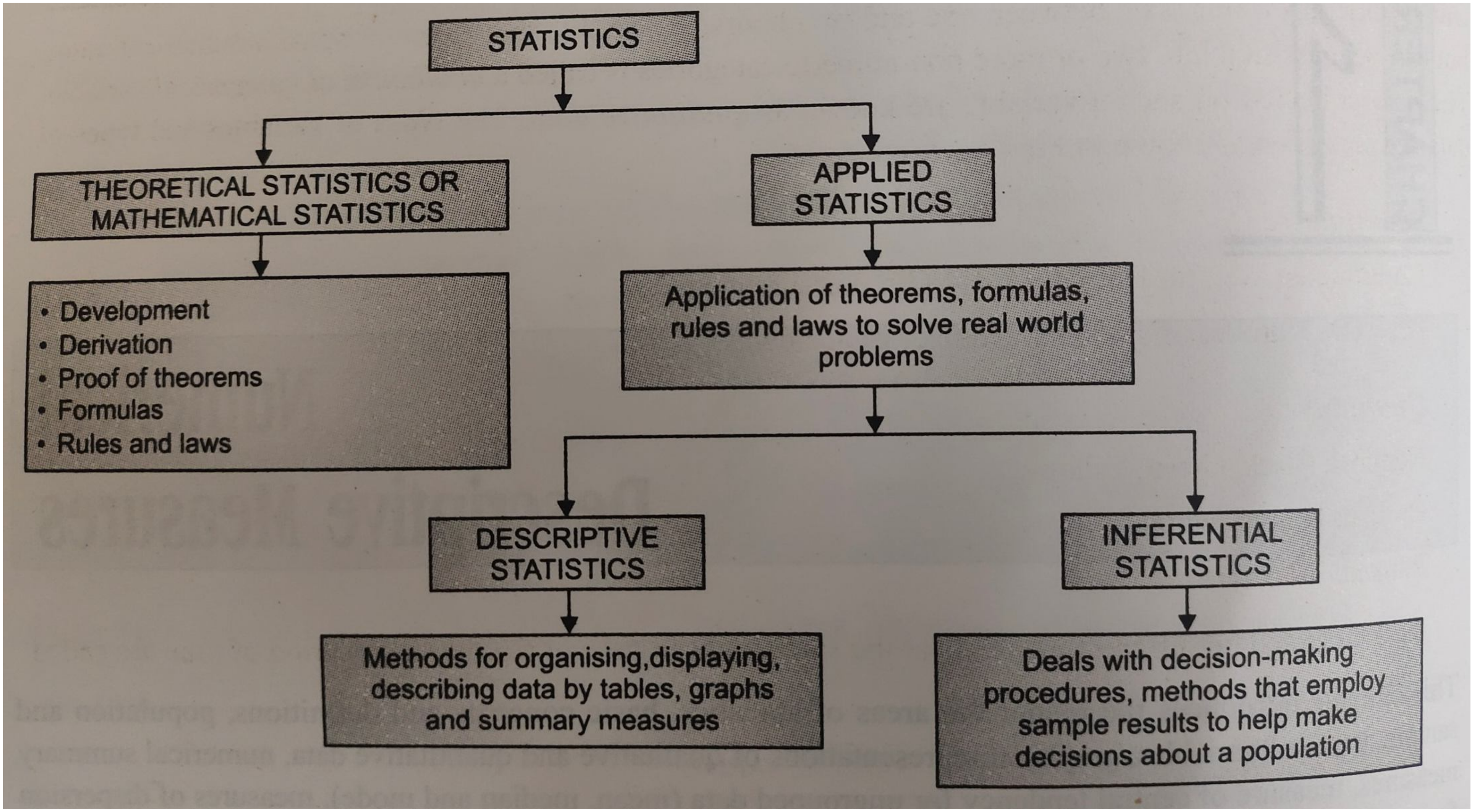
Probability Theory & Random Processes

Unit II
Random Variables
&
Probability Distribution

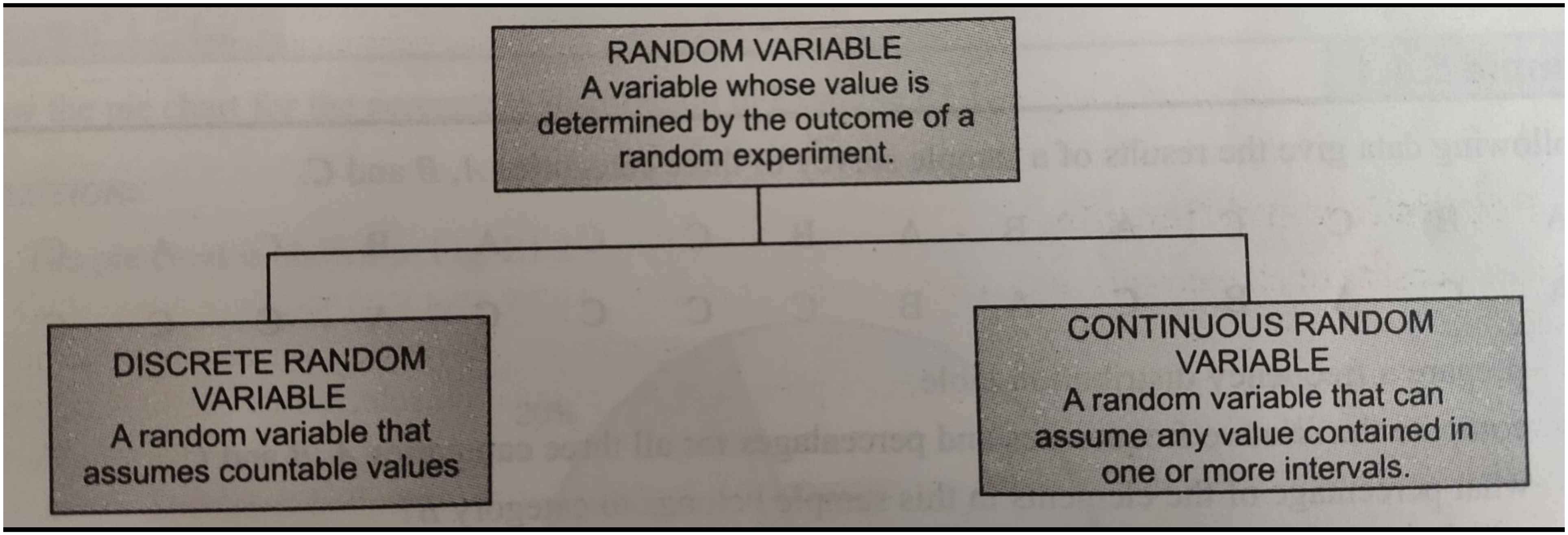
Lecture objectives

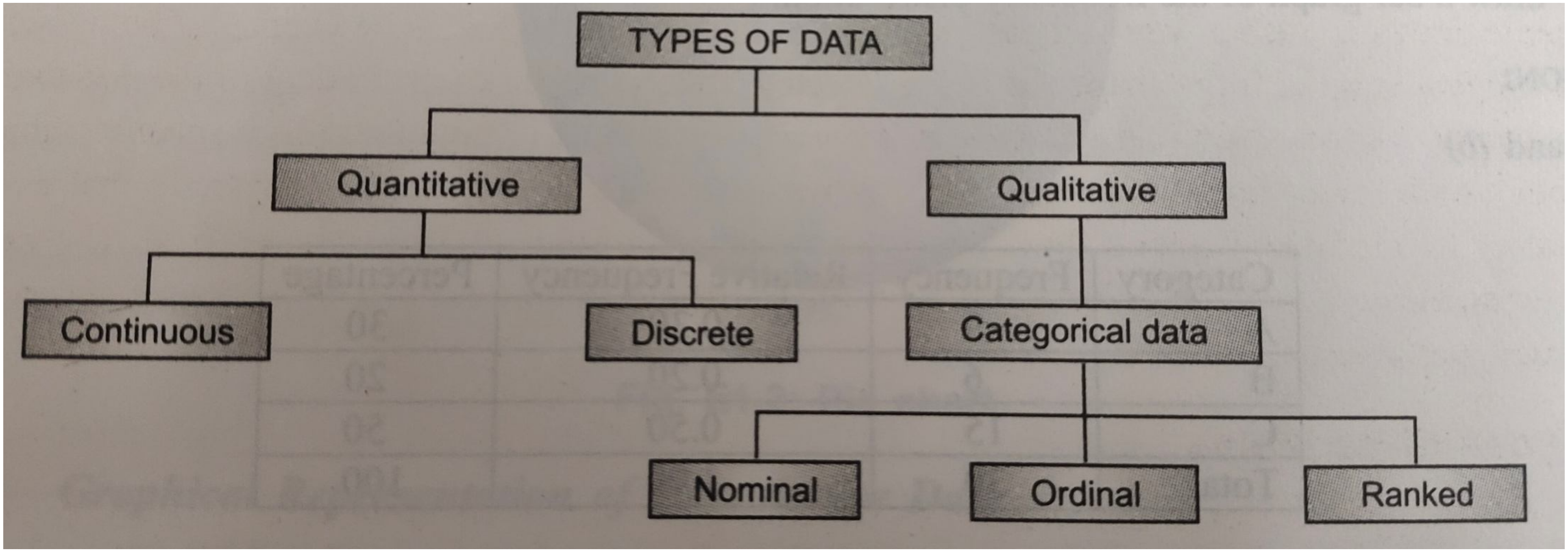
- Basic concepts of random variables (Discrete and continuous), and their mean & Standard Deviation, and probability distributions.
- Discrete probability distribution (Hypergeometric, binomial and Poisson distribution) and their mean & S.D.
- Continuous probability distribution (Normal distribution), its properties, mean and variance

Statistics: Group of methods to collect, analyze, present & interpret data to make decisions.



•The collection of all elements of interest is called population,
•The selection of few elements from population is called a sample.
•Statistics that deals with making decision, inferences prediction and forecasts about populations based on results obtained from samples.





•**Data: Categorical & Numerical (Discrete, continuous)**
•Continuous: Weight, Height
•Discrete: Number of students, Number of accidents
