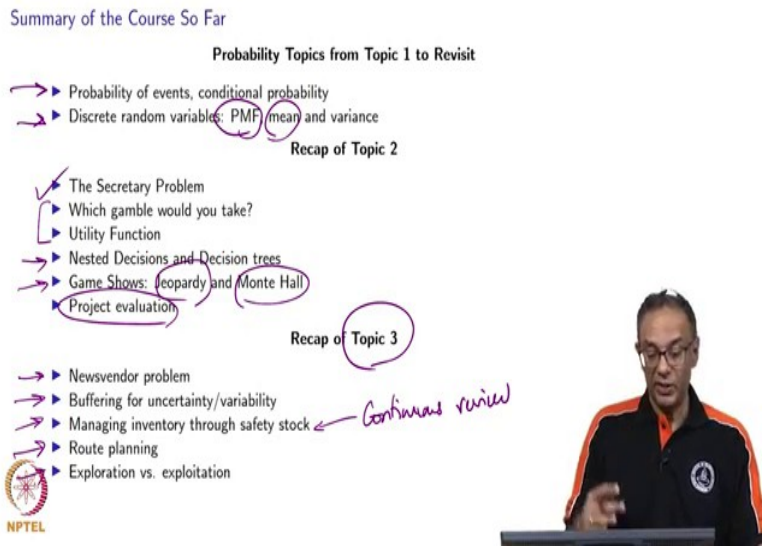
Decision Making Under Uncertainty
Prof. Natarajan Gautam
Department of Industrial and Systems Engineering
Texas A & M University, USA

Lecture - 26
Introduction to Sequential Decision Making

The next topic is Sequential Adaptive Decision Making. So, in this course on decision making under uncertainty, we have so far seen situations where we did not adapt a whole lot. And in this case we are going to be talking about, where decisions are made one after the other sequentially, then you kind of adapt depending on how they randomness gets involved.

(Refer Slide Time: 00:41)



Before we do that, I do want to summarize the course that we have seen so far. We started by talking about various probability topics and for this topic here the following are the items that are going to be important. We have to again go back to the probability of events, we will heavily emphasize conditional probability especially towards the second half. We will also talk a little bit about discrete random variable, we will not do any continuous random variables in this.

So, once again brushing up on the probability mass function and computation of mean are very important. I do not think I do any variance either. So, basically the probability mass function and mean are two topics that you should brush up. So, I would recommend brushing up them if this is a little bit unclear for you. Now, quick recap of what we did from the

previous topic: From topic 2, we looked at making one time decision such as in the secretary problem, where we figured out what is the best way to select a secretary. Then, we talked about the issue of utility function, where we were looking at three options to gamble: a, b and c with various probability and even various payoffs and figure out what we should do, but again remember that both the secretary problem and the utility function are one time decisions.

And then the next one was the nested decision, where we use the decision tree for that- Greene Cat and in that problem we had a little bit of adaptive. Because what happens is one of the strategies when they were adopted, we wanted to see if that was effective or not. If it was not effective, do something. So, there was a little bit of adaptive there and then we could take that up a notch in today's topic- topic 4.

We also looked at some other single decisions like in the game shows of jeopardy as well as in Monte Hall both those cases were one time decision about in Jeopardy how much to wager in the final Jeopardy and the Monte Hall problem whether or not you should switch in terms of which door you selected out of the three doors.

