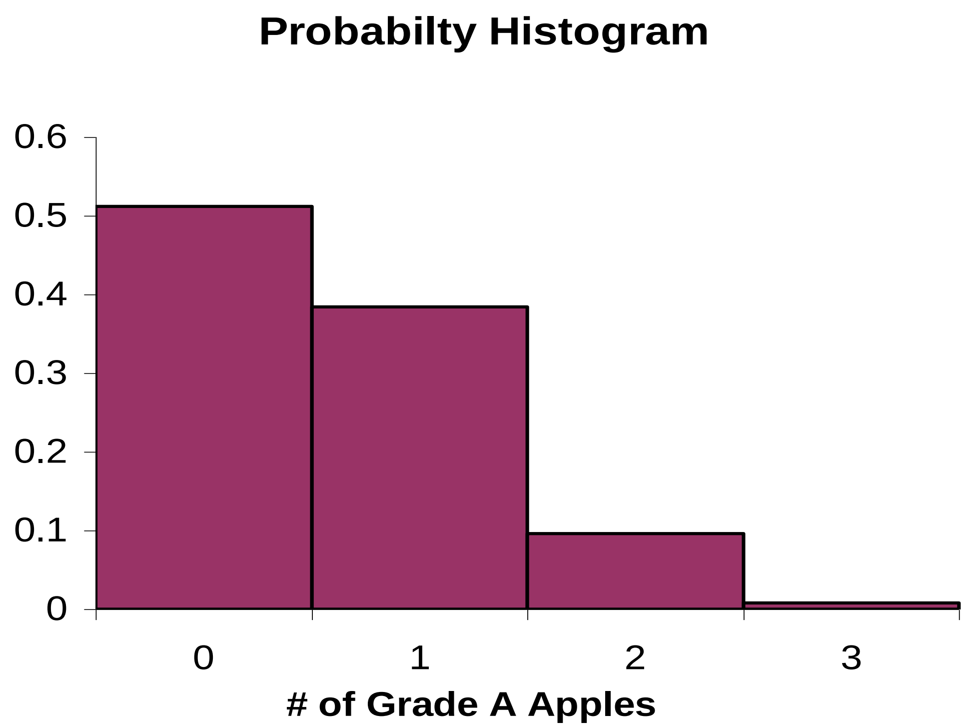
The Results in Table Form

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. $p_1 + p_2 + \ldots + p_k = 1$

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

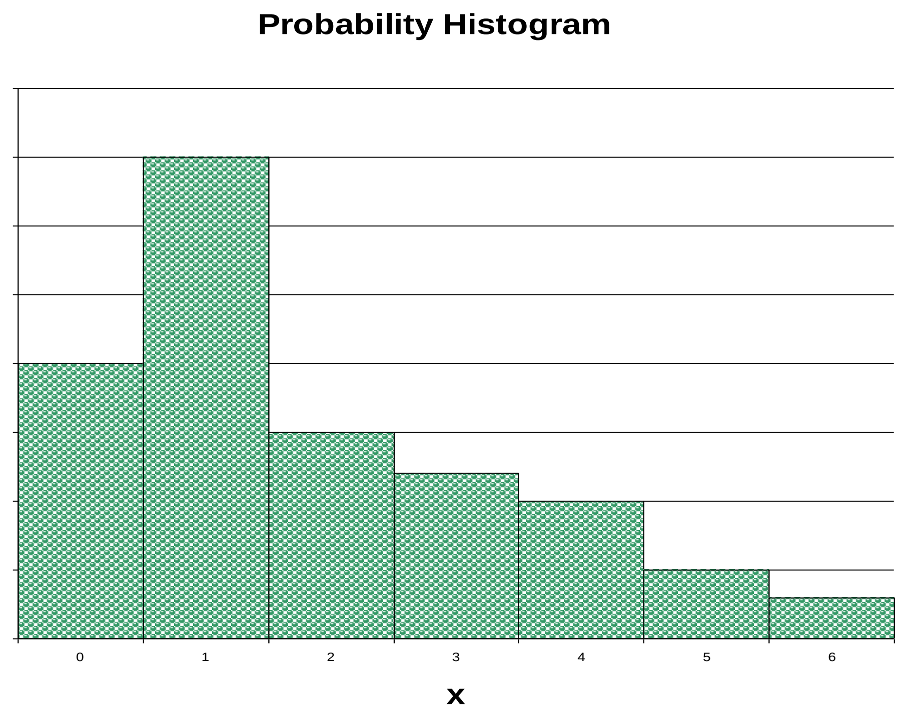
$P(X \leq 3) = 0.20 + 0.35 + 0.15 + 0.12 = \mathbf{0.82}$

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.



The **Chuck-a-LUCK** Game

Costs $1 to play   Dice game of chance

Pick any number from 1 to 6. Roll the three dice.

If you chosen number comes up, you WIN!

If your number comes up 1 time = $1

If your number comes up twice = $2

If your number comes up thrice = $3

Is this a fair game? What is the expected value?

The **Chuck-a-LUCK** Game

Costs $1 to play   Dice game of chance

Pick any number from 1 to 6. Roll the three dice

We must define the random variable …X is?

***X**: the number of occurrences of my chosen number (1 to 6) on the dice*

Now, create a probability table...

X	0	1	2	3
P(X)				

The **Chuck-a-LUCK** Game

What are the probabilities of the table?

X	0	1	2	3
P(X)	$\frac{125}{216}$	$\frac{25}{216}$	$\frac{5}{216}$	$\frac{1}{216}$

Is this a valid probability distribution ?

$$\mathrm{P}(X) = P(x_1) + P(x_2) + P(x_3) + \cdots + P(x_n) = 1$$

$$P(X) = \frac{125}{216} + \frac{25}{216} + \frac{5}{216} + \frac{1}{216} \overset{?}{\Rightarrow} 1$$

The **Chuck-a-LUCK** Game

What are the probabilities of the table?

X	0	1	2	3
P(X)	$\frac{125}{216}$	$\frac{75}{216}$	$\frac{15}{216}$	$\frac{1}{216}$

What is the Expected Value $E(X)$ or μ_x ?

$$E(X)\ or\ \mu_x = \sum x \cdot p(x)$$

$$E(X) = 0 \cdot \frac{125}{216} + 1 \cdot \frac{75}{216} + 2 \cdot \frac{15}{216} + 3 \cdot \frac{1}{216} = 0.5$$

NEXT TOPICS

Chapter 7.3 Continuous Random Variables □

Chapter 7.4 Transforming and Combining Random Variables

Chapter 7.5 Binomial and Geometric Random Variables

The HOME stretch!