•Nominal (No ranking, lowest level of measurement): Black, Green, Yellow, Marital status, Gender
•Ordinal (ranking is implied): High, medium, low, student's grades
Ranked: 1,2,3
Ratio (Highest level of measurement) differences between the measurement is a meaningful quantity, measurement have a true zero point: Age, Weight, Kelvin scale
Interval scale , differences between the measurement is a meaningful quantity, measurement do not have true zero point: Year, Temp.

Discrete and Continuous Random Variable

A Random Variable taking only some fixed values is called **Discrete Random Variable**.

Example : All the random variables defined in the examples 1,2 and 3.

A Random Variable which can take any value between two extremes is called **Continuous Random Variable**.

Example: Duration of phone call.

Time of all runners in 100 m race.

Age

In most of the practical problems **discrete variables** represent countable data such as **number of defective units in a lot.**

The **continuous variables** represent the measurable data such as all possible **heights, ages, time intervals** etc.

Random Variable

The outcomes of an experiment may be numeric or non numeric .

Example:

Numeric: Rolling a die, Marks obtained in subject, CGPA

Non Numeric: Tossing a coin, Grade in a subject, Quality

To arrive at logical conclusion, we assign numerical values to non numeric outcomes .

***Definition:* Random variable (RV)**

A function which associates a unique real number within an interval to each outcome of sample space of an experiment is known as random variable.

Definition: A random variable (abbreviatively RV) is a function that assigns a real number $X(s)$ to every element $s \in S$, where S is the sample space corresponding to a random experiment E.)

Random variable

- A variable which contains the outcomes of a chance experiment
- "Quantifying the outcomes"
- Example X= (1 = Head, 0 = Tails)
- A variable that can take on different values in the population according to some "random" mechanism
- Discrete
  - Distinct values, countable
  - Year
- Continuous
  - Mass

Examples of Random Variable

Example 1. Toss of a coin

Sample space $S = \{H, T\}$

We define random variable as $X(H) = 1$, $X(T) = 0$

Range $R = \{ 0, 1\}$

Example 2. Three coins are tossed together.

