Software Project Management
Prof. Durga Prasad Mohapatra
Department of Computer Science and Engineering
Indian Institute of Technology, Rourkela

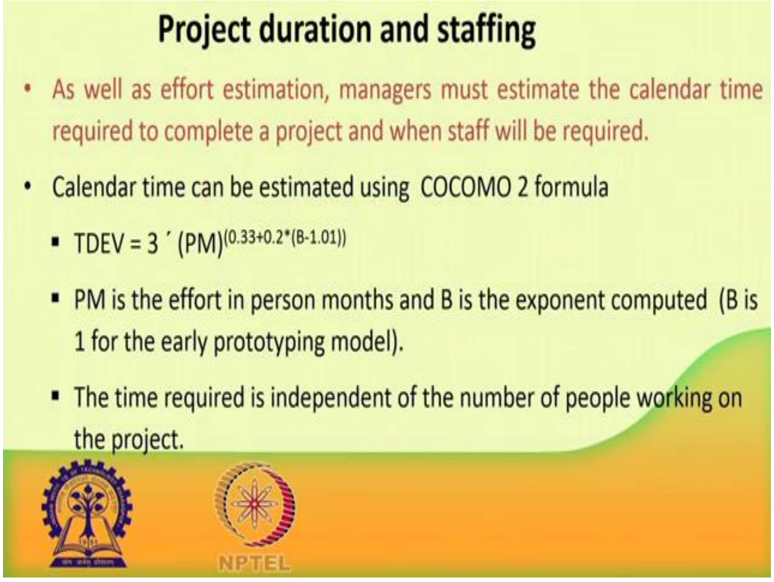
Lecture – 25
Project Estimation Techniques (Contd.)

Good afternoon. Will now today see one of the important estimation aspect of the projects that is staffing level estimation.

(Refer Slide Time: 00:24)

Here we will see two important models for staff level in estimation that is Putnam's model and Jensen's model.

(Refer Slide Time: 00:33)



So, let us see about this project duration and staffing. So, as well as effort estimation managers must estimate the calendar time required to complete a project and when the staff will be required. Let us say effort estimation it is at all very much important. Similarly, estimating the calendar time required to complete a project, as well as how many staffs will be required when they will be required these are also important aspects of a project estimation.

So, the calendar time for a given project can be estimated using COCOMO 2 formula. Here T development is; that means, calendar time is represented as

$$TDEV = 3 * (PM)^{(0.33+0.2*(B-1.01))}$$

So, here this formula is obtained by using COCOMO 2 estimation technique and here PM represents the effort in person months and B is the exponent computed, normally, B is one for early prototyping model.

The time required is independent of the number of a persons working on the project. These time developed it is does not depend upon that how many persons are working in the project. It is the time the development time required is independent of the number of people working on the project.

(Refer Slide Time: 02:05)

So, the now let us see about the staff requirement, how many staffs generally do you require for a given project? Staff required it cannot be computed just by dividing the development time by the required schedule not at all possible. Because the number of a people working on a project it varies depending on the different phases of the project.

You might require less number of people during the requirements analysis maybe again more number of people during design. And during coding again it may either increase or decrease depending on the nature of the project and in testing time again you may require more number of people. So, that is why it will vary the number of people working on a project, it will vary it will depend on the phase of the project that you are currently in. The more people who work on the project the more total effort is usually required this very obvious. If more number of people work on the project; obviously, the total effort required will be more. So, a very rapid buildup of people upon correlates with the a schedule slippage.

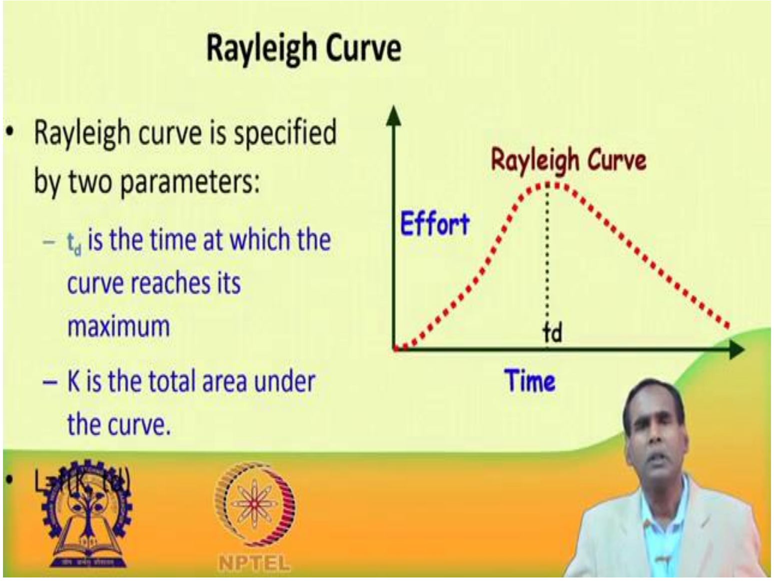
(Refer Slide Time: 03:10)

Now, let us see how to estimate the staffing level. The number of personnel required during any development project is not constant I have already told you that it may depend upon the what the phases. So, that is why the number of personal required during the development of any project is not constant.

