

**Business Analytics for Management Decision**  
**Prof. Rudra P Pradhan**  
**Vinod Gupta School of Management**  
**Indian Institute of Technology, Kharagpur**

**Lecture – 40**  
**Predictive Analytics: Simulation**

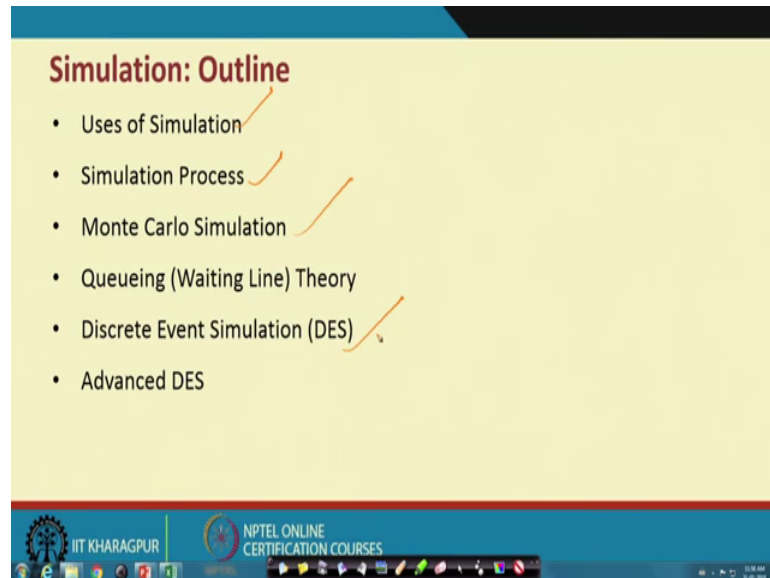
Hello everybody this is Rudra Pradhan here, welcome to B M D lecture series today we will continue with predictive analytics and that to coverage on simulation. In fact, we have already discussed couple of techniques in this particular you know unit that is on predictive analytics starting with you know association technique, causality technique, machine learnings and the data mining. So, this is also similar kind of you know predictive analytics structure through which you can understand the problems get insights from the data then develop a system through which you can do the predictions and do the kind of you know forecasting as per the problem requirement and the kind of you know management requirement.

In fact, in a real life kind of you know situations and the kind of you know business situations you will find you know there are lots of complexity, there are lots of dynamics and with respect to you know availability of you know data and the kind of you know understanding, you have to develop a kind of you know systems or you know develop a kind of you know structures or develop a kind of you know algorithm through which you can do you know better predictions, better forecasting as per the management requirement and in this context simulation is a kind of you know effective tools or you know effective technique and through which you can you know understand a particular you know business problem and then get you know better insights a after the particular you know structure and then you can go for the kind of you know prediction and you know for castings.

So, in this lectures we will highlight this particular you know component that is the simulation and then we will discuss how it is you know kind of an effective tool through which you can understand a kind of you know business problem and get the insights from the data and then connect the particular you know structure through which you can do the predictions for the particular you know problem. In the simulations lots of things are there and it depends upon a you know kind of you know situation because you know

risks and the kind of you know uncertainty are always there in a kind of you know business process.

(Refer Slide Time: 02:34)



So, that is why with respect to a particular you know structures or the kind of you know particular you know requirement, we have to develop a you know kind of you know system or you know structure. So, that you know some kind of you know better prediction and some kind of you know better insights we can actually generate for the particular you know business problem or you know particular you know business requirement. In fact, in the simulations so some of the things are you know you uses of you know simulation which I already.

In fact, you know highlighted and simulation process then Monte Carlo simulations then Quinn's theory and the discrete event simulations then advanced discrete event simulation; that means, actually like you know data mining structure simulation engine you know not a kind of in a small component it is a very big component and very attractive component for the kind of you know business environment or the kind of you know management environment because most of the business problems and the kind of you know management problems are very dynamic with you know plenty of you know risk uncertainty.


So, with the simulations we can actually streamline the particularly you know process and then developer a kind of you know system and the kind of you know structure

through which you can actually control the problems predict the problem as per the particular you know management requirement.

(Refer Slide Time: 04:05)

**Simulation**

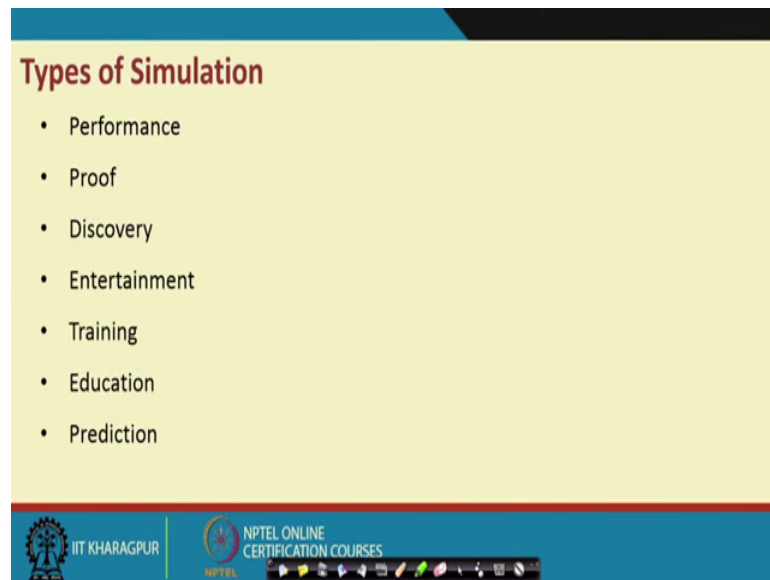
- Process of modeling reality to gain a better understanding of the phenomena or system being studied
- Simulation versus the “real world”
  - More cost effective
  - Less dangerous environment
  - Faster
  - More practical
- Does not require mathematical models or computers



IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, that is how simulation you know understanding the simulation and then use the simulation is very you know interesting for the kind of you know business analytics and first of all we would like to understand what is exactly the simulation then we connect with you know structure through which you can you know get to know how simulation can be you know very effective tools for the business requirement or the kind of you know business analytics and it is a process of modeling a reality to gain a better understanding of the kind of you know situation or system being studied.

(Refer Slide Time: 04:34)



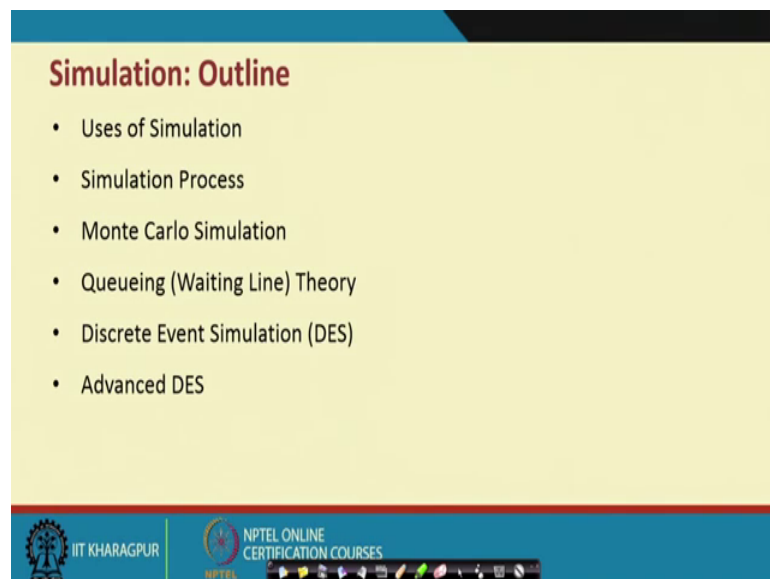
**Types of Simulation**

- Performance
- Proof
- Discovery
- Entertainment
- Training
- Education
- Prediction

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So; that means, actually a particular system to be you know understand to be kind of you know explored, but the thing is that you know you should first understand the particular you know structure through which actually the kind of you know structure you can you know develop.