$S = \{HHH, HHT, HTH, THH, THT, HTT, TTH, TTT \}$

We define the random variable X = *"Number of heads"*

$X(HHH) = 3$, $X(HHT) = 2$, $X(HTH) = 2$, $X(THH) = 2$
$X(HTT) = 1$, $X(THT) = 1$, $X(TTH) = 1$, $X(TTT) = 0$

Range $R = \{ 0, 1, 2, 3\}$

Example 3. To compute CGPA grades are assigned numerical values.

What is a distribution?

- Describes the ‘shape’ of a batch of numbers

- The characteristics of a distribution can sometimes be defined using a small number of numeric descriptors called ‘parameters’

Why distribution?

- Can serve as a basis for standardized comparison of empirical distributions
- Can help us estimate confidence intervals for inferential statistics
- Form a basis for more advanced statistical methods
  - ‘fit’ between observed distributions and certain theoretical distributions is an assumption of many statistical procedures

Probability Distributions

- The probability distribution function or probability density function (PDF) of a random variable X means the values taken by that random variable and their associated probabilities.

- PDF of a discrete r.v. (also known as PMF):

Example 1: Let the r.v. X be the number of heads obtained in two tosses of a coin.

Sample Space: {HH, HT, TH, TT}

Probabilities assigned to various outcomes in S in turn determine probabilities associated with the values of any particular RV *X. The probability distribution of X says* how the total probability of 1 is distributed among (allocated to) the various possible *X values.*

p(0) = the probability of the X value 0 = P (X =0)
p(1) = the probability of the X value 1 = P (X =1)

Probability Function

If X is a discrete RV which can take the values $x_1, x_2, x_3, \ldots$ such that $P(X = x_i) = p_i$, then p_i is called the *probability function or probability mass function or point probability function*, provided p_i $(i = 1, 2, 3, \ldots)$ satisfy the following conditions:

(i) $p_i \geq 0$, for all i, and

(ii) $\sum_i p_i = 1$

The collection of pairs $\{x_i, p_i\}$, $i = 1, 2, 3, \ldots$, is called *the probability distribution of the* RV X, which is sometimes displayed in the form of a table as given below:

$X = x_i$	$P(X = x_i)$
x_1	p_1
x_2	p_2
$\vdots$	$\vdots$
x_r	p_r
$\vdots$	$\vdots$

Example: Consider that a fair coin is tossed three times.

Then sample space is

$S = \{HHH, HHT, HTH, THH, THT, HTT, TTH, TTT\}$

We define $X =$ Number of heads in each sample point

The range is {0, 1, 2, 3}

We find that $P[X = 0] = 1/8$, $P[X = 1] = 3/8$, $P[X = 2] = 3/8$, $P[X = 3] = 1/8$

Therefore the **probability distribution** is $\{0, 1/8\}, \{1, 3/8\}, \{2, 3/8\}, \{3, 1/8\}$

The probability distribution can also be given an tabular form as

x	P[X=x}]
0	1/8
1	3/8
2	3/8
3	1/8

Discrete Distribution -- Example

Distribution of Daily Crises	
Number of Crises	Probability
0	0.37
1	0.31
2	0.18
3	0.09
4	0.04
5	0.01