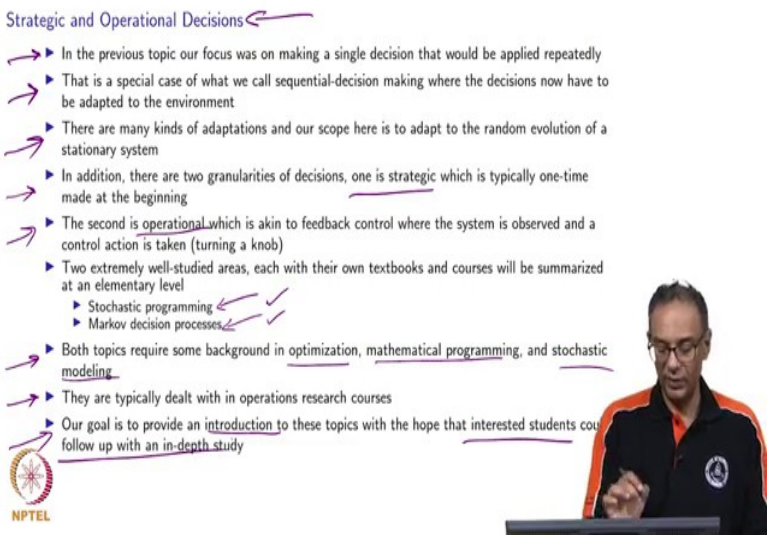
We went into PERT a little bit project evaluation again although that was a set of activities that were done sequentially. You still have one decision to make, which is how long it is going to take for you to complete the project so that you could tell your customer, when they could come and pick up what they were looking for.

Then, we moved on to topic 3 where we made a repeated decisions. We went on doing the same thing over and over again. We looked at that in the newsvendor problem. I do want to say one thing here before moving ahead. Although, we said that the demand was IID in the newsvendor problem that is not really necessary. Every day if the demand changes and as long as you know how to characterize it, you could just use that directly and then say: ok, your period is one, because you have to make sure that all your demand is satisfied for that. And then the next day you get a fresh item either a new set of flowers or in a different set of newspapers. So, so that can be used even if your demand actually changes with time.

We talked a little bit very qualitative about how do you buffer for uncertainty and variability. Then we went into some mathematical concepts for how to manage inventory using safety stock, we will again look at inventory today, but that one we were essentially looking at what is called continuous review. So, you were looking at your inventory in a continuous fashion

throughout and then you were making decisions. Today, we will be looking at inventory however periodically; we will see that today. Then we looked at after the inventory part, we looked at route planning or route planning and finally we looked at this idea of exploration versus exploitation which is a fairly popular topic today.

(Refer Slide Time: 04:45)



So, now let us move on to topic 4, and topic 4 essentially has two flavors: the first part we will focus a little bit on strategic decisions. And then with a little bit of operational type of decision. We were earlier only looking at a single decision, where you would keep repeating that decision over time. The topic 3 is also a special case of what we call a sequential decision.

So, in the newsvendor problem for example, every morning look exactly the same. Imagine in the inventory case, where next morning could be different from the previous morning. So, you were doing something a little bit different, so we did do a little bit of sequential decision making, where we adapted to be environment. And the environment itself could evolve randomly and also our decisions could result in some uncertainties, which is going to be the focus of today's topic.

Now, I do not want to say that there are two granularities of decisions. There are still one time decisions here and this is what we call strategic. So, some times this is from topic 2- the one time decision part, there is a notion of coming up with a strategic decision and then there

is something called an operational decision that is the second type of decision which is like what we call as feedback control.

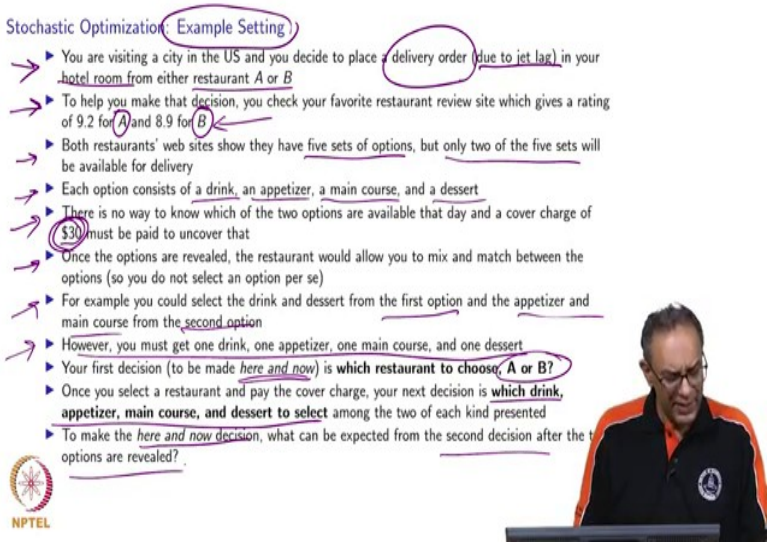
We look at something, we get feedback and then we turn them out. So, that is kind of two stage strategic decision. Now these two fall under the: “not really the strategic and operational”, but these problems fall under a very large set of books and courses available in the literature and one area is called stochastic programming.