(Refer Slide Time: 04:48)



**Simulation: Outline**

- Uses of Simulation
- Simulation Process
- Monte Carlo Simulation
- Queueing (Waiting Line) Theory
- Discrete Event Simulation (DES)
- Advanced DES

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, a simulations means you know we like to know the kind of predictive structure with you know actual scenarios so; that means, what is the actual fact the real world scenario and then how is the simulated structure; that means, we like again it is a kind of

predictive structure and through a particular you know situations and then we will compare with the actual structure with the you know predictive structure. So, more cost effective and it is more dangerous you know to the kind of you know means less dangerous to the kind of you know environment because it will give you some kind of you know setups through which you can you know monitor the process and then you know kind of you know understand the kind of you know system through which you can actually predict the business you know requirement.

So, it actually learns very fasters like you know the support vector machines or random forest. So, it is actually more practical sometimes you know because of kind of you know process which you like to develop and the kind of know technique itself understanding the insights and get you know predicted environment as per the particular you know business requirement and in fact, a the particular you know tool does not require mathematical models or the kind of you know structures.

But just you know have a simple kind of you know structure through which you can actually generate the kind of you know predictive kind of you know environment for instance actually a simple excel say you know excel spread sheet can help you to get the simulated results if you are actually you know the particular you know effect and the kind of you know insights so; that means, say again same things you understand the problems understand the insights then you know connect a particular you know model and the kind of you know tools through which you can go for better prediction better insights and the better kind of you know management requirement right.

So, with this actually I like to highlight here the types of you know simulations you know with respect to performance through discovery, entertainment, training, education, prediction so; that means, actually by the way I mean it is a very effective tools you know as per the kind of you know dynamic you know business requirement and the kind of you know real life you know reality. So, the thing is that you know simulation process the I means understanding of the simulation process and that too start with the here the kind of you know model development and then in the model development itself you know define the problems or you know put the questions or put the kind of you know structure as you know what is exactly we need.

So, that you know we can actually simulate accordingly until unless you clear about the problem the kind of you know requirement and the kind of you know need. So, we are not in a position to define the problem and we are not in a position to steam simulate the kind of you know thing.

So, that you know the thing can be analysed properly or you know it can be predicted properly. So, develop the conceptual model and then generate the data I means you know we need some kind of you know initial kind of you know inside. So, that you know you can simulate for the future requirement so; that means, to whatever techniques we have till now discussed we need actually some kind of you know historical data or some kind of you know original insights means actual insight or actual information through which actually we develop a structure with different predictive analytics tools and then we will go for the kind of you know new structure through which you can you know do the effective predictions and affective for forecasting as per the particular you know business requirement.

(Refer Slide Time: 08:44)

**Simulation Process**

- Model development ✓
  - Define the problem or question
  - Develop the conceptual model
  - Collect data
  - Build computer model
- Model validation
- Simulate and analyze output

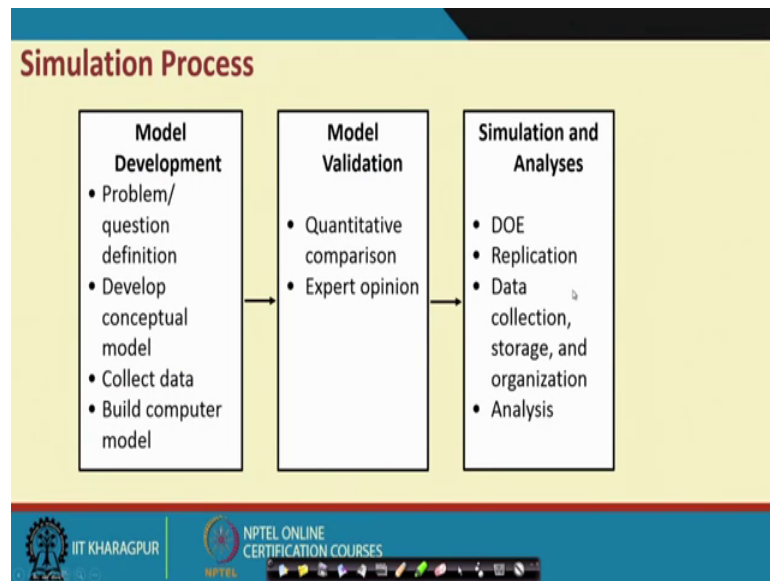
IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, we are in the processor in a simulation is a kind of in a structure through which we can get the kind of you know better structure which we will do the predictions as per the particular you know business requirement then as usual here the process is to develop the models as per the requirement as per the on you know a data availability data

visualization insights then connect with a particular you know setups then validate the particular you know setup then simulate and analyse the output.

So, this is how you know it is not something new. So, it is also unusual the kind of you know structure through which actually you can actually develop and then I know train then test validate and then finally, fix the particular you know structure as per the particular you know requirement.

(Refer Slide Time: 09:27)



And simulation process you know altogether it is of you know 3 different stages, first stage more important is the model development then, model validations, then the simulation it is like you know same techniques which you have discussed like you know neural network random forest support vector machines data mining tools like you know clustering etcetera. So, in the model development, the most important thing is the problem definitions then develop the kind of you know conceptual model and the model conceptual model you know it is not actually once you develop means it is a fixed actually it is a kind of flexible depending upon the structure the requirement it can change then final means it is a kind of a continuous search process all together.

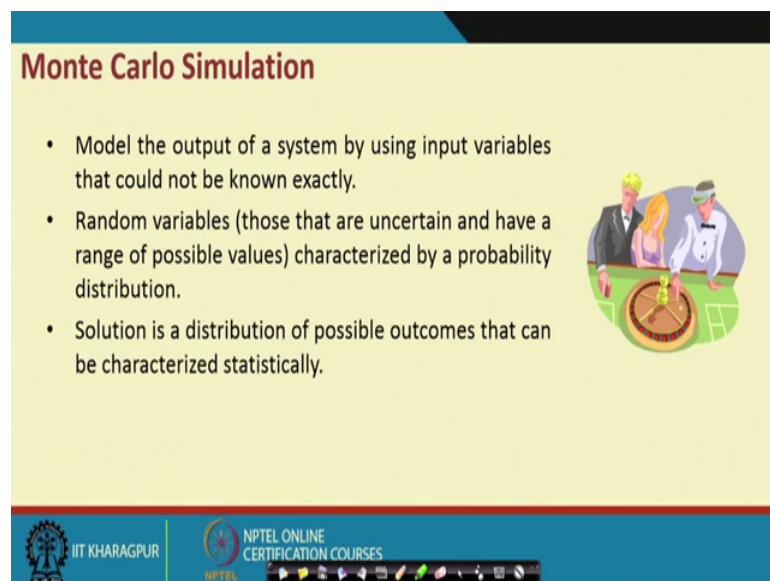
We are you know doing the kind of you know structuring, restructuring, testing, retesting, validating, revalidating then finally, fix up a kind of you know structure through which you can have a better you know system and then we will do the kind of

you know prediction as per the particular you know requirements. So, what I have mentioned the earlier.