Requirements for a Discrete Probability Function

- Probabilities are between 0 and 1, inclusively

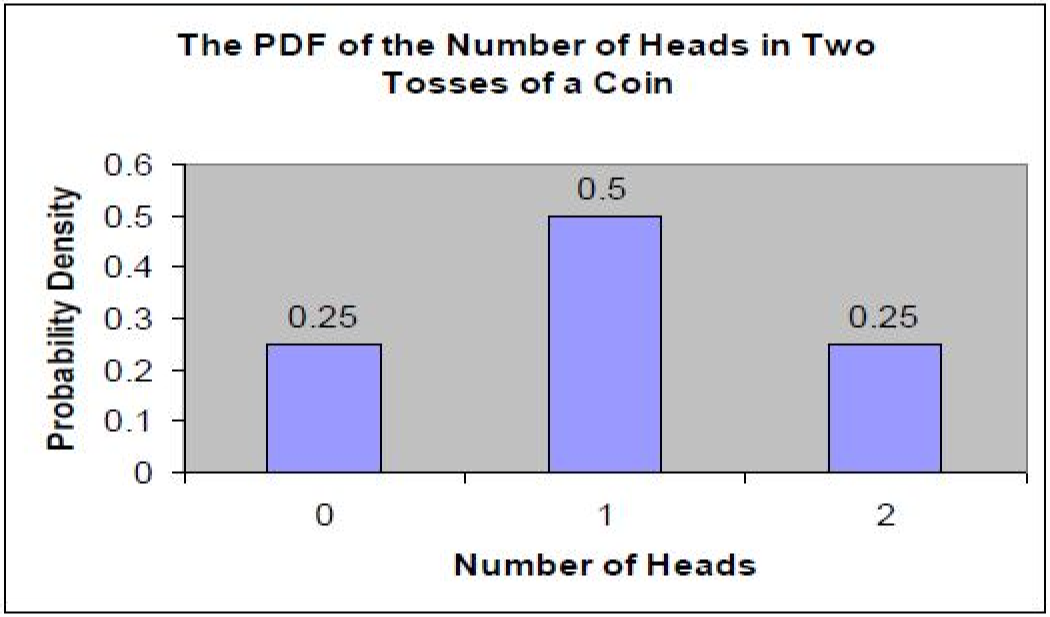
- Total of all probabilities equals 1

$$\boxed{0 \le P(X) \le 1 \quad \text{for all X}}$$

$$\boxed{\sum_{\text{over all x}} P(X) = 1}$$

PDF of Discrete r.v.

Number of Heads (X):	0	1	2	sum
PDF ($P(X)$):	¼	½	¼	1



The Cal Poly Department of Statistics has a lab with six computers reserved for statistics majors. Let X denote the number of these computers that are in use at a particular time of day. Suppose that the probability distribution of X is as given in the following table; the first row of the table lists the possible X values and the second row gives the probability of each such value.

x	0	1	2	3	4	5	6
$p(x)$	.05	.10	.15	.25	.20	.15	.10

Probability that at most 2 computers are in use is

$$P(X \leq 2) = P(X = 0 \text{ or } 1 \text{ or } 2) = p(0) + p(1) + p(2) = .05 + .10 + .15 = .30$$

Since the event *at least 3 computers are in use is complementary to at most 2 computers are in use,*

$$P(X \geq 3) = 1 - P(X \leq 2) = 1 - .30 = .70$$

The probability that between 2 and 5 computers *inclusive are in use is*

$$P(2 \leq X \leq 5) = P(X = 2, 3, 4, \text{ or } 5) = .15 + .25 + .20 + .15 = .75$$

Probability that the number of computers in use is *strictly between 2* and 5 is

$$P(2 < X < 5) = P(X = 3 \text{ or } 4) = .25 + .20 = .45$$

Six lots of components are ready to be shipped by a certain supplier. The number of defective components in each lot is as follows:

Lot	1	2	3	4	5	6
Number of defectives	0	2	0	1	2	0

One of these lots is to be randomly selected for shipment to a particular customer. Let X be the number of defectives in the selected lot. The three possible X values are 0, 1, and 2. Of the six equally likely simple events, three result in $X = 0$, one in $X = 1$, and the other two in $X = 2$. Then

$$p(0) = P(X = 0) = P(\text{lot 1 or 3 or 6 is sent}) = \frac{3}{6} = .500$$

$$p(1) = P(X = 1) = P(\text{lot 4 is sent}) = \frac{1}{6} = .167$$

$$p(2) = P(X = 2) = P(\text{lot 2 or 5 is sent}) = \frac{2}{6} = .333$$

Probability Distribution for the Random Variable *X*

A probability distribution for a discrete random variable X:

x	–8	–3	–1	0	1	4	6
$P(X = x)$	0.13	0.15	0.17	0.20	0.15	0.11	0.09

Find

a. $P(X \le 0)$ $\boxed{0.65}$

b. $P(-3 \le X \le 1)$ $\boxed{0.67}$

Cumulative Distribution Function

- The CDF of a random variable X (defined as $F(X)$) is a graph associating all possible values, or the range of possible values with $P(X \leq x)$.
- CDFs always lie between 0 and 1 i.e., $0 \leq F(X_i) \leq 1$, Where $F(X_i)$ is the CDF.

