Decision Making With Spreadsheet
Prof. Ramesh Anbanandam
Department of Management Studies
Indian Institute of Technology-Roorkee

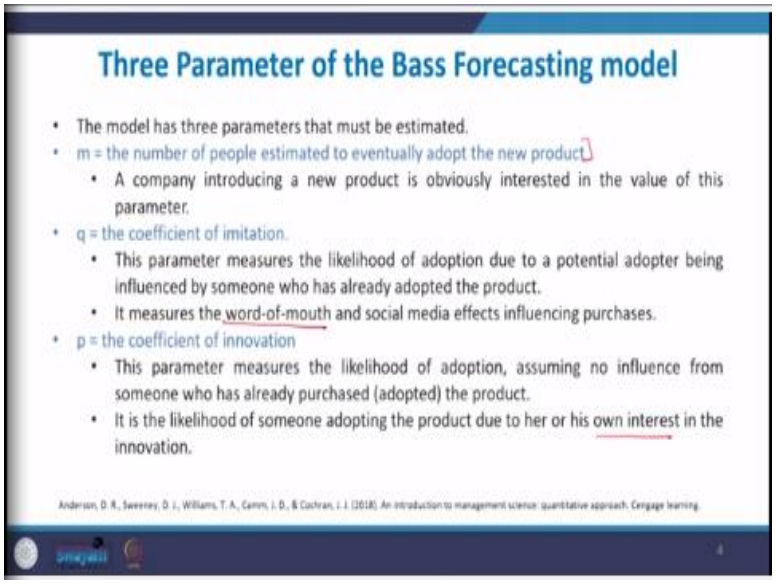
Lecture-27
Non-Linear Optimization Models-IV

Dear students, we are discussing the application of non-linear programming. In the previous class, we discussed various examples. In this class also, I am going to take another example: forecasting a new product, that is an example. So, how do we use non-linear programming to forecast the adoption pattern of a new product?

The agenda for this lecture is forecasting the adoption of a new product. This forecasting technique is called the Bass forecasting model.

I have taken this problem from the book by Anderson et al. Now, we will discuss the forecasting adoption of a new product. So, forecasting new adoptions after a product introduction is an important marketing problem. In this lecture, we can study a forecasting model developed by Frank Bass that has proven to be particularly effective in forecasting the adoption of innovative and new technologies in the marketplace.

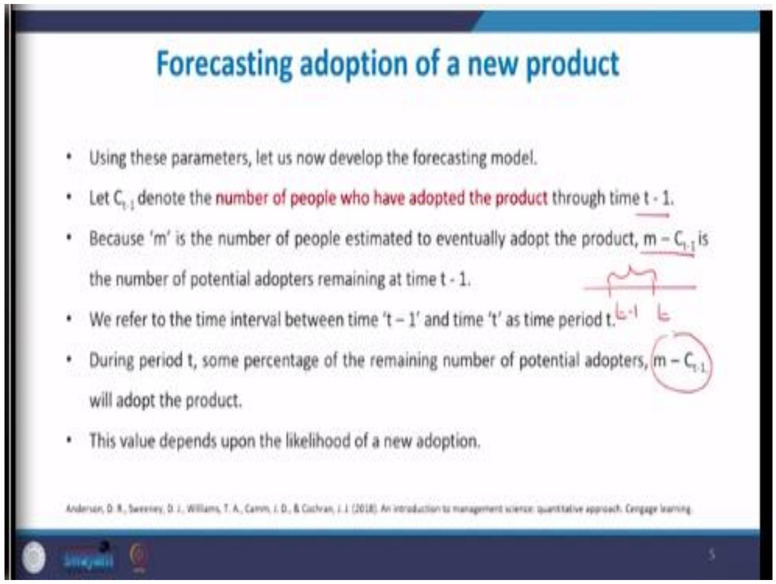
Non-linear programming is used to estimate the parameters of the Bass forecasting model. So, what are we going to do? This professor, Frank Bass, has given a forecasting model that is used for forecasting the adoption of a new product. So, what are we going to do? We are going to estimate the parameters of this Frank forecasting model. So, once we can estimate the parameters, we can estimate the forecasting of a new product that you are planning to forecast.



This model has 3 parameters, so that is 3 parameters of the Bass forecasting model that is what we are going to estimate. Here, at a time of estimating, we are going to use the concept of non-linear programming. So, this model has 3 parameters that must be estimated; what are the 3 parameters? One is m, m is the number of people estimated to eventually adopt to the new product. So, in the end, how many people are going to adopt this new product?

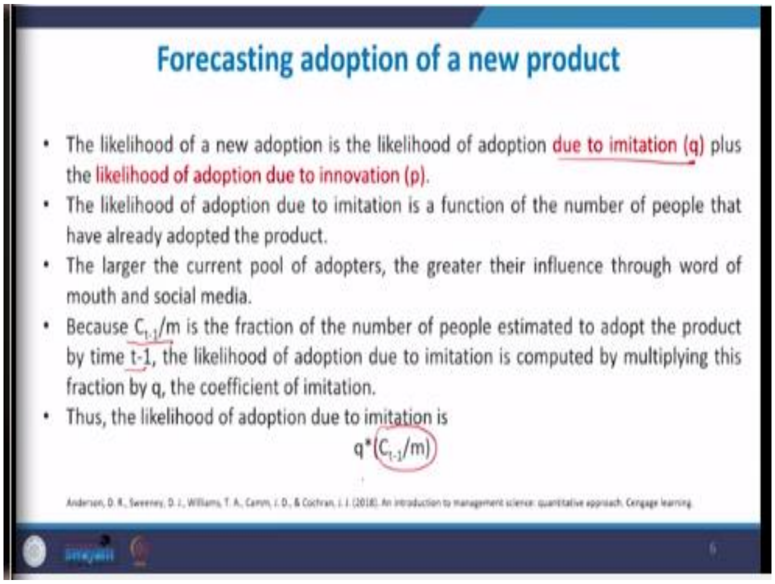
A company introducing a new product is obviously interested in the value of this parameter; they should know how many people are going to adopt this new product. The next parameter q is called the coefficient of imitation. So, this parameter, which is the coefficient of imitation, measures the likelihood of adoption due to potential adopters being influenced by someone who has already adopted the product; it is the coefficient of imitation. So, this parameter says how many people are buying the product who were influenced by others who have already purchased this.