It will may be less at the initial time; that means, this requirement specific analysis and specification, then again more number of people will be required during design. And the highest number of people may be required during the testing time, so that is why it is not constant it varies.

Now, then in 1958 he has analyzed he had analyzed many R and D projects, and observed the followings; what he observed? He observed that the Rayleigh curve represents the number of full time personnel required at any time. That means, the number the number of full time personnel required at any time for a project can be represented by using a Rayleigh curve.

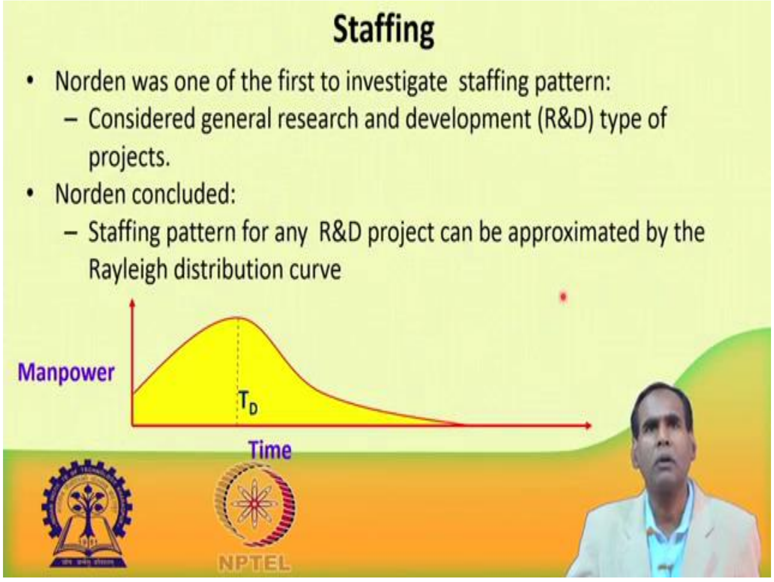
(Refer Slide Time: 04:12)



So, let us see what is a Rayleigh curve? I hope you must have known Rayleigh curve earlier see this is the form of a Rayleigh curve. Here in the x axis we have taken the time and y axis we have taken the effort. You can see as the time moves the effort also it gradually increases and there will be time come peak period and after which even if the time increases the effort will slight it will fall down it will decrease.

Rayleigh curve is specified by two important parameters; one is the t d, another is K, t d is the time at which the curve reaches its maximum we can see. So, the td is the time at which the curve reaches its maximum and the K is the total area under the curve, so total area under the curve is represented by K. So, Rayleigh curve is specified by using the two these two parameters td and K.

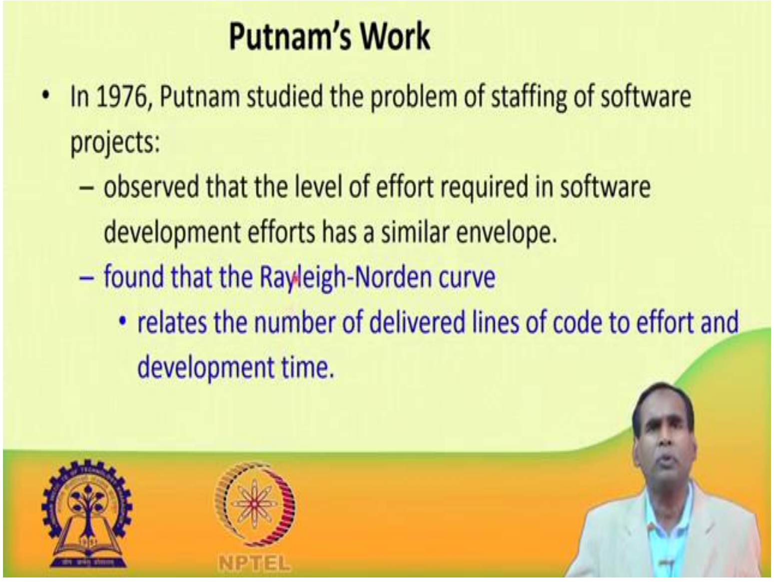
(Refer Slide Time: 05:09)



Norden was one of the first persons or to investigate the staffing pattern. He considered some general R and D kind of a projects, he concluded that this staffing pattern for any R and D project can be approximated by the Rayleigh distribution curve.

So, he has observed several projects and then what concluded that this staffing pattern then; that means, the number of staffs required for developing a project it will follow on R and D project ok. Sorry the number of people required for developing an R and D project it can be approximated by a Rayleigh distribution curve as shown below, and Rayleigh distribution curve is specified by two parameters T d and K.