Discrete Distribution Function

Definition : The distribution function of a discrete random variable X with probability (mass) function $p(x_i) = p_i \, , \, i = 1,2,3,......$ is defined as

$$F(x) = P[X \le x] = \sum_{i:x_i \le x} p_i$$

Note that $p_i = p(x_i) = P[X = x_i] = F(x_i) - F(x_{i-1})$

Example 1 : If X is a discrete random variable with probability function

X = x	0	1	2	3
P[X=x]	1/8	3/8	3/8	1/8

Find P[X ≤ 2]

Answer :

$$F(x) = P[X \le x] = \sum_{i:x_i \le 2} p_i$$

$$= \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$$

Example 1 : If X is a discrete random variable with probability function

$X = x$	-2	-1	0	1
$P[X=x]$	0.4	k	0.2	0.3

Find k, $P[X<0]$, $P[0 \leq X]$ and distribution function of X.

Answer :

We know that $\Sigma p(x) = 1$, therefore

$0.4 + k + 0.2 + 0.3 = 1$

$k = 0.1$

$P[X<0] = p(-2)+p(-1) = 0.4 + 0.1 = 0.5$

$P[0 \leq X] = p(0) + p(1) = 0.2 + 0.3 = 0.5$

The distribution function of X

$$F(x) = \begin{cases} 0, & x < -2 \\ 0.4 & -2 \leq x < -1 \\ 0.5 & -1 \leq x < 0 \\ 0.7 & 0 \leq x < 1 \\ 1.0 & 1 \leq x \end{cases}$$

A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory. The accompanying table gives the distribution of $Y =$ the amount of memory in a purchased drive:

y	1	2	4	8	16
$p(y)$	.05	.10	.35	.40	.10

Let's first determine $F(y)$ for each of the five possible values of Y:

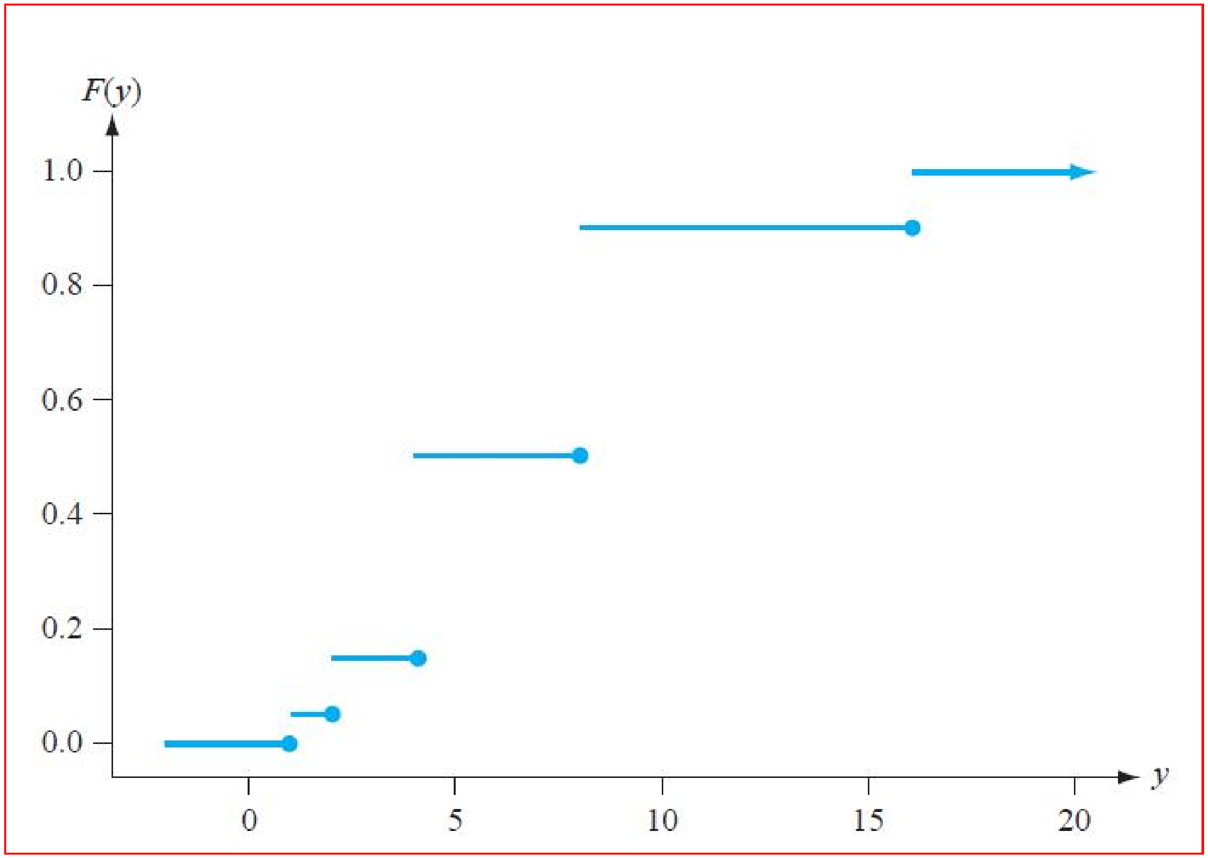
$$\begin{aligned}
F(1) &= P(Y \le 1) = P(Y = 1) = p(1) = .05 \\
F(2) &= P(Y \le 2) = P(Y = 1 \text{ or } 2) = p(1) + p(2) = .15 \\
F(4) &= P(Y \le 4) = P(Y = 1 \text{ or } 2 \text{ or } 4) = p(1) + p(2) + p(4) = .50 \\
F(8) &= P(Y \le 8) = p(1) + p(2) + p(4) + p(8) = .90 \\
F(16) &= P(Y \le 16) = 1
\end{aligned}$$