So, the first three or four lectures that we will see today will be on the topic of stochastic programming, then the last field will be on Markov decision processes. So, these are two topics that were pretty much independently evolved over time, but there is very little crossing between the two areas. However, they address somewhat similar problems; we will we will talk about those. The stochastic programming has this notion of strategic decisions and then follow the operational, in the Markov decision process will tend to be a lot more operational.

So, that is kind of how it is. However, both topics in some sense are one, which require tremendous amount of tools in optimization, in mathematical programming, and also in stochastic model. We are not assuming that level by teaching the course. So, therefore, we will keep the topic of stochastic programming and Markov decision processes at a very introductory fundamental level. I do want to clarify that, I do not expect you to become experts on this topic (Refer Time: 07:53), but you will get a flavor of what is going on.

These topics are typically dealt with in upper level operations research courses. Our goal is to provide some type of an introduction to these topics with the hope. One of the topic I will squeeze in how simple dialogue, but I just want to give you an introductory lecture. However, if you are interested I would recommend that you take up an in depth study.

(Refer Slide Time: 08:23)



I am going to give you an example for stochastic optimization. Now, this is a bizarre example in some sense, but I wanted to create an example of something which you know is complicated enough that you have to think about it carefully and also fix what we are looking at and also is simple enough that you do not have to do very complex things like mathematical programming (Refer Time: 08:48).

Consider the following things, let us say you are visiting a city in the US. So, you are going to see a lot of dollars signs (Refer Time: 08:57) in this example and you are in a hotel room and you have two restaurant options: A and B. You can order food in both places and they will deliver it to you. Well you might ask why do you want to place a delivery order. Well it could be several reasons- may be so cold, you do not want to go out there (Refer Time: 09:19) or more commonly you could have just landed in US, you are guaranteed to have jet lag because of which you just sit in your room and have the food get delivered.

Let us pretend in that sense (Refer Time: 09:31). Now, what you do is you look up your favorite restaurant reviews right. There are many of them or I sometimes use tripadvisor if I go and order something and you go check it out and then they will give you a number.

Now, tripadvisor does not do ratings out of ten, they rate it out of 5. So, this is a special site which cannot be named no it is not like (Refer Time: 09:58) you do not worry about it, it is neither a big source. Turns out that the two restaurants have reviews of 9.2 on their site and

8.9 respectively. So, 9.2 for A and 8.9 for B and you want to first select the restaurant and then place an order.

Now, both restaurant are very similar, this is purely for academic reasons, i.e., I want to give you an example. Now, they have five sets of options, I will again explain this in the in the next slide. But, remember that there are five sets of options, you will see all five options when you go to the website, but when you are ready to make a purchase, you will only see two out of the five set of options. Now, each of the two options or each of the five options as a matter of fact has a drink, appetizer, main course and a dessert. However, you have to commit to the restaurant in order to actually see the two choices.