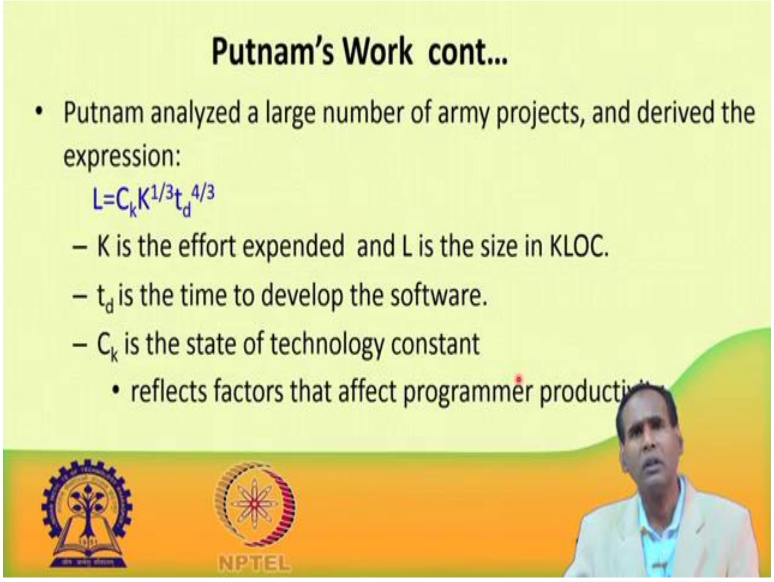
(Refer Slide Time: 06:04)



Then Putnam came, so Putnam has also done some work, he has studied the problem of staffing of software projects. He has also what investigated many projects in 1976 Putnam studied the problem of staffing of software projects. He observed that the level of effort required in software development projects has a similar envelope ok. He observed he has observed that the level of effort required in software development has also a similar effort because previously Putnam this Norden had studied normally what some of the what R and D projects.

So, then Putnam had studied as some what this problem of staffing of software projects, he observed that the level of effort required in software development has also a similar envelope that happens in case of R and D projects. He also found that the he also found that the Rayleigh Norden curve it relates the number of delivered lines of code to the effort and development time.

(Refer Slide Time: 07:26)



So; that means, you will see this Rayleigh Norden curve it also relates the number of the delivered lines of code to what to effort and development time. On Putnam analyzed he has studied a large number of army projects and derived the expression for L. And L is specified as

$$L=C_kK^{1/3}t_d^{4/3}$$

Where K is the effort expended or the export the effort required and L is the size in terms of KLOC, t d is the time to develop the software, C k is the is a constant it is known as the state of technology constant.

This C k which is known as technology constant it reflects the different factors that affect the programmer productivity, as you have seen in earlier cases COCOMO model etcetera, there are many factors they affect the programmer productivity. Here C k also represents the different factors that the that affect the programmers productivity.

(Refer Slide Time: 08:25)

Putnam's work has established something Putnam has stated the followings. That after conducting several experiments he has observed that C k the value of C k is equal to 2 for poor very poor development environment where no software engineering methodology, no proper methodology used whatever poor documentation is there.

And no reviews there or very poor reviews there for that the value of ck may be set to 2, but what where or the organizations where good software development and environment is there. Software engineering principles are used, good programming languages are used etcetera, then the value of C k may be set to 8 and if the environment is excellent.

So, using various current software engineering practices, modern programming practices uses what you can see now automated tools case tools etcetera and then the value of ck may be set to 11. Or very small number of engineers are needed at the beginning of the project carrier for carrying out planning and specification.

We can see at the beginning only very small number of manpower are required for carrying out the planning and specification. That means, requirement analysis and specification ok, planning and requirements specification and analysis activities. As the progress as the project progresses more detailed work is required. See the time is passing are the project progresses more number of people are required because what more detailed work is required. That is why the number of engineers slowly increases and reaches peak as the time progresses as more number of work is required. Now, let us see this peak time represents which phase of software development.

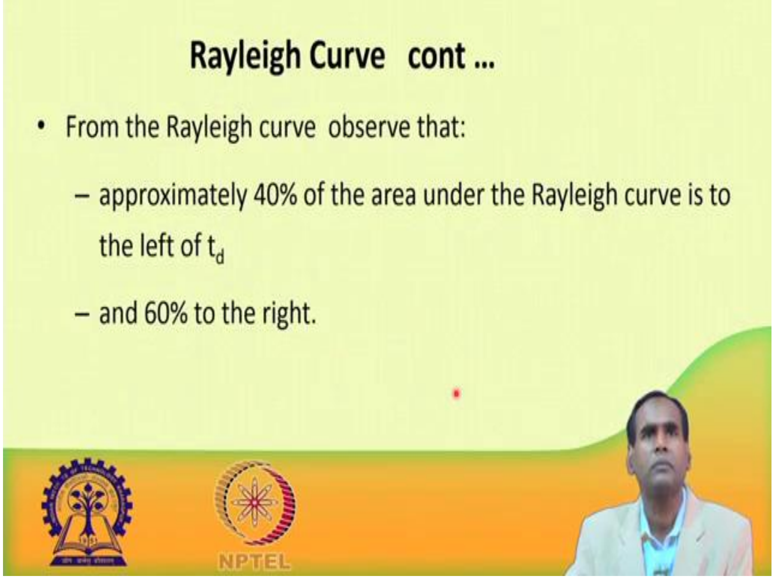
(Refer Slide Time: 10:21)

Putnam observed that the time at which the Rayleigh curve reaches it is maximum value it corresponds to the system testing after product is released. So, you see initially I have told this is the during the project planning or this what requirement analysis specification analysis phase; then design will come, then coding will come.