It measures the word of mouth and social media effects influencing purchases. So, I buy the product, I go and write in the review so that other people may read my review, and then they may be interested. They also will buy; that is called coefficient of imitation. Then the third parameter is the coefficient of innovation; this parameter measures the likelihood of adoption, assuming no influence from someone who has already purchased or adopted the product. It is the likelihood of someone adopting the product due to his own interest in innovation. So, these are the three parameters in the Bass forecasting model.



Using these parameters, we can now develop the forecasting model. Let C_{t-1}; here, t represents the current period, and (t – 1) is the previous period. That is C_{t-1} denotes the number of people who have adopted the product through time (t – 1). Because 'm' is the number of people estimated to eventually adopt the product, so m - C_{t-1} is the number of potential adopters remaining at time t - 1. So, m is similar to our population, so C_{t-1} is the number of people who have adopted, so the remaining people are m - C_{t-1}.

We refer to the time interval between time (t – 1) and the time t as a time period, so it is like this. So, this is t, so this is (t – 1), so during this interval, how many people have adopted the product? That is the interval t. During the period t, some percentage of the remaining number of potential adopters, that is, m - C_{t-1}, will adopt the product. So, this value depends upon the likelihood of a new adoption. So, m - C_{t-1} this value depends upon the likelihood of a new adoption.



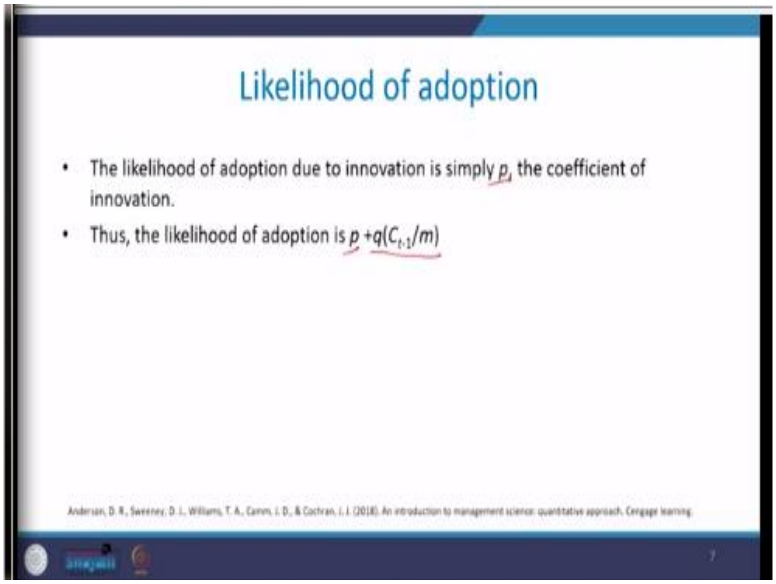
The likelihood of a new adoption is the likelihood of adoption due to imitation plus the likelihood of adoption due to innovation. So, there are two possibilities: someone will adopt the product, one due to imitation or due to innovation. The likelihood of adoption due to imitation is a function of the number of people who have already adopted the product. Obviously, the q represents there is a likelihood of adoption due to imitation in the function of a number of people that have already adopted the product.

The larger the current pool of adopters the greater their influence through word of mouth and social media. Because the C_{t-1} divided by m this term is the fraction of the number of people estimated to adopt the product by time (t – 1), the likelihood of adoption due to imitation is computed by multiplying this fraction by q the coefficient of imitation. Thus, the likelihood of adoption due to imitation is there are 2 term is there, one is this many people.

And this term is multiplied by this q, which is nothing but your coefficient of imitation. You see that this is proportional to this C_{t-1} divided by m; it is like probability. So, that probability is multiplied by this term q. If there are more proportions, the product will be more, so the likelihood of adoption due to imitation is q multiplied by C_{t-1} divided by m, so this q will estimate it.

Thus, the likelihood of adoption due to imitation is

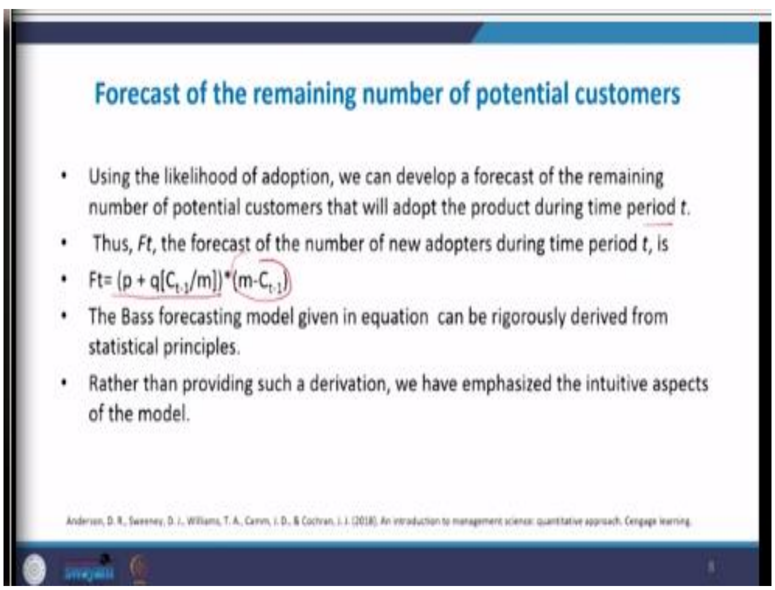
$$q*(C_{t-1}/m)$$



The likelihood of adoption due to innovation is simply p, the coefficient of innovation. Thus, the likelihood of adoption is p, what is the p? Coefficient of innovation plus this coefficient of imitation multiplied by C t - 1 upon m, = (C_{t-1}/m)

so this will be the overall likelihood of adoption, so p + q C t - 1 upon m.