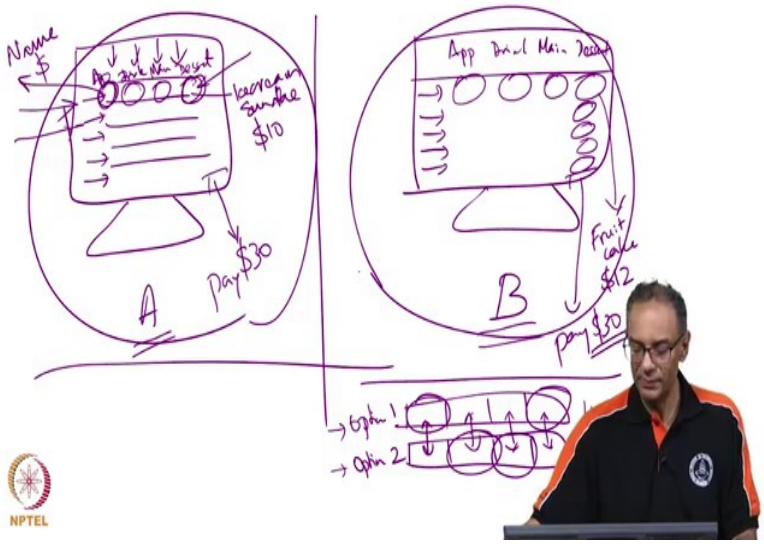
So, what you will see first- you will see all five choices and then you have to commit by paying 30 dollars. And once you pay that, it will tell you which two out of the five choices that are available that day for delivery. Now, once those two choices are revealed, the restaurant will allow you to mix and match; you do not have to select one choice or another choice completely.

You can pick one from one choice and another. For example, you can pick the drink and dessert from the first choice and the appetizer and main course in the second choice. But, you must get one of each, you must get one drink one appetizer, one main course, one dessert. So, your first decision which is to be made, what is called: “here and now”. This is a word that is often which the stochastic optimization, you make “here and now” decision, which is which restaurant you choose.

Now that is important because if you do not decide that, you wait- you are unnecessarily wasting the 30 dollar that you paid for cover charge. So, you decide the restaurant that is looking at your options with a few others. Once you decide which restaurant and you pay the 30 dollar for your choice, then your next decision is what appetizer, what drink, what main course and what dessert to select.

So, there is a two stages with this. Now, first we will talk about the “here and now” decision and also what should we do in the second decision. Once the options are revealed are the two things are we are going to look in. Let me just summarize what we just saw one more time.

(Refer Slide Time: 12:44)



So, essentially what happens is let us say we go to the two websites; I'm drawing computer screen here. I am drawing two screen, but this is just to show you that there will be two restaurants- restaurant A and restaurant B. You will of course, see one at a time. First is the restaurant A, it will list five options: 1, 2, 3, 4 and 5 and there will be an example that will be appetizer, then a drink, then a main course, then a dessert. What you will see here is you will see: option 1, option 2, option 3, option 4 and option 5.

For example, the name of the appetizer will be written and you will also see the price for it. So, what you see is- you will see the name and you will see the dollar amount, the dollar the price. So, you will essentially see 5 times 4, 20 number here, 20 dollar amount and the item as well. I'm not going to the item.

Because it is gets really complex- already 20 things, but think of the following: let us say for example, you could write something like you know like an ice cream sundae. For example, ice cream sundae for dollar 10, so that could be something written in the dessert. Now, that is the first option and then the second and then the third, fourth and fifth.

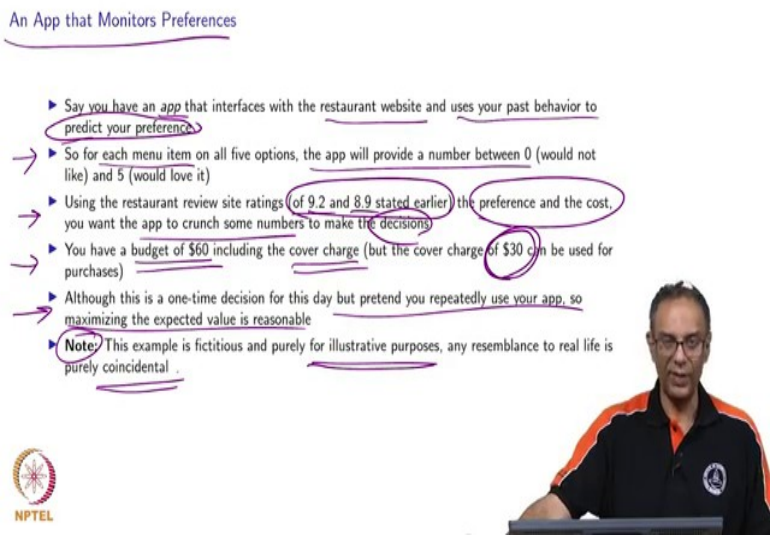
Similarly, for restaurant B as well, you will have an appetizer, a drink, a main course and with dessert and you will have the 5 choices. Now, again it will tell you what is the name of the item and what price it is. So, for example, here each of possible items something like fruit cake for dollar 12 something like that.