So, the problem definitions and develop the conceptual model then have the data and then connect with a kind of you know structure through which you know it is not necessarily the computer you know some structure through which actually you can actually generate the particular you know insights then you know model validations that mean in the kind of you know of quantitative comparisons then sometimes you know you can actually connect or you know valuate through experts opinions and because we have a kind of qualitative you know forecasting structures through which actually we can quantify and develop a quantitative models with the help of you know ting information or in actual information.

Then numerically or in the modeling structure you can also validate and then again still to make you know more effective more efficient you can also connect within expert's opinion and then finally, simulate and then analyse. So, there are lots of you know again process DOE, replication that is a design of experiments then, data collections, storage and organize and then finally, you can you know analyse the particular you know process.

(Refer Slide Time: 11:53)



**Monte Carlo Simulation**

- Model the output of a system by using input variables that could not be known exactly.
- Random variables (those that are uncertain and have a range of possible values) characterized by a probability distribution.
- Solution is a distribution of possible outcomes that can be characterized statistically.

The slide features an illustration of a roulette wheel with a croupier and a player, symbolizing randomness and probability. The footer includes the IIT Kharagpur logo and the text 'NPTEL ONLINE CERTIFICATION COURSES'.

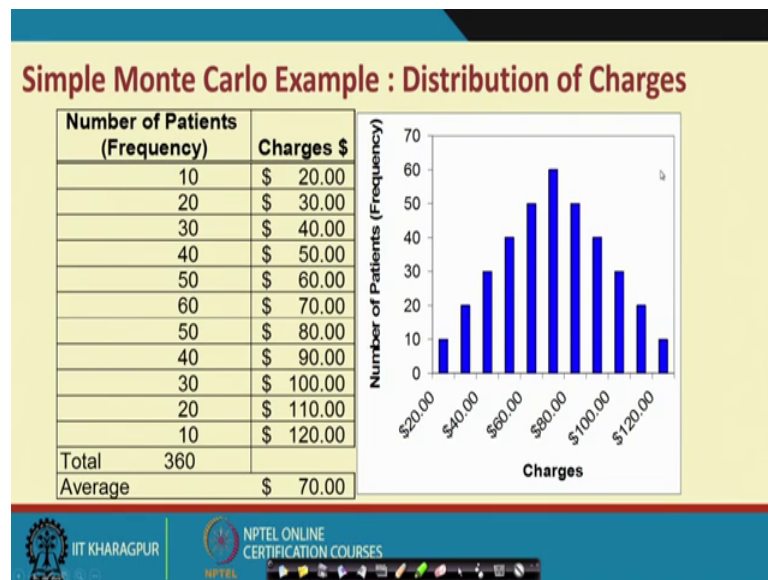
And Monte Carlo simulation is a one of the such you know kind of you know structure through which you know we can do the a you know simulations and then means with the



actual information we get little bit insights and develop the kind of know model and generate or you know simulate for the future requirement.

So, ultimately again, every times there is a input kind of you know things and then the kind of you know response variable the target variable that is the output through to need to be actually predicted and for that actually we are looking for a kind of you know structure or you know kind of an algorithm or you know system through which we can actually simulate a you know the future requirement of a business. So, random variables you know usually a technical use for the kind of you know means a kind of you know probability distribution structure through which you will do the prediction and solution is a distribution of possible outcomes that can be generalized you know statistically.

(Refer Slide Time: 12:50)



This is a simple example and we have a number of patients that is it with respect to frequency and the kind of you knows charges. So, patients are you know paying and they are you know 20000. So, it is it you know 10 and then 30000 then, 20 like that we have a plenty of you know data and then just you plot and then you can get a kind of you know insight and the idea of about this plotting actually to get to know a particular you know distributions.

Now, one of the basic requirement of you know simulation is to connect since it is a with respect to you know risk and uncertainty then; obviously, 2 major components for simulation for this particular predictive analytics is the probability and probability

distribution. We have already gone through so, many in distribution starting with you know binomial you know Poisson distribution, normal distributions generating distributions. So, like that you know a uniform distribution. So, we have lots of you know distribution generally the simulation is actually done through the particular you know distribution.

So, having a particular you know structure or you know actual information of a particular problem. So, we like to first you know get insight and understand the system and then take you know how these information are actually fitted with a particular you know distribution. Once you pick up a particular distributions with respect to the actual availability or actual information availability then you know easily it through that particular you know distribution you can actually simulate for the future requirement that is the basic [FL] about the simulation and that is how first you understand then you know having the means the thing is that you know without having the initial insights or you know initial you know informations that is how the you know data is a kind of you know one of the first and requirement for the simulations and then the kind of you know management requirement.

So, it has actually 2 step process altogether first step you having a the actual information you understand the particular you know structure check how these information are you know connected with a particular you know distribution, then using that particular distribution you can simulate for the future requirement and that too the kind of you know management requirement or the problem requirement.

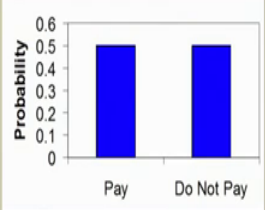
(Refer Slide Time: 15:11)

### Simple Monte Carlo Example

- Fifty percent of the clinic's patients do not pay for their services, and it is equally likely that they will pay or not pay.
- The payment per patient is modeled by:  

$$\text{Probability of payment} \times \text{Charges/patient} = \text{Payment/patient}$$
- A deterministic solution to this problem would be:  

$$0.5 \times \$70/\text{patient} = \$35 \text{ per patient}$$



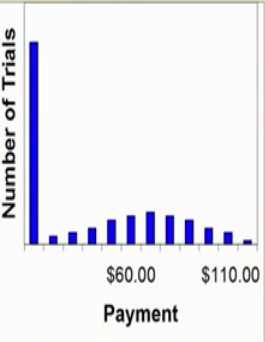
IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, this is another way of you know visualization. So, it is a kind of, you know so; that means, here we like to have a payment structure and some are paying not paying and then we have we like to know how much probability is you know for the paying case and non-paying case. So, this is a simple case of you know structuring.

(Refer Slide Time: 15:33)

### Simple Monte Carlo Example : Payment Distribution

Trial #	Coin Flip	Payment	Die Total	Charges	Patient Payment
1	H	1	7	\$ 70.00	\$ 70.00
2	T	0	10	\$ 100.00	\$ -
3	H	1	8	\$ 80.00	\$ 80.00
4	T	0	8	\$ 80.00	\$ -
5	H	1	9	\$ 90.00	\$ 90.00
6	T	0	8	\$ 80.00	\$ -
7	H	1	7	\$ 70.00	\$ 70.00
8	T	0	10	\$ 100.00	\$ -
9	H	1	9	\$ 90.00	\$ 90.00
10	T	0	10	\$ 100.00	\$ -



IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

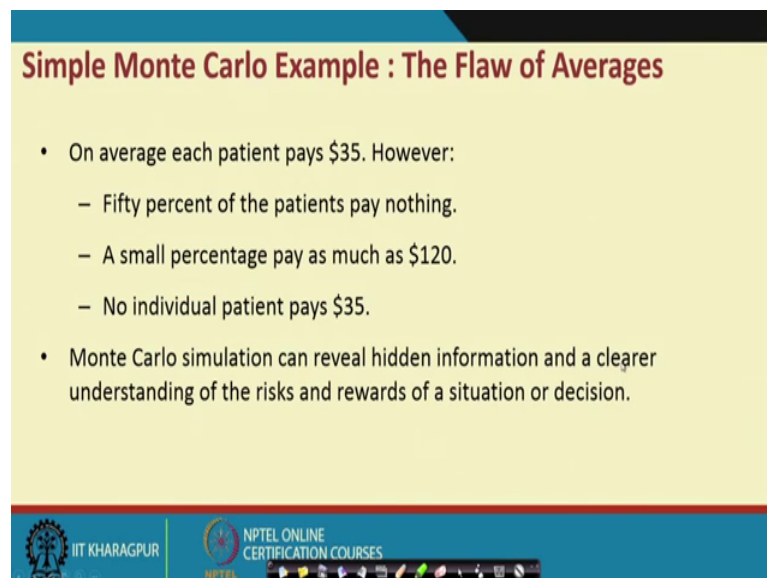
So, I will take you to a kind of an excel sheet and show you how once you get a little bit insights and little bit you know structure how simulation can help you to generate you

know kind of you know future requirement or the kind of enough problem requirement or the kind of you know management requirement.