$= p + q(C_{t-1}/m)$

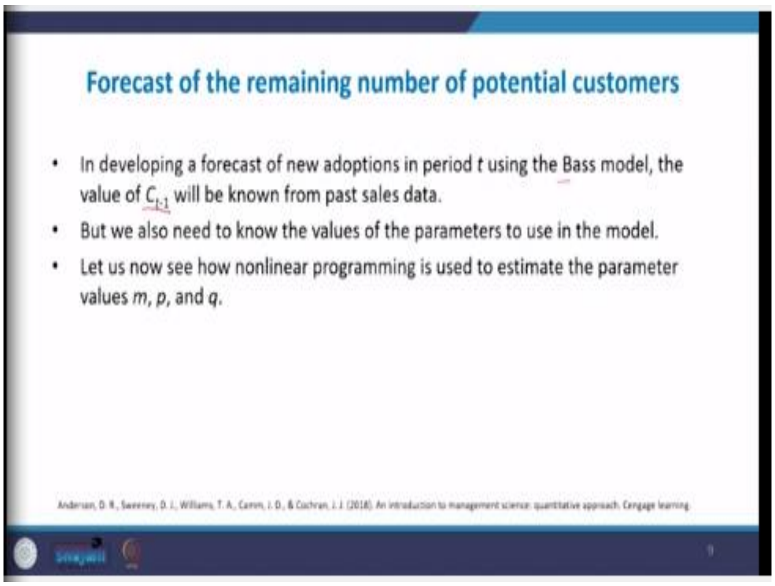


Now we will go for forecasting the remaining number of potential customers. Using the likelihood of adoption, we can develop a forecast of the remaining number of potential customers that will adopt the product during time t. Ft the forecast of the number of new adopters during time t is how we are doing?

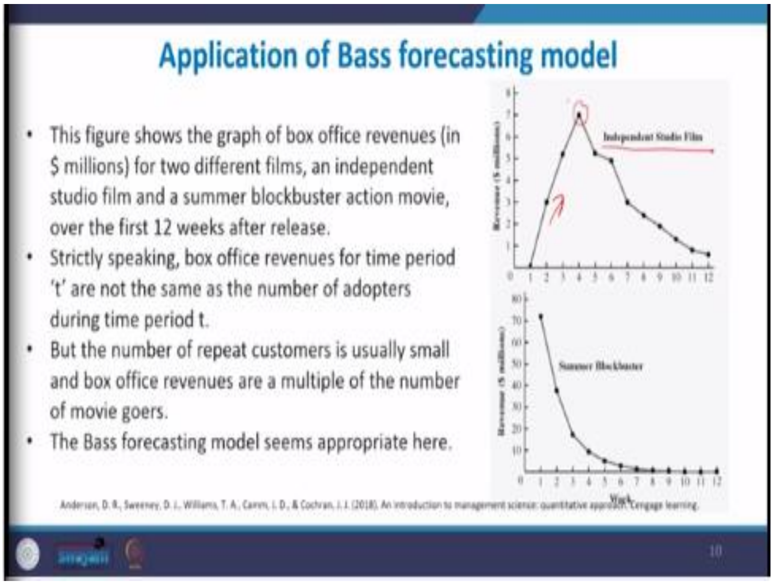
- $Ft = (p + q[C_{t-1}/m]) * (m - C_{t-1})$

So, this was the total number of adopters; this has the remaining people who have not adopted.

So, when you multiply that, we can see how many people are going to adopt the product. This model $p + q(C_{t-1}/m)$ multiplied by $(m - C_{t-1})$ this equation was given intuitively. The same Bass forecasting model given the equation can be rigorously derived from statistical principles. But we have brought this formula in an intuitive manner, so what have we done? Rather than providing such a derivation, we have emphasized the intuitive aspect of the model. So, you can intuitively get convinced by looking at this model Ft.



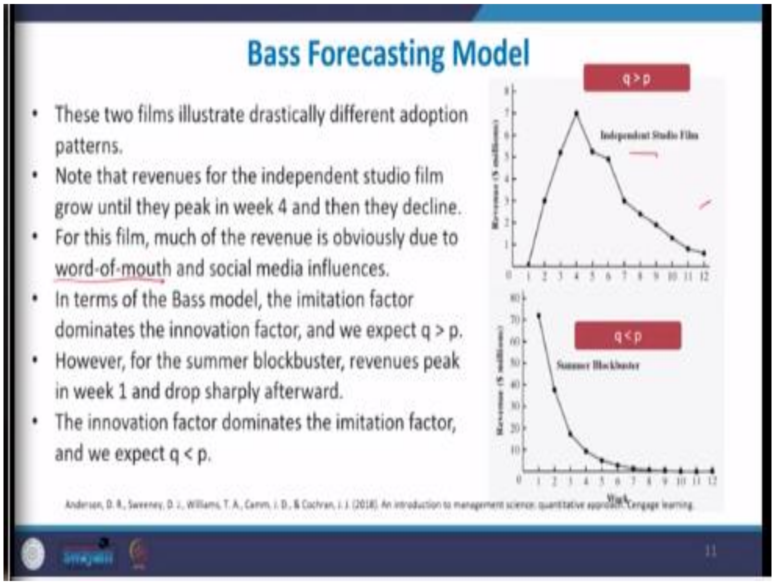
In developing a forecast of a new adoption in period t using the Bass model, the value of C_{t-1} will be known from the past sales data. What C_{t-1}? People who have already adopted the product, but we also need to know the values of parameters to use in this model; what are they? m, p and q. Let us now see how non-linear programming is used to estimate the parameter values of m, p, and q.



Now, I am going to explain some examples. With the help of that example, we are going to estimate the parameters. Now, this figure, which is on the right-hand side, shows the graph of box office revenue in a million dollars of 2 different films, an independent studio film and a summer blockbuster action movie, over the first 12 weeks after release. Strictly speaking, box office revenue for time period t is not the same as the number of adopters during the time period t.

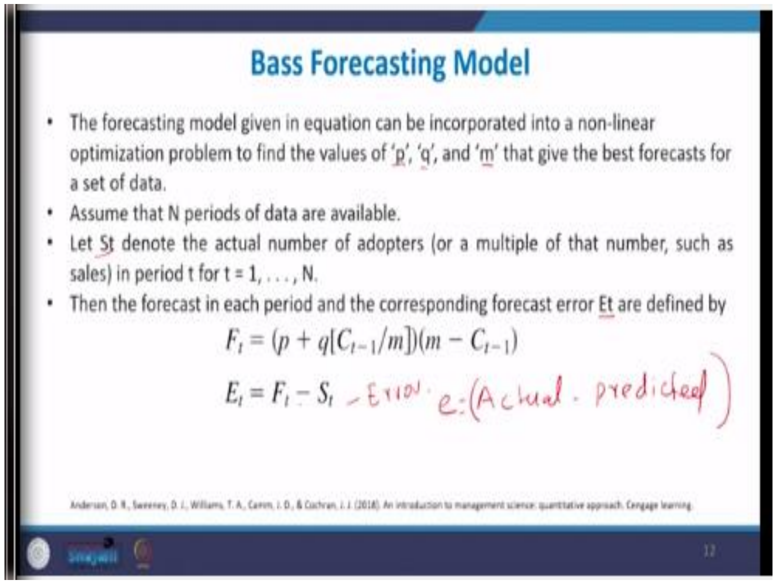
However, the number of repeat customers is usually small, and the box office revenues are a multiple of the number of moviegoers. So, the Bass forecasting model seems appropriate here. So, what are we going to see now? We are going to estimate how many people are going to watch the movies. There are two kinds of patterns, so product 1, for example, see the independent studio films; you see the y-axis is the revenues. Initially, the revenue increases, it goes top again, it decreases.