So, this peak period T D will correspond to the testing phase that using system testing phase. Putnam observed that the time at which the Rayleigh curve reaches it is maximum value it roughly corresponds to the testing phase, the system testing phase and product release.

Then after system testing is over, the number of project staff it falls gradually till the product installation and delivery. So, after this is over see this is the system testing this T D this point corresponds to the system testing phase. And after the testing phase is there over then gradually the manpower required will be fall down, because now the system will be implemented at this side and then gradually the number of person will be at falling down that is what he says. So, Putnam observed that, after system testing is over the number of a project staff it gradually falls down till the product installation and delivery it can still the product installation over and delivery comes.

(Refer Slide Time: 11:46)

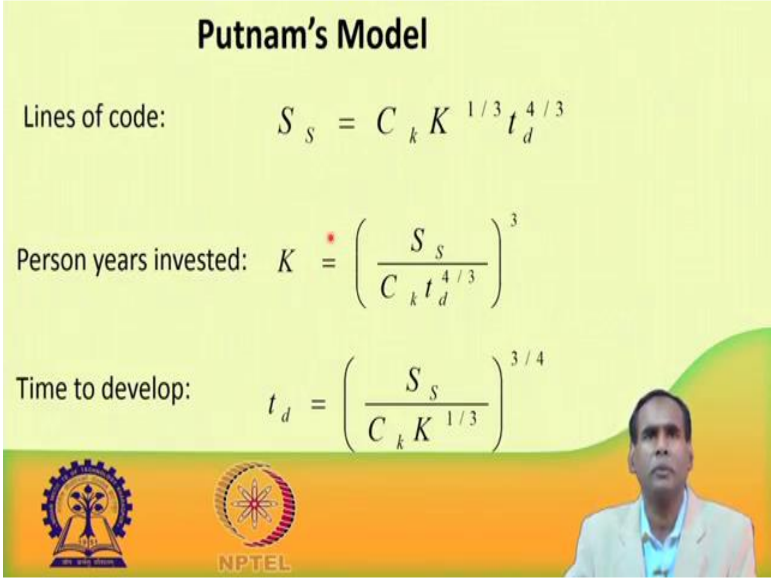


From the Rayleigh curve it is also observed that approximately 40 percent of the area on the under the curve is to the left of the T D and 60 percent to the right of the T D. You see it is also observed that almost 40 percent of the area what lies to the left of the curve and 60 percent lies to the right of the curve sorry to the left of the T D. So, 40 percent lies to the left of this point T D and 60 percent area lies to the right of this point T D.

(Refer Slide Time: 12:29)

Now, let us see Putnam's model can be used to estimate what? Putnam's model can be used to estimate the lines of code which is represented as Ss. It can also be used to estimate the person years invested the time to develop and a constant called as technology coefficient constant C k.

(Refer Slide Time: 12:59)



Putnam has given these formulas. So, that the lines of code can be computed by using the formula

$$S_S = C_k K^{1/3} t_d^{4/3}$$

the meanings of the term solid who have told you earlier. From this formula you can easily find out the value of K; that means, the person years invested.

$$K = \left(\frac{S_S}{C_k t_d^{4/3}} \right)^3$$

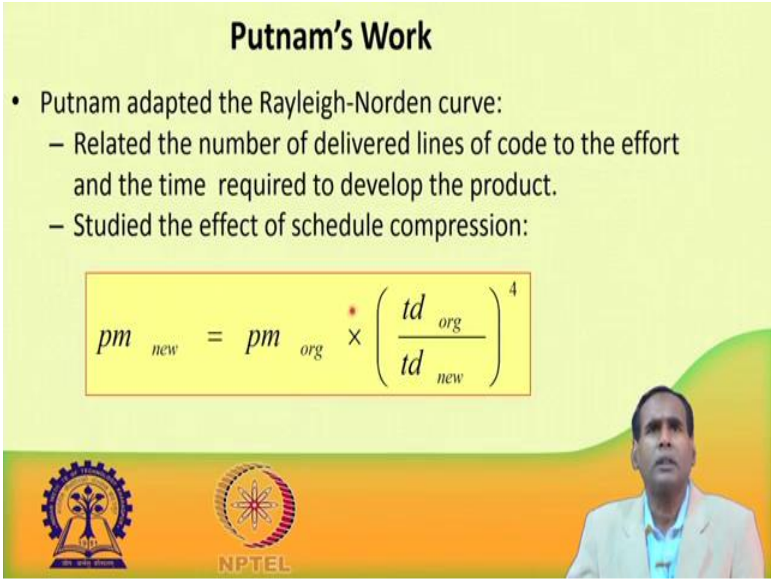
That means, you can say this is the effort required maybe in terms of person month. This will be equal to S s, S s is the lines of code divided by C k into to the power 4 to the power or 4 by 3. Also from this equation lines of code, you can easily derive the value of td

$$t_d = \left(\frac{S_S}{C_k K^{1/3}} \right)^{3/4}$$

which is equal to S s divided by C k into k to the power 1 by 3.