So, here it is another kind of an example similar kind of an example with you know more features and more you know variables and here the payment structures then you know this is with respect to payment then you know die total charges you know how much payment with a different kind of you know clustering and then against again what you know I have suggested earlier that you know with a particular you know informations with a particular in item. So, you can first you know check the kind of know be you know means visualize first check the particular you know distribution then again use this particular distribution for the simulation it can be in a kind of in a bivariate setup and it can be with you know multivariate setup.

Since we are now connecting with you know payment with the number of trials so; that means, it is a 2 variable case we can plot and then you can check, but when we have a multiple kind of you know structure then we can mathematically check and connect their models and then finally, can you know predict the kind of you know requirement right. So, this is how the kind of you know simulation structure and likewise you know.

(Refer Slide Time: 17:00)



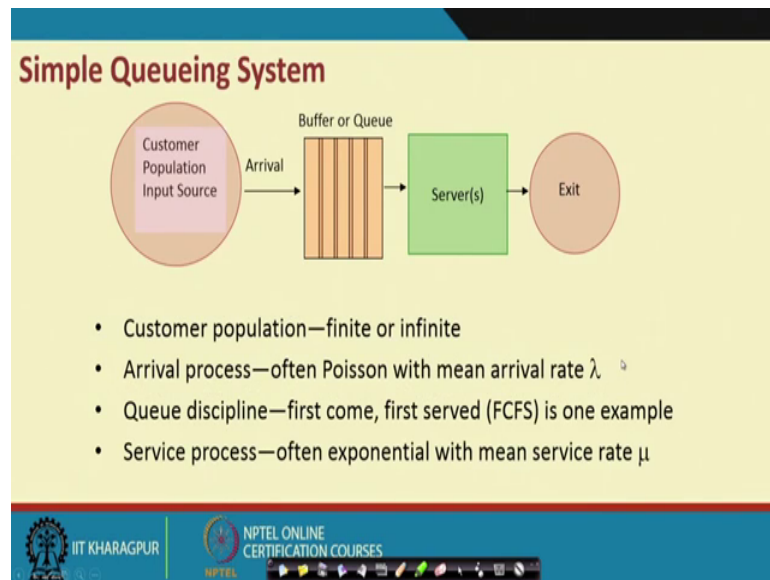
**Simple Monte Carlo Example : The Flaw of Averages**

- On average each patient pays \$35. However:
  - Fifty percent of the patients pay nothing.
  - A small percentage pay as much as \$120.
  - No individual patient pays \$35.
- Monte Carlo simulation can reveal hidden information and a clearer understanding of the risks and rewards of a situation or decision.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

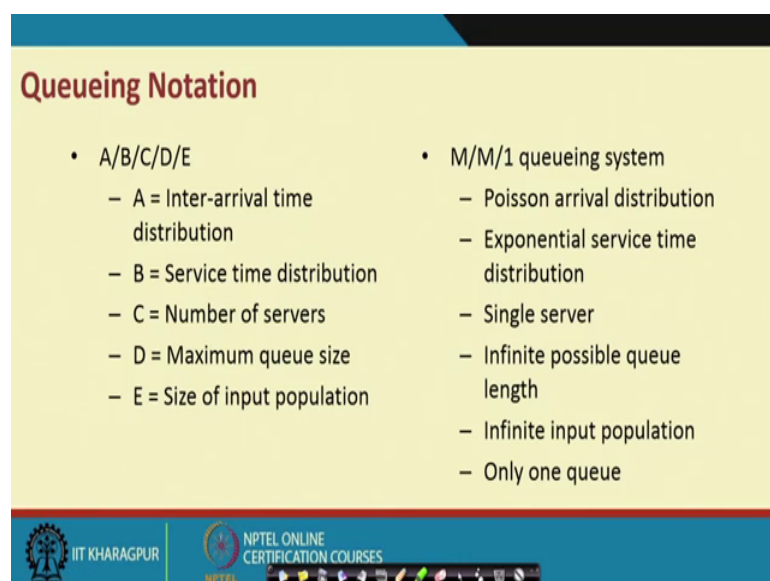
So, this is how the analysis. So, once you get some kind of an insights then; obviously, you can actually develop a structure through which you can do the kind of you know predictions and do the kind of you know simulation.

(Refer Slide Time: 17:06)



So, simple there is a Quinn's theory through which actually you can also do the similar kind of you know prediction and similar kind of you know kind of you know simulation. So, here it is a question of you knows or means to understand this particular you know technique typically. So, we need actually understand the particular you know system like customer population then it is it may be finite and infinite then arrival process so; that means, we need actually the kind of you know requirement. So, once you know the requirement then we can follow the kind of particular follow the particular you know structure through which you can do the simulation.

(Refer Slide Time: 17:55)



So, in this case the best you know the particular you know system is followed by a Poisson distributions and as through Quinn's theory you first you know identify the kind of requirement then we the particular theory has a different kind of you know models.

(Refer Slide Time: 18:09)

**Queueing Solutions (M/M/1,  $\lambda < \mu$ )**

Capacity utilization = Percentage of time the server is busy

$$= \rho = \frac{\lambda}{\mu} = \frac{\text{mean arrival rate}}{\text{mean service rate}}$$

$$= \frac{1/\text{mean time between arrivals}}{1/\text{mean service time}} = \frac{\text{mean service time}}{\text{mean time between arrivals}}$$

Average total number of customers in the system

= Arrival rate  $\times$  time in the system

$$L_s = \frac{\lambda}{\mu - \lambda} = \lambda W_s$$

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, this is what the usual model. So, the simple actually to find out the capacity utilization which is actually percentage of time the server the busy with respect to that particular you know problem. So, then we like to find out the capacity utilization which actually the ratio between mean arrival rate mean service rate and again. So, with respect to some you know mathematics we can calculate the particular you know system of course, actually when will you do for you know more kind of you know future requirement and the kind of you know simulated values.

(Refer Slide Time: 18:37)

**Queueing Solutions (M/M/1,  $\lambda < \mu$ )**

Average waiting time in the queue

$$= W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Average time in the system

= Average waiting time in the queue + Average service time

$$= W_s = W_q + \frac{1}{\mu} = \frac{1}{\mu - \lambda}$$

Average length of the queue (or average number in the queue)

$$= L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \left(\frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu - \lambda}\right)$$

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, manual it is not possible, but just we have to follow the kind of you know distribution and the kind of you know structure through which actually we just get connect with you know inputs and then the kind of you know distributions and after that it will help you to generate these simulated results as per the particular you know problem requirement and the kind of you know business requirement.