But you see the other movie that is a summer blockbuster. Initially, the revenue is very high, more people are watching, then when the time increases week increases, the revenue decreases because all the population might have watched. Here, word of mouth plays an important role so somebody goes to watch a movie and says that this movie is good, so he, somebody, and other people also come and watch, so the revenue peaks.



These two films illustrate drastically different adoption patterns. Note that revenue from independent studio films here grows until it peaks in week 4. See, up to week 4, the revenue is more, and then they decline. So, for this film, much of the revenue is obviously due to word of mouth and social media influences. That is why initially, the revenue is less, then it goes up, and then it comes back again. In terms of the Bass model, the imitation factor dominates the innovation factor. So, we expect the q to be greater than p, which is why we are getting this kind of pattern.

However, for the summer, blockbuster revenue peaked in week 1 and dropped sharply afterward. The innovation factor dominates the imitation factor, and we expect q to be less than p. So, in these 2 figures, we understand the top one, where the imitation factor is more, imitation factor is dominating word of mouth is dominating. In the second one, where the innovation factor is dominating, our job is to estimate the value of p and q.



Now, we will go for the Bass forecasting model. The forecasting model given in the equation can be incorporated into a non-linear optimization problem to find the values of p, q, and m that give the best forecast for a set of data. Where p is the innovation factor, q is the imitation factor, and m is, we can say, the number of people eventually who are going to adopt the product. Assume that the N periods of data are available.

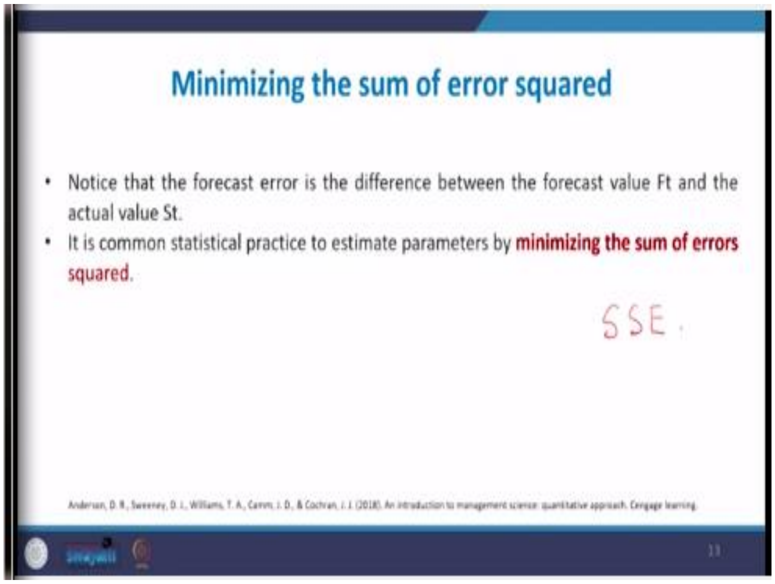
Assume that N periods of data are available.

Let St denote the actual number of adopters (or a multiple of that number, such as sales) in period t for t = 1, . . . , N.

Then the forecast in each period and the corresponding forecast error Et are defined by

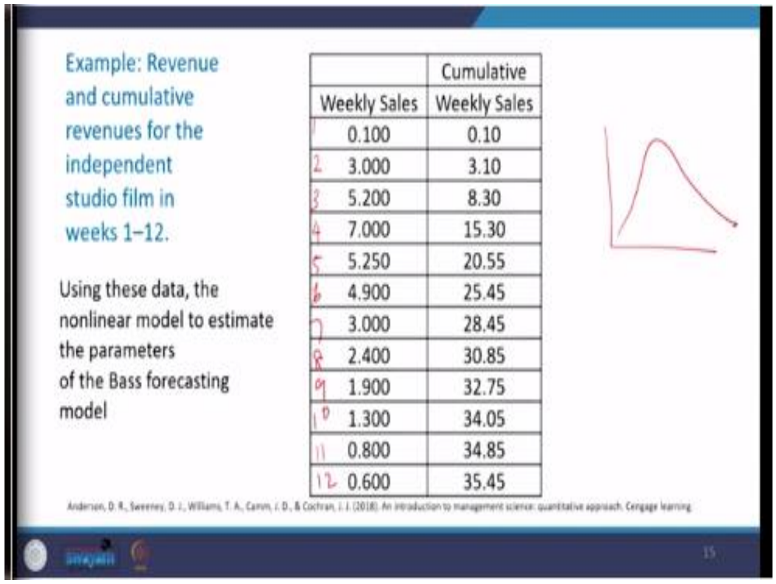
$$F_t = (p + q[C_{t-1}/m])(m - C_{t-1})$$