So, you have that as a dessert so that is the first option. Now, when there is a second, third, fourth and fifth. Then what happened is here if you decide to go with restaurant A, you pay dollar 30 or here the restaurant B you pay dollar 30. So, you pay the 30 dollar cover cost using the card. Once you do the payment, depend on which restaurant you picked, let us see you picked restaurant B. Once you make the payment, it will go and it will give you only two options: option 1 and option 2, give two options.

Now, what happens is- you can mix and match. You could take the appetizers here, drink from here, the main course from here, and this dessert. You could do that all right. Now, so that is essentially what we can do in here and I have summarized that, it is not like you have to pick one after the other.

But, once you pay the 30 dollars, you will be revealed with this. Now, you have a choice you can pick either of these appetizers, either of these drinks, either of these main courses, either of these desserts. But, your first decision here and now is to figure out whether you pick restaurant A, or you pick restaurant B- that is your first choice.

(Refer Slide Time: 16:11)



Now, once you do that, now this part is something that I am sure will come up very soon in life. There is an app that will monitor your preferences, your tastes, and your likes because you do not want to sit down and do all these calculations. So, you want an app that will go to these websites, get those information and do some numbers on this. So, now this app talks to the restaurant website, uses the past behavior to predict what are your preferences; so that is

what the app does. We are not going to go into algorithm to do the predictions. But, I am just telling you that this app have been looking at what you do in the past.

Some of you this might scare you; I am not saying that it will happen in your future, I am just saying that this is something that you can see it happen. Now, turns out that what this will do is- if you look at all the menu items. So, it will basically we are here, we have not yet paid the 30 dollars, so the app will look at this or this, it will look at both of those and it will make a decision. It will look at all the items that are there, now it is going to provide a rating out of 5: 0, 1, 2, 3, 4 and 5, it knows whenever it sees for example, whenever it sees this ice cream sundae, they look at it and say: ok you really like ice cream sundae, I should give you high rating for that.

Fruit cake, ah! I'm not a big fan, maybe I will give you a low number. So, it knows your preference, now this there is a difference here- the ratings that we saw here are based out of what other users are telling. These ratings are based on your preference. So, what this app needs to do is: use these rating that we saw before as well as your preferences in the past and the app needs to crunch some numbers. So, this will be based out of what we call stochastic programming and tell you: do this. Now, there is one more catch- you have a budget of 60 dollars and that covers a cover charge, what I mean by that is so you can spend up to 60 dollars, but just 30 dollars that you already paid can be used towards that.

So, for example if you are traveling on business and your company says: well we can only reimburse you for 60 dollar. So, it is ok, I am not going to eat for more than 60 dollars that is the situation- that they have a budget 60 dollars and the power charge is also included in that. So, that is what for them like an incentive to pick the restaurant, they do not want you to pick restaurant A or restaurant B and then go in there you find your two choices and: ah! I don't like this, I am going to the other restaurant, and most people will not do that because you have already paid the 30 dollars.

Now, turns out that this although is a onetime decision. However, hint of using this app over and over again, wherever you go, you use this app. So, sometimes you are repeating, so it is good to think of maximizing some type of expected payoff and do that or minimizing expected cost. Now, I do want to say a quick note before I stop to say that this example is purely fictitious; very seriously this is not a joke. I am doing this only for illustrative

purposes. If by any chance there is really such an app or there is really such a restaurant, it is purely coincidental, I have no idea of a real life situation like this.

Thank you.