(Refer Slide Time: 19:07)

**VVH M/M/1 Queue Example**

- **Goal:** Only one patient waiting in line for the MRI
- **Data:**
  - Mean service rate ( $\mu$ ) is four patients/hour and is exponentially distributed
  - Arrivals follow a Poisson distribution and the mean arrival rate is three patients/hour ( $\lambda$ )

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, in this case you know the means; obviously, what I mentioned you know the objective must be very clear actually. So, if you do not know the kind of you know

response variables or the kind of you know objective requirement and the kind of you know attributes through which you can predict the response variables or simulate the response or the kind of you know objective.

So, then you are not in a position to do the kind of you know analysis so; that means, we technically get to know the insights first you understand the insight the problem structure the kind of you know objectives then with the help of you know availability and the kind of you know understanding you can develop a system and then develop a kind of you know simulation through which actually you generalize the problem as per the particular you know requirement and the kind of you know management requirement.

(Refer Slide Time: 19:56)

**VVH M/M/1 Queue Example**

If one customer arrives every 20 minutes and it takes 15 minutes to perform MRI, the MRI will be busy 75% of the time.

Capacity utilization of MRI = %age of time MRI is busy

$$= \rho = \frac{1/\mu}{1/\lambda} = \frac{15 \text{ minutes}}{20 \text{ minutes}} = 75\%$$
$$= \rho = \frac{\lambda}{\mu} = \frac{3}{4} = 75\%$$

The slide includes logos for IIT KHARAGPUR and NPTEL ONLINE CERTIFICATION COURSES at the bottom.

So, this is similar kind of you know problems with the different understanding and different kind of you know structuring.




(Refer Slide Time: 19:58)

### VVH M/M/1 Queue Example

Average time waiting in line  
$$= W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{3}{4(4 - 3)} = \frac{3}{4} = 0.75 \text{ hours}$$

Average time in the system  
$$= W_s = \frac{1}{\mu - \lambda} = \frac{1}{4 - 3} = 1 \text{ hour}$$

Average total number of patients in the system  
$$= L_s = \frac{\lambda}{\mu - \lambda} = \frac{3}{4 - 3} = 3 \text{ patients}$$
  
= Arrival rate  $\times$  Time in the system  
= 3 patients/hour  $\times$  1 hour = 3 patients




So that means

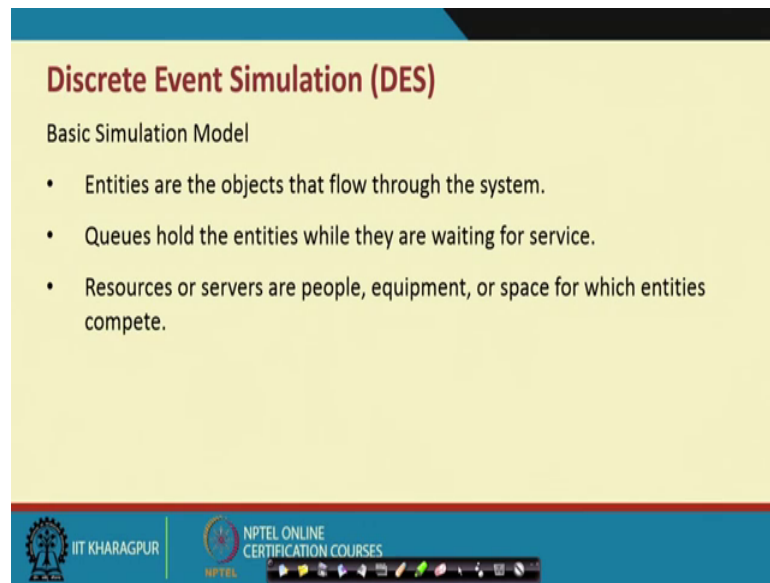
(Refer Slide Time: 20:01)

### VVH M/M/1 Queue Example

- Average number of patients waiting in line =  
$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \left(\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu - \lambda}\right) = \left(\frac{3}{4}\right)\left(\frac{3}{4 - 3}\right) = \frac{3^2}{4(4 - 3)} = \frac{9}{4} = 2.25 \text{ patients}$$
- VVH needs to decrease the utilization,  $\rho = \lambda/\mu$ , of the MRI process
- VVH can
  - Increase the service rate ( $\mu$ )
  - Decrease the arrival rate ( $\lambda$ )
  - Do a combination of both



(Refer Slide Time: 20:04)



**Discrete Event Simulation (DES)**

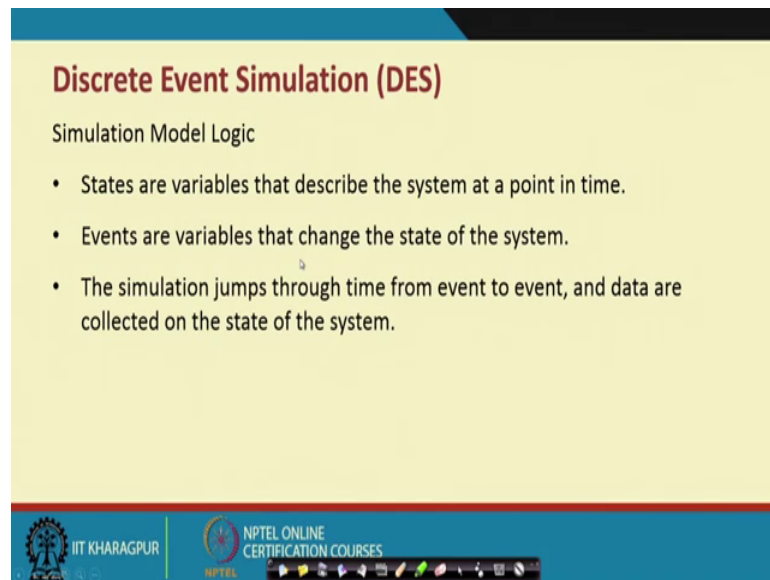
Basic Simulation Model

- Entities are the objects that flow through the system.
- Queues hold the entities while they are waiting for service.
- Resources or servers are people, equipment, or space for which entities compete.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

Technically we have already discussed the Monte Carlo simulations and the Quinn's theory through which you can do this summation; that means, these are the different process through which you can understand the problem then develop a structure and then go for the kind of you know simulation the kind of you know future forecasted requirement as per the management you need and the third kind of you know structure in the simulation is called as a discrete event simulation and this is the basic simulation models which and you know there is a kind of you know objects that flow through the system and it is again through Quinn's theory like you know entities where they are waiting for a kind of you know for a service and; obviously, there are certain attributes through which actually we like to actually understand the problem and then and get the insights for the kind of you know simulation requirement and so.

(Refer Slide Time: 20:59)



**Discrete Event Simulation (DES)**

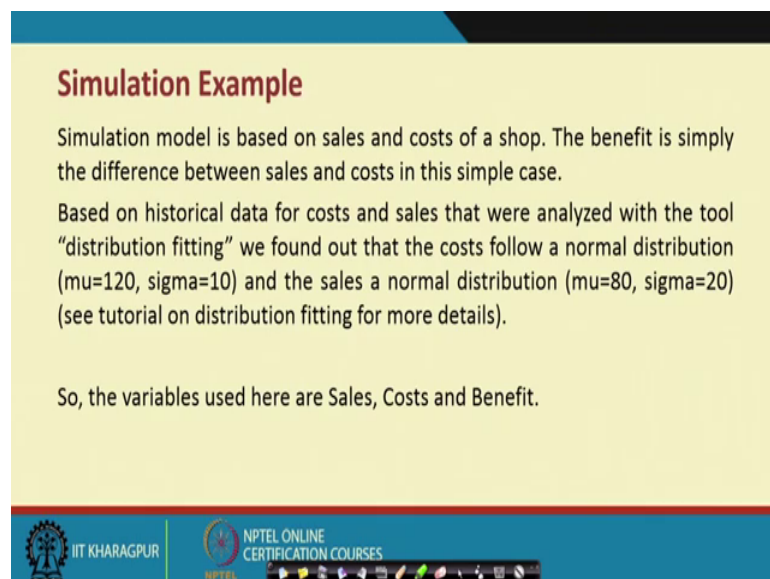
Simulation Model Logic

- States are variables that describe the system at a point in time.
- Events are variables that change the state of the system.
- The simulation jumps through time from event to event, and data are collected on the state of the system.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

These are all you know basic understanding about this discrete event simulation and what I will do actually only I mean say again so, in order to understand the particular you know structure.