$$E_t = F_t - S_t$$



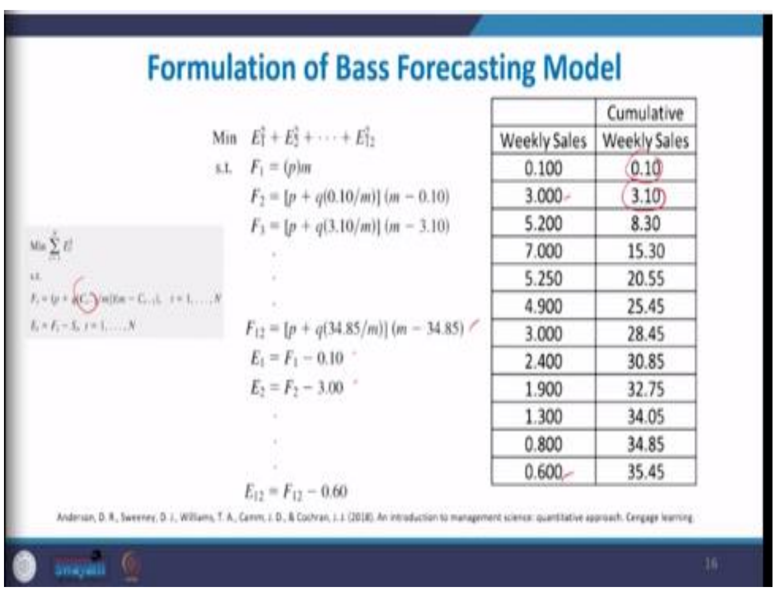
Here is what we are going to do for the forecasting model. We are using the concept of minimizing the sum of error squared. Notice that the forecast error is the difference between the forecast value Ft and the actual value St. It is a common statistical practice to estimate parameters by minimizing the sum of the error squares. In the regression context, we will say SSE minimizing error sum of square; if it is smaller, then we can say this model is more accurate.

Let us see the complete forecasting model; where there is an objective function, there is a constraint. Here, the objective function is minimizing the sum of the square of the error; here, what is a constraint? Ft equal to our forecasting model $F_t = (p + q[C_{t-1}/m])*(m-C_{t-1})$, here the error term is predicted value minus actual value. Because equations 1 and 2 both contain non-linear terms 1 and 2, this model is a non-linear minimization problem; what are we minimizing here? We are going to minimize the sum of the square of the error. But you should remember here the decision variables are p, q, and m.



Weekly Sales	Cumulative Weekly Sales
0.100	0.10
3.000	3.10
5.200	8.30
7.000	15.30
5.250	20.55
4.900	25.45
3.000	28.45
2.400	30.85
1.900	32.75
1.300	34.05
0.800	34.85
0.600	35.45

Now, I have taken an example, also taken from the book by Anderson et al., Revenue and cumulative revenues for the independent studio film in weeks 1 to 12. We have seen the pattern, what was the adoption factor there? It is like this, so what do we have to do? Using this data, the non-linear model to estimate the parameters of the Bass forecasting model will be discussed. What data is given? Weekly sales are given 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, so twelve weeks is there and it is sales is given. Then we found the cumulative sales 0.1, 3.1, 3.1 + 5.2, 8.3 up to 12 weeks.



Weekly Sales	Cumulative Weekly Sales
0.100	0.10
3.000	3.10
5.200	8.30
7.000	15.30
5.250	20.55
4.900	25.45
3.000	28.45
2.400	30.85
1.900	32.75
1.300	34.05
0.800	34.85
0.600	35.45

Now I will discuss the formulation of Bass forecasting model. There are 12 weeks, so there will be an error term for 12 weeks: E1 square + E2 square up to E12 square. For week 1, the F1 = p into m, what is the p? It is the innovation factor; there will not be any q, and there will not be any imitation factor. Because that person is coming to watch the movie, not by listening to somebody's reviews, he himself comes.

$$
\begin{aligned}
\text{Min} \quad & E_1^2 + E_2^2 + \cdots + E_{12}^2 \\
\text{s.t.} \quad & F_1 = (p)m \\
& F_2 = [p + q(0.10/m)]\,(m - 0.10) \\
& F_3 = [p + q(3.10/m)]\,(m - 3.10) \\
& \quad \vdots \\
& F_{12} = [p + q(34.85/m)]\,(m - 34.85) \\
& E_1 = F_1 - 0.10 \\
& E_2 = F_2 - 3.00 \\
& \quad \vdots \\
& E_{12} = F_{12} - 0.60
\end{aligned}
$$

The error for week 1 is F 1 minus actual data is 0.1; actual sales in week 2 errors are F2 - 3, and for week 12, F12 - 0.6.

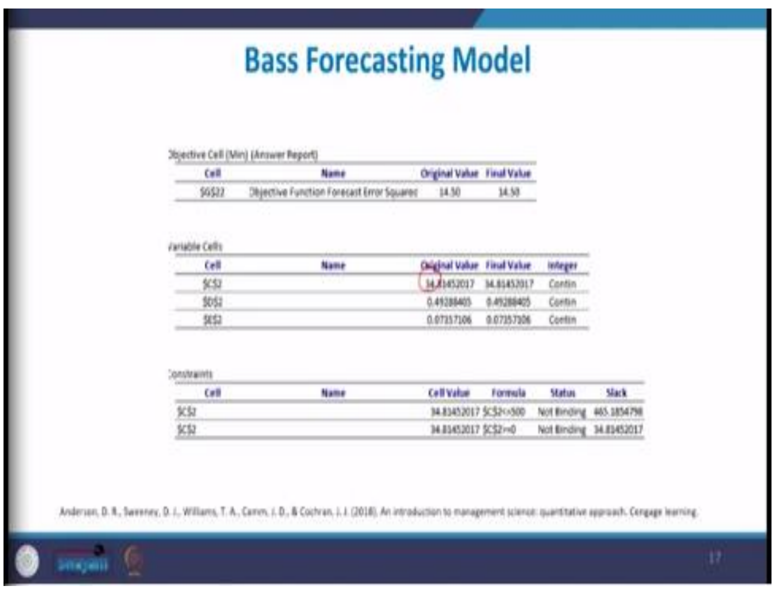
$$
\begin{aligned}
& \text{Min} \sum_{t=1}^{N} E_t^2 \\
& \text{s.t.} \\
& F_t = (p + q[C_{t-1}/m])(m - C_{t-1}), \quad t = 1, \ldots, N \\
& E_t = F_t - S_t, \; t = 1, \ldots, N
\end{aligned}
$$

I am going to solve this model with the help of a solver. Now, I am going to solve this non-linear problem with the help of a solver. So, I have written C2, where the value of m was my final answer, and I will show you how I got this answer. Then D2 is the value of q, then C2 is the value of p, which is the coefficient of innovation.