Now for any other number y, $F(y)$ will equal the value of F at the closest possible value of Y to the left of y. For example,

$$\begin{aligned}
F(2.7) &= P(Y \le 2.7) = P(Y \le 2) = F(2) = .15 \\
F(7.999) &= P(Y \le 7.999) = P(Y \le 4) = F(4) = .50
\end{aligned}$$

Hence, a graph of cdf is:

$$F(y) = \begin{cases} 0 & y < 1 \\ .05 & 1 \le y < 2 \\ .15 & 2 \le y < 4 \\ .50 & 4 \le y < 8 \\ .90 & 8 \le y < 16 \\ 1 & 16 \le y \end{cases}$$



For X a discrete rv, the graph of $F(x)$ will have a jump at every possible value of X and will be flat between possible values. Such a graph is called a **step function.**

Probability Density function & Cumulative Distribution Function

***Definition* :** Consider a continuous function *f(x)*, *f(x)* ≥ 0

$$P[X \leq x] = F(x) = \int_{-\infty}^{x} f(x)dx$$

The *f(x)* is known as **Probability Density Function (PDF)** of continuous RV *X*.

F(x) is the **Cumulative Distribution Function** of the RV.

$$\frac{d}{dx}F(x) = f(x) \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

Hence,

$$P[a < X < b] = \int_{a}^{b} f(x)dx$$

$$P[X = x] = 0$$

Probability of continuous random variable at a point is zero.

Example 1 : Verify if whether $f(x)=\begin{cases}|x|, -1\le x\le 1\\ 0, \text{otherwise}\end{cases}$ can be a PDF of a continuous random variable.

Answer : For *f(x)* to be a PDF it must satisfy

$$f(x)\ge 0, \forall x \qquad \int_{-\infty}^{\infty} f(x)dx = 1$$

First condition is satisfied since $|x| \ge 0$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{1}|x|dx = 2\int_{0}^{1}|x|dx = 2\int_{0}^{1}xdx = 2\left|\frac{x^2}{2}\right|_0^1 = 1$$

Hence, second condition is also satisfied.