(Refer Slide Time: 21:08)



**Simulation Example**

Simulation model is based on sales and costs of a shop. The benefit is simply the difference between sales and costs in this simple case.

Based on historical data for costs and sales that were analyzed with the tool “distribution fitting” we found out that the costs follow a normal distribution ( $\mu=120$ ,  $\sigma=10$ ) and the sales a normal distribution ( $\mu=80$ ,  $\sigma=20$ ) (see tutorial on distribution fitting for more details).

So, the variables used here are Sales, Costs and Benefit.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

So, I will take you know 3 variables and here the 3 variables are you know sales, cost and benefit and with a particular you know understanding and you know particular structure and we like to you know check how simulation and you know can generate certain kind of you know future requirement as per the problem need and you know



once you know acquainted with the particular you know structure then you connect the actual insights and then get the you know inbuilt insights then you know finally, go for the kind of you know prediction and the requirements. So, now, in this case we like to you know go for you know simulation that too we have a lots of you know simulation techniques.

So, we like to it because one of the most important technique is the Monte Carlo simulations and we like to use that particular you know technique through which you generate the simulated figure and you see basic requirement is here. So, there are sales cost and benefit and some data which we have highlighted here and then again you can actually develop the simulation of a particular problem with you know any software's which like you know which you have already highlighted.

You can go through you know means you can solve the problem by math lab you can solve the problem through R software's and you can go through kind of you know other software like you know (Refer Time: 24:16) data and against in this case actually we are using here excel start which is a similar kind of you know package whatever we have already highlighted and same way we will also highlight this problem and get to know how simulation can help the particular you know business process to get some kind of you know effective you know kind of you know scenario through which you can do the management decision and in this case going through this data for the particular you know simulation packets.

So, it will be appearing in the advanced futures and here we have a plenty of you know tools like you know conjoint analysis decision, time series analysis Monte Carlo. So, likewise we have a couple of you know advanced you know analytics tools through which you can you know analyse the problem get insights and then think about the prediction about the future. So, since we are you know our focus is on you know simulation that too in Monte Carlo simulation. So, you just click here and then since there are 3 variables. So, one of the first hand requirement of this particular you know structure is to define a particular distribution that what I have already mentioned that you know understand the problem then connect that problem with a particular distribution.

So, connect with a particular distribution means it is not arbitrary choice you must have a kind of you know insights to check that you know whether this particular problem is

fitted with you know binomial distribution, Poisson distribution or normal distribution and let us assume that this is connected with a normal distribution. So, first we check the, you know kind of you know structure. So, let us start with the particular sales value first and then go to the advanced feature then go to the defined distribution. So, which I have defined here actually normal distributions. So, put then again go to the cost side and then again go to the excel start advanced features choose Monte Carlo simulations again define the discussion for the cost and again we allow for the normal distributions and then finally, you go to the benefit which is actually our requirement then again go to the Monte Carlo simulations.

So, now here we would like to actually go we like to have define the kind of in a result variables because we like to actually here simulate the benefit structure with respect to sales and cost and again. So, we define structure you have to pick up a particular you know set up and finally, you go to the excel set excel start and then advanced futures again pick up the Monte Carlo simulations and then run the model so; that means, what we have done we you know we have actually system here benefits with respect to sales and cost and we like to generate the you know simulation for the benefit structure and that too with respect to sales and cost and for that we have actually used the normal distribution structure and as per the normal distribution requirement, we need actually population parameters that is the mean and standard deviation.

So, giving a mean and standard deviations as a kind of you know proxy which you can you know assume or we can actually get through a kind of you know market informations and the kind of you know expert structure and then you know expert structure through it you know we can actually design and then go for the kind of you know simulation.

So, now you just put and then it will be run the models and finally you to get the kind of you know simulated value and this is what the simulated value sees here.

(Refer Slide Time: 27:51)

The screenshot shows the XLSTAT interface with simulation results. The 'Simulation results (Distributions)' section is expanded to show 'Descriptive statistics' for 'Costs' and 'Sales'.

Statistic	Costs	Sales
Nbr. of obs	1000	1000
Nbr. of mis	0	0
Sum of wgt	1000	1000
Minimum	14 961	79 934
Maximum	150 561	151 339
Freq. of ms	1	1
Freq. of ms	1	1

This is how statistic and cost and sales and these are all again descriptive statistics starting with you know number of observations and minimum and maximum.

(Refer Slide Time: 27:58)

The screenshot shows the XLSTAT interface with simulation results. The 'Simulation results (Distributions)' section is expanded to show 'Descriptive statistics' for 'Costs' and 'Sales'.

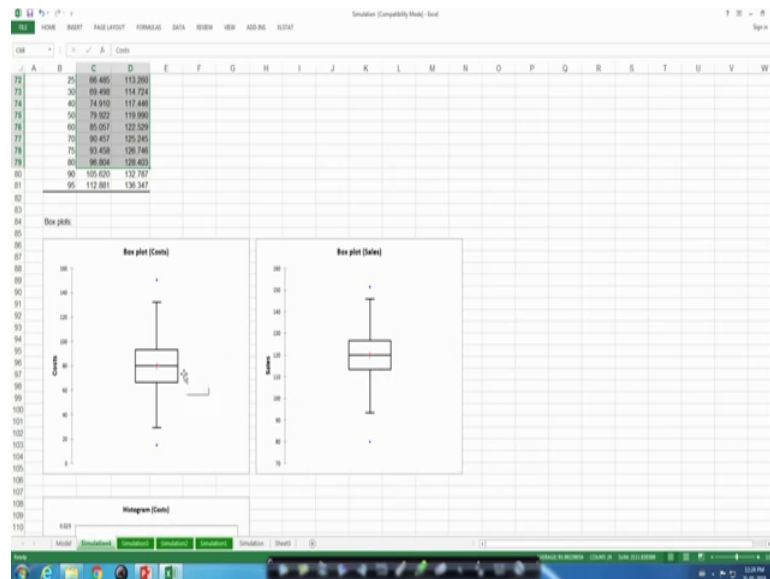
Statistic	Costs	Sales
Nbr. of obs	1000	1000
Nbr. of mis	0	0
Sum of wgt	1000	1000
Minimum	14 961	79 934
Maximum	150 561	151 339
Freq. of ms	1	1
Freq. of ms	1	1
Range	135 600	71 405
1st Quartile	66 508	113 279
Median	79 960	119 969
3rd Quartile	93 458	136 748
Sum	80007 865	119968 405
Mean	80 008	119 969
Variance (d)	400 111	109 729
Variance (s)	400 511	100 830
Standard d	20 003	10 476
Standard s	20 013	10 241
Variation (s)	0 250	0 084
Skewness	0 010	-0 041
Skewness (s)	0 010	-0 041
Kurtosis (d)	0 000	0 000
Kurtosis (s)	0 006	0 131
Kurtosis (f)	0 012	0 138
Standard e	0 633	0 318

In fact, we in the simulation we put you know 1,000 structure that can be actually changed you put in a 5,000 to you know 100,000. So, automatically the simulated figures will be coming. So, these are all various you know descriptive statistics. So, this is again like you know previous discussion about the random for a support vector machines and the data mining. So, descriptive statistic will be the kind of you know mandatory

requirement to report to know understand the actual insight and the actual kind of you know scenario then the particular technique can be actually developed and then compared with the actual scenario with the predicted scenario.