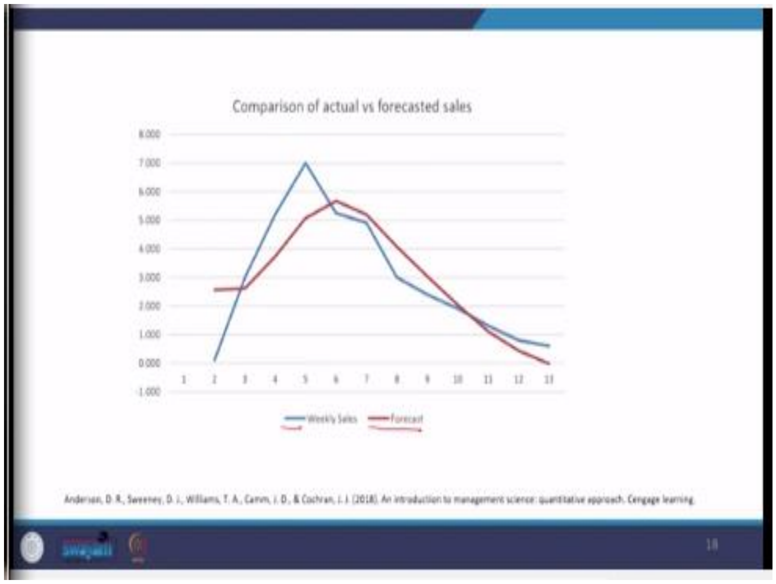
Now how I am forecasting, suppose if we go E9, E9 is E2 multiplied by C2, look at the equation which I have written. Because F1 = p into m, then F 2 the column which is down p + q multiplied by 0.10, 0.10 is where the D9 is there, upon m, m is my C2 multiplied by m - 0.10, m is C2 - D9 like that I have written for all 12 weeks. Then what is the error? You go to F9, and F9 is the predicted value - the actual value that is C9; what is C9?

C9 is the actual value - the predicted value is E9, similarly for, say, F10, when you go for F10, it is C10 - E10. So, what have I done in the G column? I have squared the error. I have added all the squares, so that is our G21, G21 is I have added this squared value. So, now when I go to data and solver, you see this is a minimization problem. My objective function is the G22 sum of the square of the error that has to be minimized.

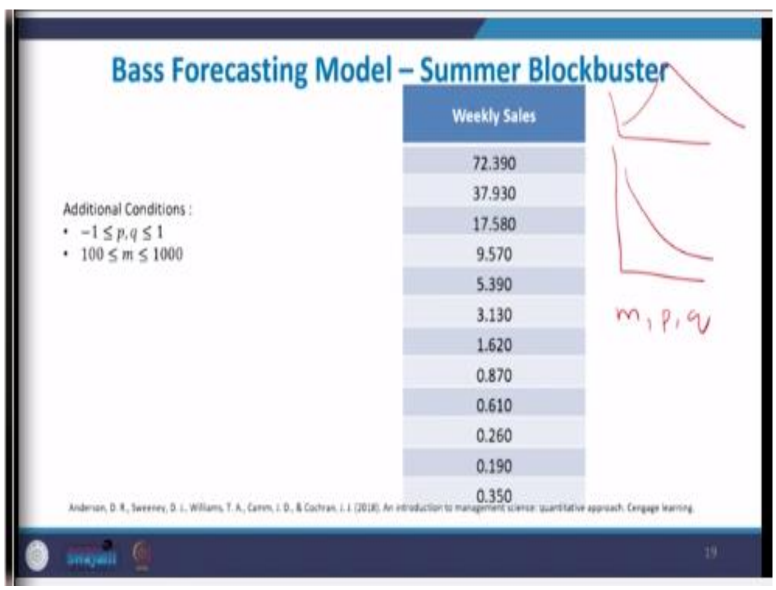
This is a very important point to remember: here, the changing cell is our value of m, p, and q. And the C2 is greater than or equal to 0. So, when I solved it using CRG non-linear, I got this result. So, you can go back; the result is the value of m is 34, q is 0.49, and p is 0.07. Here m, I did not put any upper limit on the total number of people eventually who are going to adopt it. If you want, you can put the approximate population size that will be a more accurate answer. So, now I have taken a screenshot of this answer in my presentation.



So, the value of m is 34, and the values of p and q are given. One point you should remember is that this model is neither convex nor concave, so every time, you get a different result. So, that means that you may not get the same answer because we are getting a local optima solution, which is the local minimum solution; it is not the global minimum solution.



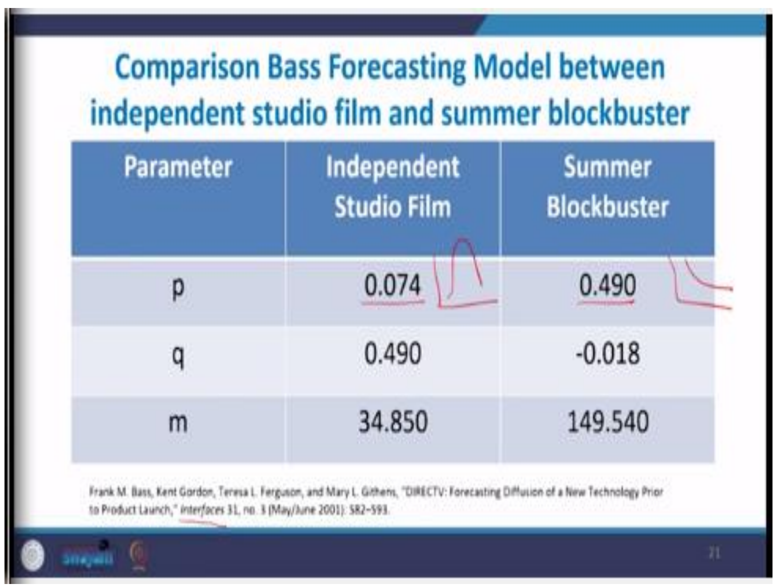
When I forecasted and compared the error, you see that there is a similar pattern, so this is the forecast value. The blue one says the actual value; you see that the red one is the forecasted value, and there is a similar pattern between the actual value and the forecasted value.



Now, there is another problem; another example is forecasting or finding the value of m, p, and q for summer blockbusters. So, this data is given weekly sales data; it is like the same dataset. So, this dataset is given, but here, the adoption pattern is different; you see that here, the adoption pattern is like this; previously, the adoption pattern was like this. So, for this type of model also, we can find out the value of m, p, and q.