Therefore, f(x) can be a PDF for a random variable

Example 2 : Probability density function of a random variable is given by

$$f(x)=\begin{cases} cxe^{-x}, x>0 \\ 0, x\le 0 \end{cases}$$

Find the value of c and Cumulative distribution function (CDF) of x.

Answer:

If *f(x)* is a PDF, Then: $\int_{-\infty}^{\infty} f(x)dx=1$

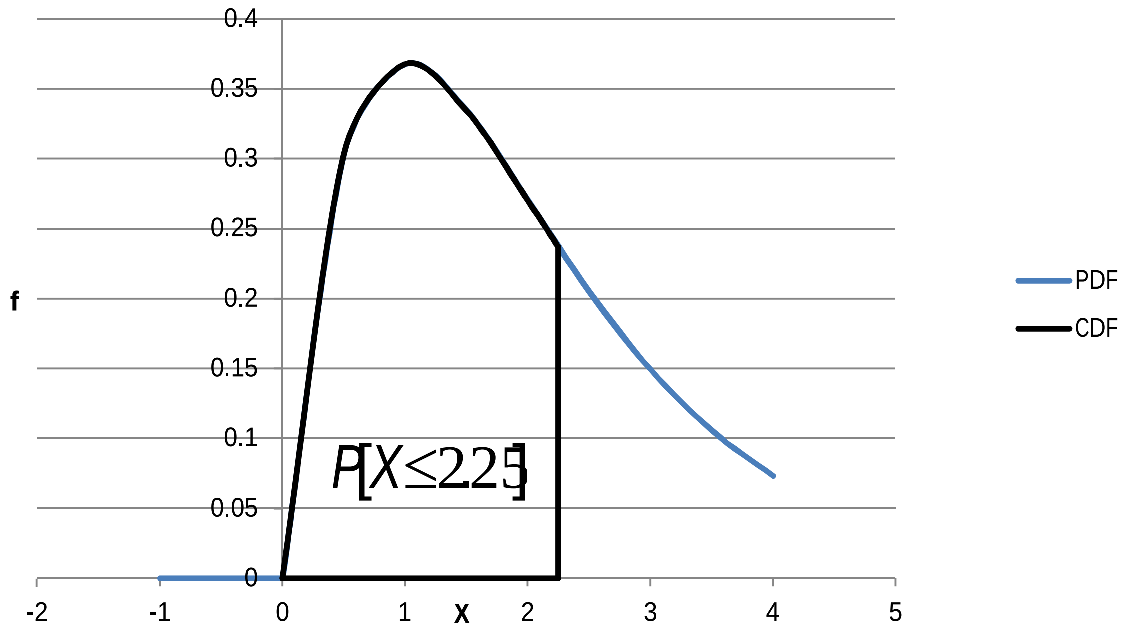
$$\int_0^{\infty} cxe^{-x}dx=1 \qquad c\left[x\frac{e^{-x}}{-1}-\int 1\frac{e^{-x}}{-1}dx\right]_0^{\infty}=1 \qquad c\left[x\frac{e^{-x}}{-1}+\frac{e^{-x}}{-1}\right]_0^{\infty}=1 \qquad c=1$$

The CDF of x = $F(x)$ =$P[X\le x]$ = $\int_{-\infty}^{x} f(x)dx=\int_0^{x} xe^{-x}dx=\left[x\frac{e^{-x}}{-1}+\frac{e^{-x}}{-1}\right]_0^{x}$

$$F(x)=\begin{cases} 1-(1+x)e^{-x}, x>0 \\ 0, \text{otherwise} \end{cases}$$

Continuous Random Variable

PDF: $f(x) = \begin{cases} xe^{-x}, x > 0 \\ 0, \ x \leq 0 \end{cases}$ CDF: $F(x) = P[X \leq x] = \begin{cases} 1-(1+x)e^{-x}, x > 0 \\ 0, \text{otherwise} \end{cases}$



Graph for PDF and CDF

PDF: $f(x) = \begin{cases} xe^{-x}, x > 0 \\ 0, \ x \le 0 \end{cases}$

CDF: $F(x) = P[X \le x] = \begin{cases} 1-(1+x)e^{-x}, x > 0 \\ 0, \text{otherwise} \end{cases}$