So, descriptive statistic will help you to understand the actual you know effect and connect with a particular you know distribution particularly for the simulation then it will help you lot to go for the means choosing a best distributions then for you know simulation if you are the choice of particular distribution is wrong then the simulated figure will be also wrong so, as a result the management decision cannot be effective.

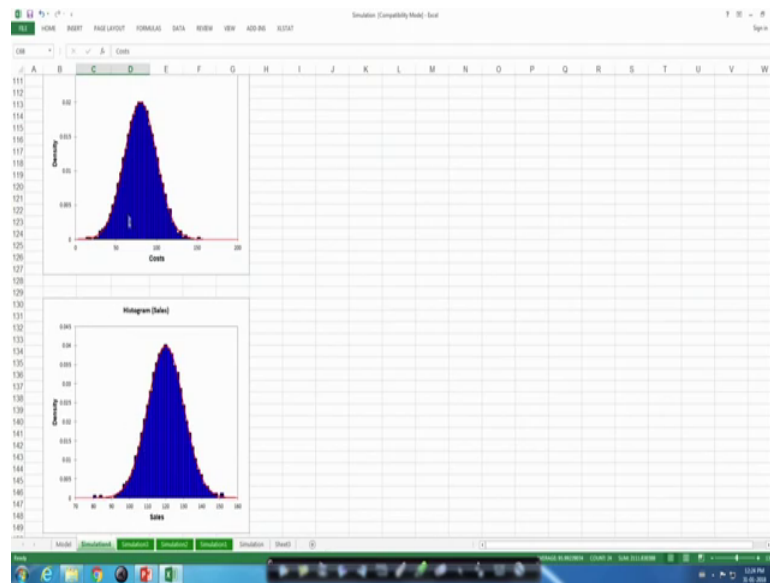
(Refer Slide Time: 29:19)



That is why this should be actually you know first connected and understand and this is how the you know kind of you know cost and the sales structure and this is how the box diagram will give you some kind you know indication about the model validations.

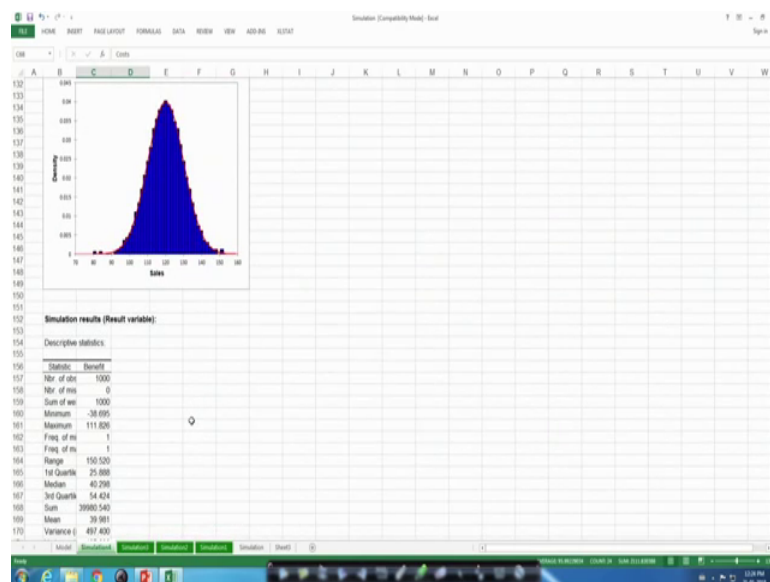


(Refer Slide Time: 29:29)



And this is how since we choose actually normal distribution and this is how the cost structure and this is how the sales structure and it is normally distributed and as a result the most important thing is actually the simulated results.

(Refer Slide Time: 29:40)



This is what the simulated results and we simulated for you know benefits that too with respect to sales and cost and the simulated figures are available here. So, these are all actually simulated figures.

(Refer Slide Time: 30:00)

The screenshot shows an Excel spreadsheet with the following data:

**Simulation results (Correlation matrix):**

Correlation matrix (Spearman)

Variables	Credits	Sales	Benefit
Credits	1	0.027	-0.884
Sales	0.027	1	0.413
Benefit	-0.884	0.413	1

Values in bold are different from 0 with a significance level alpha=0.05

**Simulation details:**

Simulation	Distributions			Result variable
	Credits	Sales	Benefit	
1	113.781	104.519	-9.261	
2	85.273	99.888	14.615	
3	81.802	120.988	38.179	
4	116.343	128.567	12.223	
5	88.174	130.432	42.243	
6	95.074	129.984	34.919	
7	77.984	114.847	36.963	
8	76.642	136.839	60.187	
9	90.913	113.288	22.375	
10	85.723	123.151	37.432	
11	110.987	114.152	4.938	
12	85.018	101.880	18.442	
13	87.580	113.188	45.625	
14	108.674	138.743	30.076	
15	94.802	123.724	28.622	
16	109.186	117.942	8.777	
17	80.254	122.830	62.576	
18	77.944	120.689	42.755	
19	69.721	120.112	50.391	
20	75.504	100.856	25.351	
21	62.331	131.962	69.631	
22	77.586	127.657	51.148	

So, let us see here. So, this is actually particular you know distribution which is followed by normal distribution and this is what the result variable so; that means, technically. So, the simulation generate now it is a these are all you know future requirement and this will be available for actually 1000 because we ask the software to run the model for you know 1000. So, if you change that particular you know structure then if you choose you know 500 then it will give you 500 simulated figure, if you choose you 1000 it will give you 1000 you know simulated figure again you see here.

So, this typical understanding here is that you know if you go to the first end data. So, it is with respect to cost and you know sales and that too the kind of, you know benefits CS. So, technically it is the benefit which depends upon you know sales and cost and what we have done actually we understand fast the cost structure connect with you a particular distribution and sales structure which we can understand and connect with a particular distribution.

Then now finally, benefit is a predicted and you know simulated with the help of you know cost and sales where we are assuming a normal distribution structure and then one of the most important thing here requirement is here you know fix the or assume a particular you know a structure that is the requirement of you know parameters value that is a kind of you know for the normal distribution we need actually mean and standard deviations. So, once you get mean and standard deviations and then the kind of

you know variables you know insights and the kind of you know variable understanding then you can actually simulate for the you know requirement problem requirement.

So; that means, in this case here the kind of you know benefit is actually predicted with the help of you know cost and sales and since we have fixed the distribution effects to the particular data structure. So, the software will help you to give you the forecasted figure of you know benefits that too with respect to cost and sales that too for you know future one thousand units altogether. So, now, if we change the particular you know simulated you know requirement then; obviously, this value can be actually also generated more.

So, this is actually big kind of you know structure I am just highlighting so, that you can see more effectively. So, these are all actually simulated figures. So, this is starting with the first unit. So, this is one this is. So, this is how the first units this is what the simulation and in the simulation this is the main results through which we are actually looking as per the particular you know problem requirement and management requirement and. So, it is for 1000 unit because we have fixed the particular you know software or allow the softwares to generate simulation figure for 1000 unit if you change then you know it will give you means if you change to you know or 2000 or 3000 then again you will get you know simulated figure of you know 3000 you know items.