So, given any, so lines of code there are you can see that four things, S s C k K and t d. So, if three parameters are given C k, K and t d are given easily you can find out the lines of code easily you can estimate the lines of code. So, similarly if the values of S s C k and t d are given you can easily estimate the value of K. And if the values of S s ,C k and K they are given you can easily estimate the time to develop that is td .

(Refer Slide Time: 14:14)



Putnam adopted this Rayleigh Norden's curve, he has related the number of delivered lines of code to what to the effort and the type ok. From the Rayleigh Norden curve he has related the he has established an relation between the number of delivered lines of code and the effort and the time required to develop the product. He has also studied the effect of schedule compression on the effort as well as the time. So, instead in general you can say he has studied the effect of scheduled compression on the cost.

So, he has established the followings

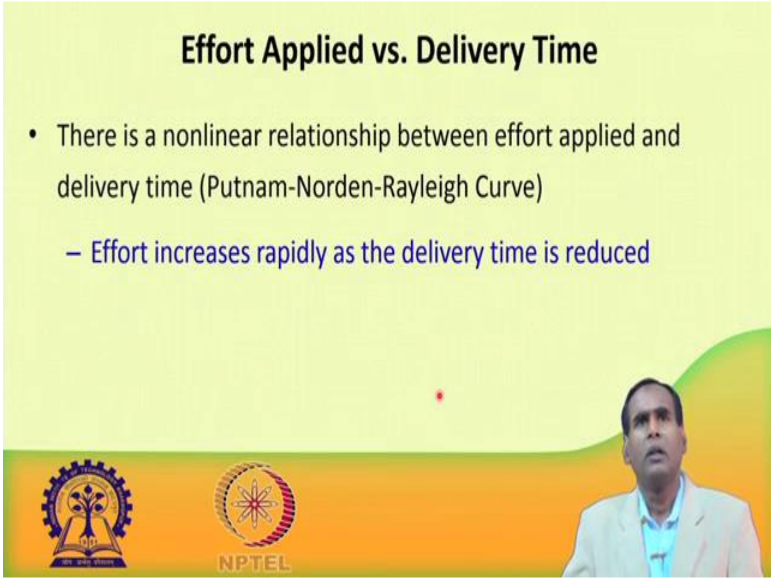
$$pm_{new} = pm_{org} \times \left(\frac{td_{org}}{td_{new}} \right)^4$$

Here td is this schedule the time to delivery or org means what originally you have planned to develop the what software, td new just the new time. Suppose you have estimated that 12 months will be required to develop the project.

Then the customer comes some urgency, due to some urgency they are requesting you delivery it within 6 months. So, now, td org is 12 this the original schedule and new schedule is suppose now 6. So, pm org means the effort required for the original schedule and here pm new means the effort that will be required for the new schedule, because now according to the new schedule time has been reduced.

So, if time will be reduced how much what more effort you have to put to develop the software that is represented here. So, pm new; that means, the new effort revised effort will be equal to the original effort into what the original time to delivery original time divided by the new time.

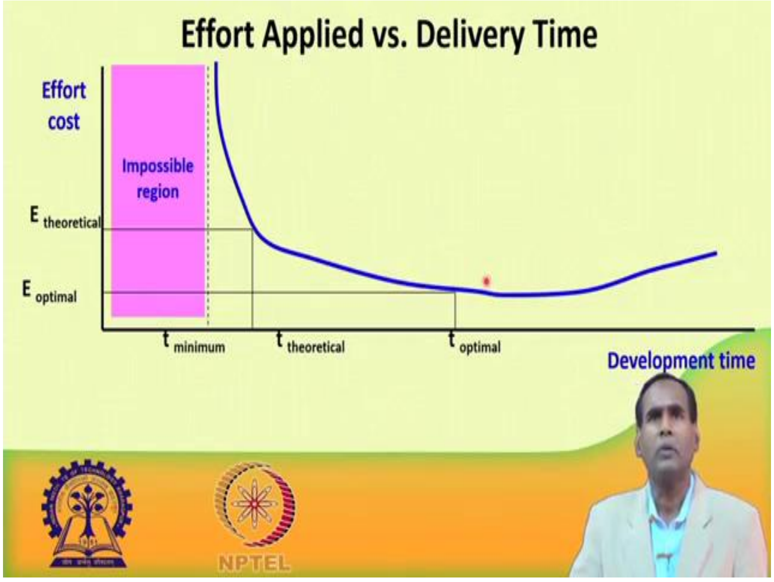
(Refer Slide Time: 16:10)



So, now let us see the effort applied versus the delivery time. There is a non-linear relationship between effort apply and delivery time according to put Putnam Norden Rayleigh curve ok. So, according to do according to the Putnam Norden Rayleigh curve, you can observe that there is a non-linear relationship between the effort that will be applied on the delivered time.

So, what does it mean that the effort increases rapidly as the delivery time is reduced not linearly. The if the time is increased the effort that does not increase just linearly, it increases very rapidly I will show one example. So, if the delivered time is reduced ok, here you can see that if the delivery time is reduced the effort increases very rapidly.