So, when I run this model, I am also going to solve it with the help of a solver. Now, this is an excellent model for a summer blockbuster movie. So, what has happened? It is the same thing, only the pattern of the data is different from the weekly sales. You see, initially, it is 72, then it keeps on decreasing; all other models are the same. For example, forecast E9, see the F2, E2 - C2 the same thing that you have done previously.

Here also, we are finding the error, then squaring the error and summing the error. So, when I go to data, when I click on the solver, I see the requirement. Here, I am minimizing the sum of the square of the error, so I want to get the changing cell as m, p, q. Here, I am keeping a cut-off. So, the value of m is between 1000 and 100 because these are other probabilities, so I kept less than or equal to 1. So, when I solve it, I get the value of m is 149, the value of q is 0.01, and the value of p is 0.48. So, why have I solved it? I am going to compare the previous model with this model.



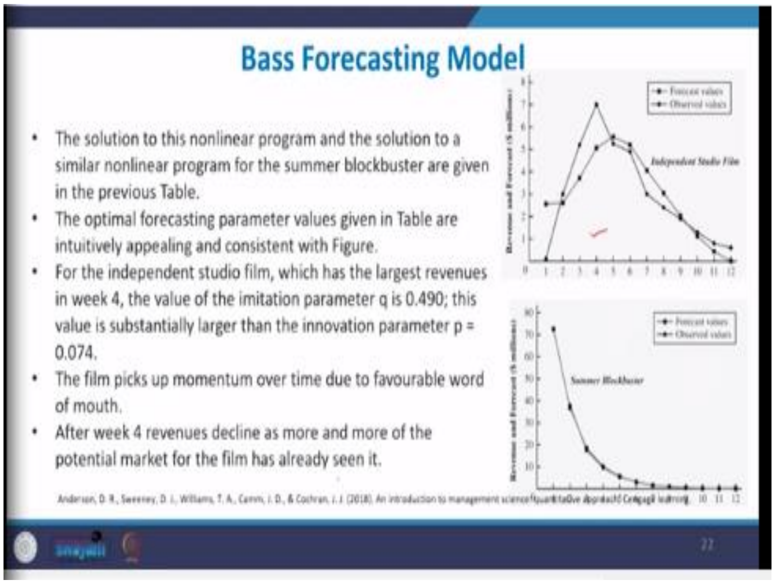
Comparison Bass Forecasting Model between independent studio film and summer blockbuster

Parameter	Independent Studio Film	Summer Blockbuster
p	0.074	0.490
q	0.490	-0.018
m	34.850	149.540

Frank M. Bass, Kent Gordon, Teresa L. Ferguson, and Mary L. Githens, "DIRECTV: Forecasting Diffusion of a New Technology Prior to Product Launch," *Interfaces* 31, no. 3 (May/June 2001): S82–S93.

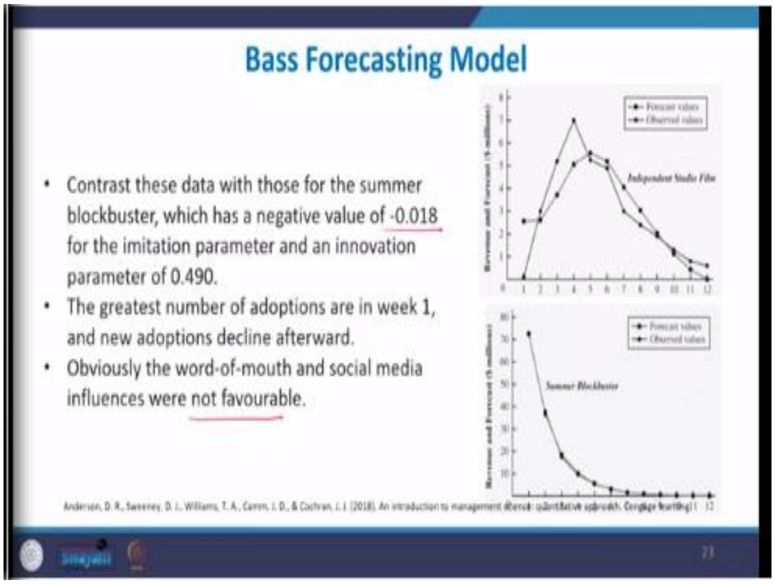
So, what has happened? Now you see the pattern, the comparison of Bass forecasting model between independent studio film and summer blockbuster. You see that the independent studio film the value of p is smaller, what is p? The innovation factor is smaller but for a summer blockbuster innovation factor is high, that is why the pattern was like this, here the pattern was like this.

You see the q, for independent studio films the q the imitation factor is more but for summer blockbuster the imitation factor is less. And the m we can put it, this m we can predict, this will change based on our population size. This is a very famous model; the reference for this model is Frank M Bass et al., See a DIRECTV forecasting diffusion of new technology prior to product launch, this paper is published in interfaces. So, the mathematics behind this model has come from this research paper.

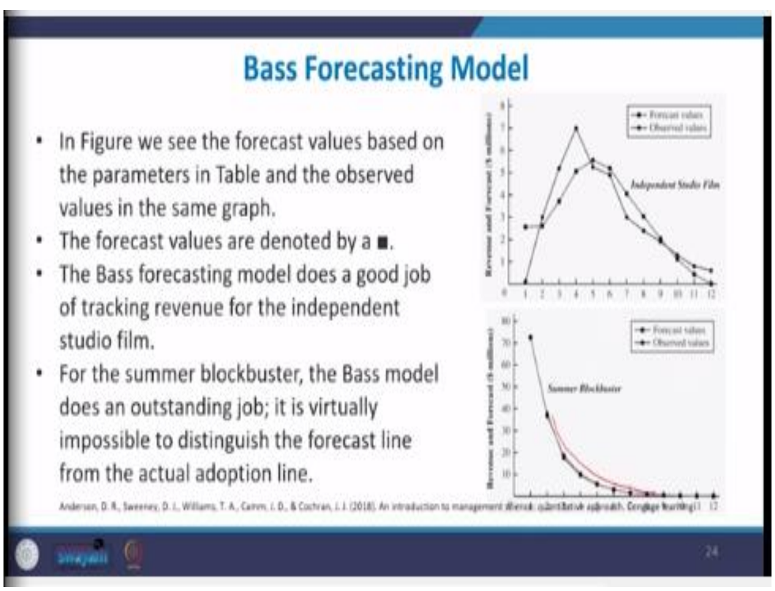


The solution to this non-linear program and the solution to a similar non-linear program for the summer blockbuster are given in the previous table; we have seen the comparison. The optimal forecasting parameter values given in the tables are intuitively appealing and consistent with the figure. For the independent studio film, which is this one, the top, which has the largest revenue in week 4, and the value of the imitation parameter is 0.490.