So, as a result you know what we actually you know try to address here that you know simulation is a kind of in a effective tool through which you can get some kind of you know structure understand the insight and you know define a kind of you know structure or develop a kind of you know structure for the future requirement and; obviously, over the time you can actually compare which the actual structure or in actual fact, but you know means it is a part of the game and simulated figure is always for you know future kind of you know requirement; that means, in this contest. So, we have actually predicted for you know 1000 units you know forecasted you know structure, but how is the actual structure happening.

So, that needs to be a compared and then finally, you can you know check the validations; that means, if your problem understanding is very clear and the choice of the distribution is very clear then the difference between the actual happenings and the kind of you know predicted happenings may not be actually highly distance. So, the effective

prediction and effective simulation means we can call that you know effective prediction and effective simulation for the particular you know business requirement. If the difference between the actual and predicted will be a you know minimum as much as possible actually that is how we you know most of the techniques we use actually the kind of you know prediction indicators various prediction indicators just to compare the actual scenario and the predicted scenario and then check what particular you know methodology or you know framework gives you know highly minimum kind of you know bias. So, that you know we can declare that you know the model is very effective for the particular you know problem and the kind of you know prediction and the kind of you know management requirement.

(Refer Slide Time: 35:09)

### Simulation Example: Result

Model constructs identified by XLSTAT:			
Name	Cell	Default cell value	Data type Distribution Parameters
Costs	[Simulation 80.000]		Continuous Normal $\mu = 80$ , $\sigma = 20$
Sales	[Simulation 120.000]		Continuous Normal $\mu = 120$ , $\sigma = 10$

Result variable:		
Name	Cell	Formula
Benefit	[Simulation B2-B3]	

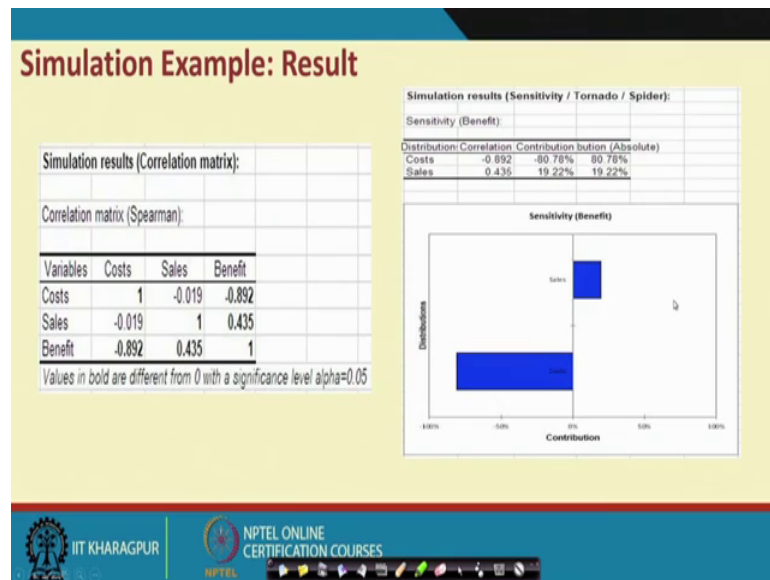
Quantiles:	
%	Benefit
5	2.435
10	10.752
20	21.157
25	24.505
30	27.674
40	34.416
50	40.616
60	46.250
70	52.655
75	55.971
80	59.535
90	68.499
95	76.720

Statistic	Costs	Sales
Nbr. of obs	1000	1000
Nbr. of mis	0	0
Sum of weights	1000	1000
Minimum	11.798	88.464
Maximum	140.724	149.763
Freq. of min	1	1
Freq. of max	1	1
Range	128.927	61.299
1st Quartile	66.533	113.257
Median	79.999	120.001
3rd Quartile	93.468	126.745
Sum	79982.874	119998.911
Mean	79.983	119.999
Variance (s)	400.375	99.681
Variance (p)	400.776	99.680
Standard error	20.009	9.979
Standard error	20.019	9.984
Variation coefficient	0.250	0.083
Skewness	-0.023	-0.005
Skewness	-0.023	-0.005
Skewness	0.000	0.000
Kurtosis (F)	-0.001	-0.064
Kurtosis (P)	0.005	-0.058
Standard error	0.633	0.316
Lower bound	78.741	119.379
Upper bound	81.225	120.618
Standard error	0.077	0.077
Standard error	0.155	0.155
Mean absolute deviation	15.962	7.972
Median absolute deviation	13.498	6.746
Geometric mean	77.191	119.580

So, these are the kind of innocent things which you have already highlighted and.

(Refer Slide Time: 35:12)



This is how the cost and sales benefits and that is the correlation matrix so. In fact, you know the correlation says that you know cost and benefits are you know highly correlated, but sales and cost are you know low correlated and again benefit in sales is having high correlated anyway. So, our job is here easy to predict the kind of you know benefit structure with respect to sales and cost of course, when we analyse this kind of you know problems in between there may be certain issues you know with respect to relationship between independent variables you know the kind of you know multi coordinate issue the kind of you know data issue the kind of you know model issue. So, what is actually the requirement that you know you understand the data structure then understand the problem theoretically logically convinced as per the particular you know requirement.

Then you connect a particular you know models or choose a particular model as per the particular you know problem structures with respect to variables availability and the data availability. Then come with you a kind of you know a model structure and we can validate after that and finally, fix the kind of you know model which can give you effective prediction and you know effective kind of you know requirement and which is actually the kind of you know management requirement.

(Refer Slide Time: 36:35)

### Simulation Example: Result

Simulation details:

Simulation	Distributions		Result variable
	Costs	Sales	Benefit
1	87 847	134 575	46 729
2	52 520	128 063	75 542
3	64 282	116 051	51 769
4	70 665	139 360	68 694
5	63 135	116 703	53 568
6	101 964	121 555	19 621
7	60 998	125 990	64 992
8	108 731	135 253	26 522
9	44 416	104 862	60 446
10	100 687	102 868	2 181
11	64 329	112 888	48 559
12	49 373	119 627	70 253
13	84 109	98 187	14 078
14	51 833	121 692	69 858
15	78 294	123 245	44 951
16	93 845	120 477	26 633
17	92 217	127 923	35 706
18	88 621	137 631	48 910
19	70 159	148 237	78 068
20	72 166	113 629	41 463
21	80 986	114 506	33 520
22	79 913	116 655	35 752
23	92 336	110 671	18 234
24	88 292	109 084	20 792
25	66 935	128 224	61 289
26	78 631	123 416	44 785
27	104 762	113 267	8 504
28	98 625	124 327	25 802

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

(Refer Slide Time: 36:37)

### Summary

- Simulation is a powerful tool for modeling processes and systems to evaluate choices and opportunities.
- Simulation can be used in conjunction with other initiatives such as Lean and Six Sigma to enable continuous improvement of systems and processes.

IIT KHARAGPUR | NPTEL ONLINE CERTIFICATION COURSES

And then accordingly we can go for some kind of, you know effective decisions. So; that means, there are a number of you know areas where simulation can be applied and that is actually you know that is how it is very useful for the kind of you know problem with a requirement and the kind of you know management requirement and all together in the predictive analytics we have discussed you know lots of techniques. In fact, particular technique requirement is different and I mean say sometimes we can choose a particular technique as per the problem requirement. So, altogether we get to know various tools predictive analytics tools through which some management problem can be analysed can

be predicted and accordingly management decision can be taken into consideration with this we will stop here.

Thank you very much, have a nice day.