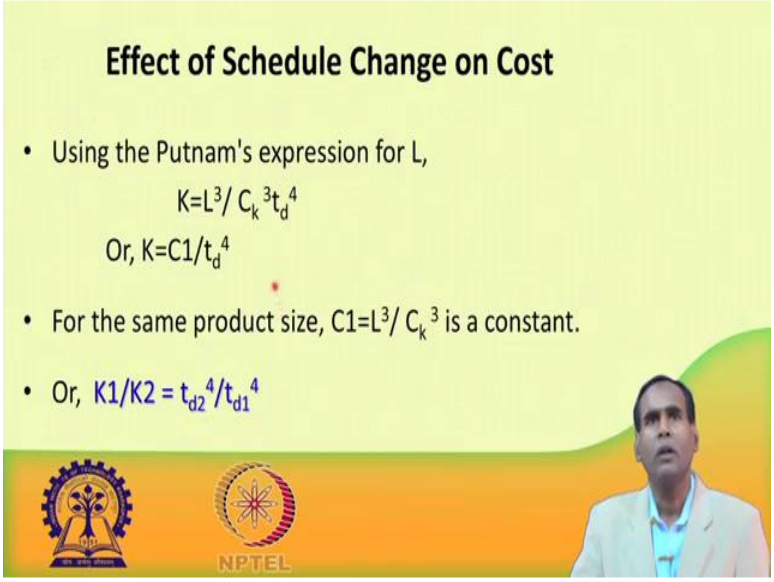
(Refer Slide Time: 17:02)



Let us see how? So, initially you can say this is the optimal time; that means, if this is the time to delivery the optimal report is required will be required it is given at this point. Now, you see there is the compression suppose, so theoretically this delivered time it is reduced this is t theoretical. Then what will happen how much is the required the effort is required what increases tremendously.

And still if you will compress the time, again you reduce the time; suppose at this point ok. This point then you will say that then it is almost impossible region you cannot what get your work done you cannot complete the project. So, there is certain point up to which you can reduce the time, you can reduce the schedule. Beyond that even if you will schedule how many even if how many more number of personnel you hire you may not complete the project let us see.

(Refer Slide Time: 18:05)



The conclusion says that what is the effect of schedule change on the cost? Let us see we have already seen by using Putnam's expression for L, L we have already seen earlier the value of L is equal to ok. The value of L is specified like this using Putnam's equation, from this you can find out the value of the C k or the value of K let us see.

By using that equation you can find out the value of K which is equal to

$$K=L^3/ C_k^{\,3}t_d^{\,4}$$

And you can see that if will simplify, you will get K is equal to

$$K=C1/t_d^{\,4}$$

some constant into 2 to the power 4. So, C1 is either constant and this will be C1 will be coming to be L cube by C k cube.

So, if you will what simply this statement you will get K is equal to how much? L cube by C k cube in into 1 by t d to the power 4. So, this L k cube by C k cube is a constant ah, so for the same product size C this L cube by C k cube against C k to the power 3 is a constant and this is represented as C k here. So, finally, K is equal to C 1 to the power ok.

So, final k is equal to C 1 to divided by t d to the power 4. Now, if you will take say a now on say there are two timings two schedules you will take t d 1 which one the original schedule you have estimated. Now, due to some urgency the customer is requesting you to deliver it before that time say that is t d 2, then you can get the equation you can get two values here K1 and the K2. So, then you take the ratio K1 by K2 is equal to

$$K1/K2 = t_{d2}^4/t_{d1}^4$$

this will be now reverse t d 2 to the power 4 by t d 1 4.

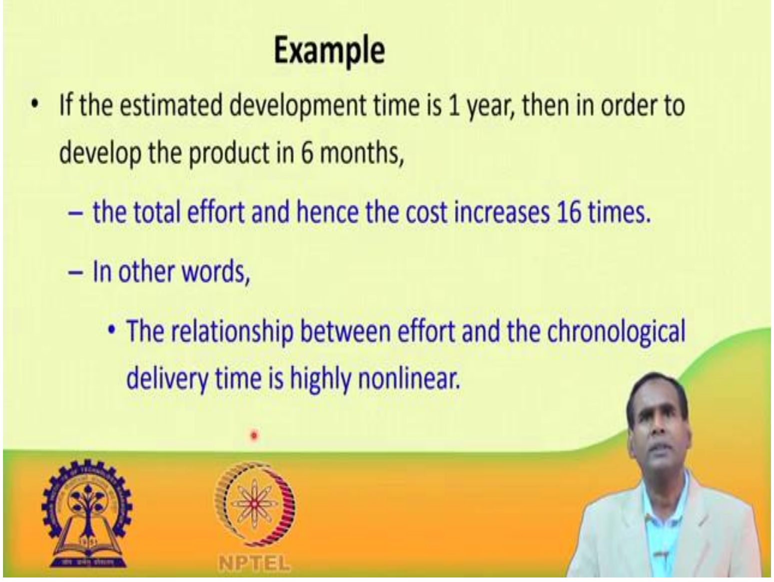
So, this is just a simplification what I have done, here I have found out the values for 2 times t d 1 and t d 2. And accordingly I will get two values K 1 and K 2. So, divide K 1 by K 2 it will be simplified as t d 2 to the power 4 by t d 1 to the power 4. So, what does it mean then what is it is what inference we can draw from this statement.