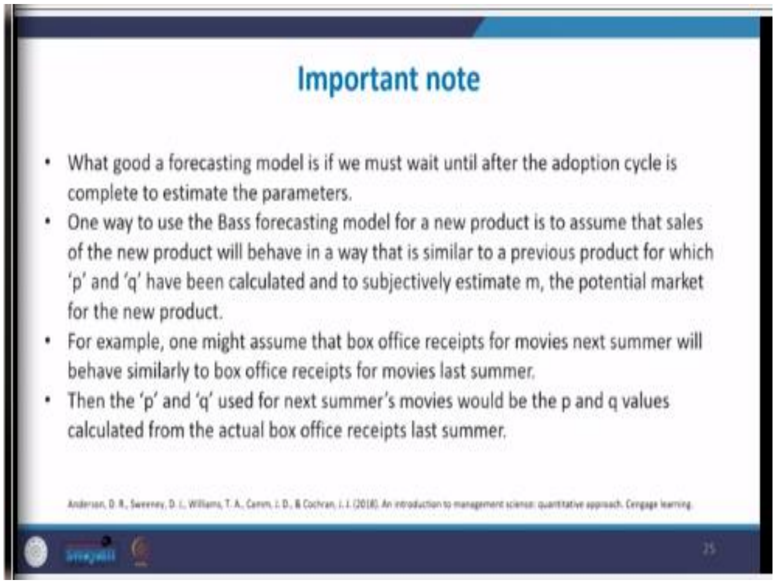
This value is substantially larger than the innovation parameter p, the value of p is only 0.074. So, what do we understand? The film picks up momentum over time due to favourable word of mouth. After 4 weeks, the revenue declined as more and more of the potential market for the film has already seen it.



Contrast these data with those of Summer Blockbuster, which has a negative value of -0.018 for the imitation parameter and the innovation parameter is 0.494. So, that means for the summer blockbuster movie, the innovation parameter is high, but the imitation factor is very low. So, what do we understand? The greatest number of adoptions are in week 1, and new adoptions decline afterward. Obviously, the word of mouth and social media influences were not favorable where summer blockbusters, initially there were more viewers, but after that, it kept on declining.

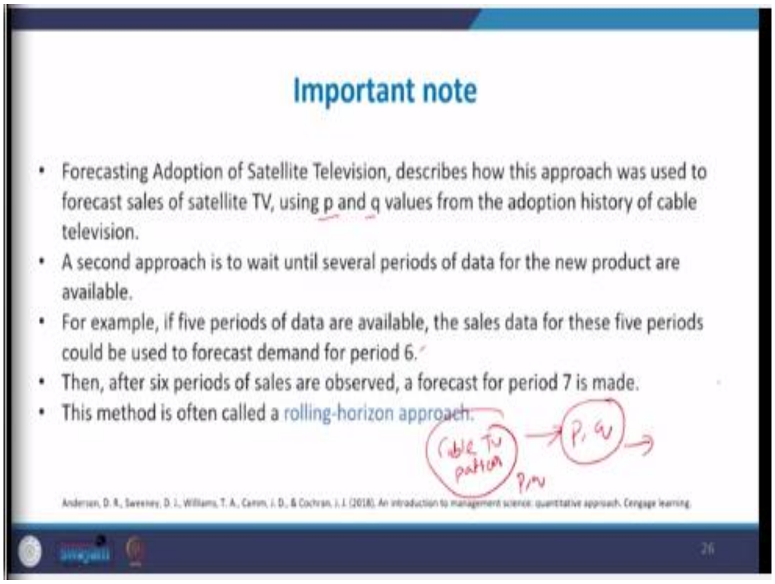


In the figure, we see the forecast value based on the parameters in the table and the observed values in the same graph. So, what do we understand? The forecast values are already denoted by this square. The Bass forecasting model does a good job of tracking revenue for independent studio films; you see that this revenue pattern is a very nice pattern we have seen in Excel also. For summer blockbusters, the Bass model does an outstanding job; it is virtually impossible to distinguish the forecast line from the actual adoption line. You see that here, both completely overlapping this Bass forecasting model.



So, the important point you should remember is what good a forecast model is if we must wait until after the adoption cycle is complete to estimate the parameters. You may ask this question: how do we get these parameters m, p, and q until the adoption cycle is complete? So, one way to use this Bass forecasting model for a new product is to assume that the sales of the new product will behave in a way that is similar to the previous product for which the p and q have been calculated and to subjectively estimate m, the potential market for the new product.

For example, one might assume that the box office receipts for movies next summer will behave similarly to box office receipts for movies last summer. So, the value of m, p, and q can be compared with our previous models or the products that have a similar nature. Then, the p and q used for next summer's movies would be the p and q values calculated from the actual box office receipts from last summer.



Forecasting the adoption of satellite television describes how this approach was used to forecast sales of satellite TV, using p and q values from the adoption history of cable television. So, what I am trying to say here is, from the adoption pattern of cable TV, they got the values of p and q. With the help of this, p and q, they have estimated the adoption pattern of satellite television. So, the values of p and q were estimated from an activity that is similar to what we are currently forecasting.

The second approach is to wait until several periods of data for the new product are available. For example, if 5 periods of data were available, the sale of data for these 5 periods could be used to forecast demand for period 6. Then, after 6 periods of sales are observed, the forecast for period 7 is made. So this method is often called the rolling-horizon approach. So what we understand from this.

The value of p and q innovation factors and imitation factor are estimated from the previous product which has similar adoption factors. For example, what has done? So, from the cable TV adoption pattern we got the value of p and q. So, this p and q can be used for forecasting the adoption pattern of satellite TV. In this lecture, I have explained another application of non-linear programming, that is, the application was forecasting the adoption pattern of a new product.

So, in this forecasting model, we have estimated 2 parameters, that is, a value of p and q; p is the coefficient of innovation, and q is the coefficient of imitation. After estimating these p and q, these parameters can be used for forecasting a similar product that has a similar adoption pattern. In the next class, I will explain another example of or another application of non-linear programming; thank you very much.