(Refer Slide Time: 20:25)

You can see that a relatively small compression in delivery schedule can result in substantial penalty on the human effort. See if you are reducing then how much time it is increasing almost 4 times, if you will reduce the job or sorry if you will reduce the time, so the delivered time. Then the effort that you have to put and hence the cost will be increasing very rapidly. So, this can result in, so if a relatively only a very small compression in schedule or in delivery schedule is required, it can result in a substantial penalty on the human effort.

You have to spend much more human effort, you have to spend much more cost. Also observed that the benefits can be gained by using fewer people over a somewhat longer time span, if you can use only few people for a long amount of time then you will get more profit.

(Refer Slide Time: 21:25)



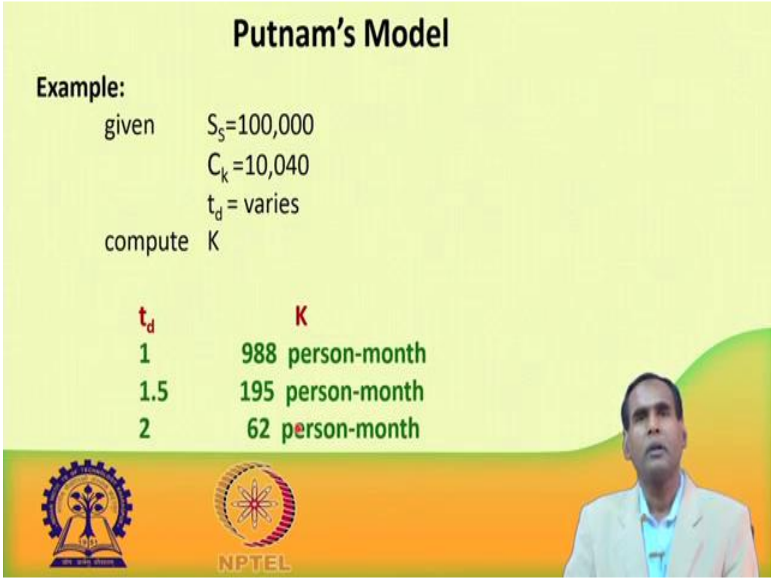
Let us take quickly a small example, if the estimated development time is 1 year then in order to develop the product in 6 months, so what will the consequence on the effort or the cost ok? So, original development time was 1 year, now the customer is coming and requesting you to develop the product in 6 months due to some emergency.

So, now let us see what will be it is effect on the effort and the cost? So, what I will do now I will put the values in the equation, what is the equation? Here you can see that originally t d 1 year; that means, 12 month ok. Then t d 2 is equal is equal to how much? New time what is what you can say that, so original time, so originally it was how much? 12 months or 1 year or 12 months and then it is subsequently reduced to how much? 6 months.

So, how much you will get? 12 by 6, so almost 2 to the power 4 that is 16 times. So, the effort or the cost will be increased by 16 times if the development time is just reduced to half, that is what I want to say. So, if the estimated developed time is 1 year and then in order to develop the product in 6 months; that means, just half you are reducing the time half. The total effort and hence the cost will increase how much? 2 to the power 4 that is

16 times. In other words, we can say that the relationship between the effort and the chronological delivery time is highly non-linear.

(Refer Slide Time: 23:10)



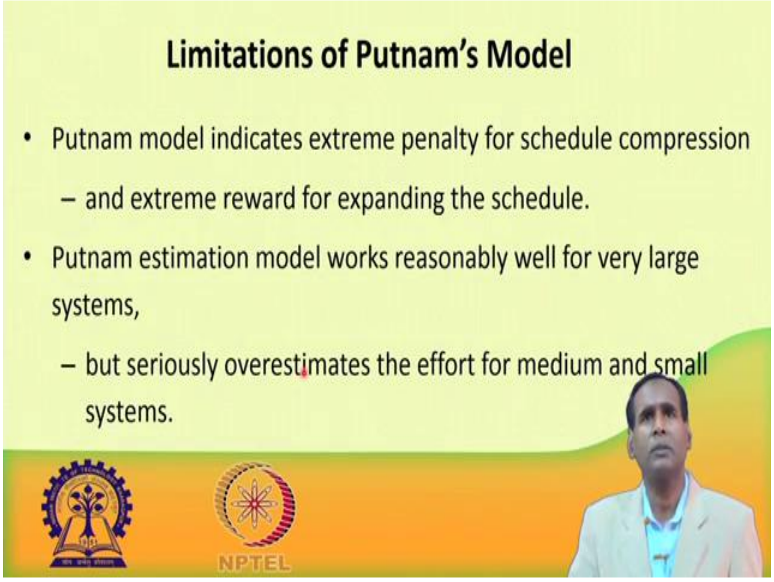
So, let us take a small example and apply Putnam's model for estimating this size effort etcetera. Given that the size S s ok, I hope S s is the size, so given that S s is equal to 100000 lines of code loc, or C k is equal to 10040 and t d varies it may be what 1 month 2 month or whatever that. So, t d is varies or maybe you can take yes t d you have taken in a months. So, now, you have to compute K for t d maybe 1 month, you are getting how much? Putting the values S C K be the values the since I have already shown you earlier that is the equations here.

So, S s, C k K and t d there, we want to compute the K which is the person years invested or effort in person months. So, K is equal to what? So, this is the basically effort k, so S s is given Ck is given t d is given. So, now, we can easily compute the value of K. So, for 1 month you see that for 1 month by putting this value of S s and this value of Ck.

That means, S s equal 100000 and C k is equal to what 10040, then the value of K is coming to be 988 person month, please put the values in the previous equation you will get this. Similarly, for td is equal to 1.5 months you will get this is the value of K; that means, this is the value of effort required, for td delivery time 2 months the value of effort is 62 person month. Given these given the these values for the parameters S s and

C k and t d is equal to 2. In this way you can compute the effort required by using Putnam's model.

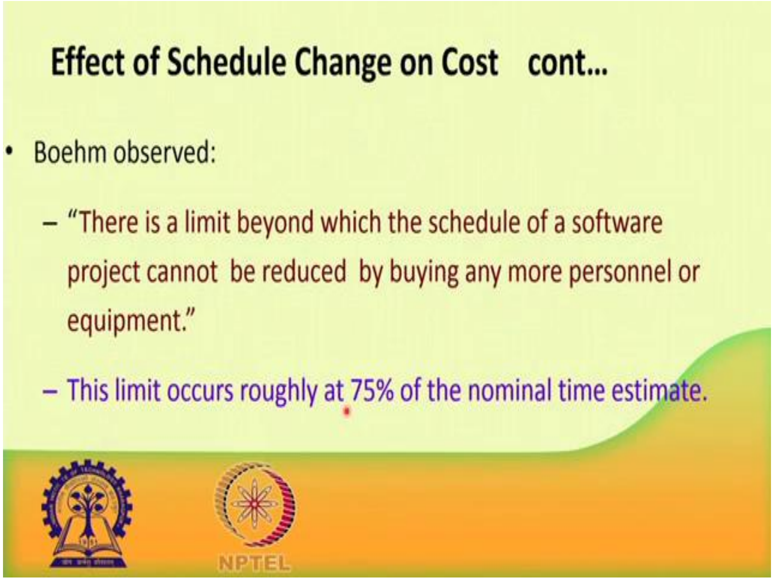
(Refer Slide Time: 25:12)



Now, let us quickly see what is the limitation of Putnam's model? Putnam's model indicates extreme penalty for schedule compression, that we have already seen that if this time or schedule is compressed a half. So, almost we have to what the effort will be increased by 16 times, the cost will be may be increased accordingly or what because the cost maybe increased 16 times.

And extreme reward for expanding the schedule, if you can go on expanding the schedule then your profit will be more you will get more benefit. Putnam's estimation model it works reasonably well for very large system. So, if you are system is a very large, your project is very large, then it will work perfectly the estimations will be more accurate. But seriously overestimate the effort for medium and small scale systems. So, if you are using just a small scale small scale system or a medium scale system, then on the estimates will not be accurate they will be overestimated.

(Refer Slide Time: 26:16)

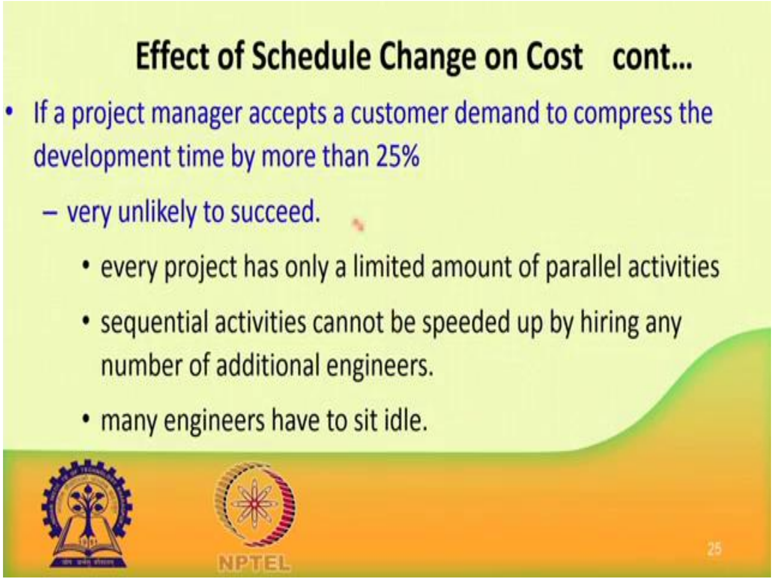


So, in order to what here you can see that there is a limit beyond which the schedule of your software project cannot be reduced by buying any more personnel or equipment, I have already told you in this graph that. So, this is the optimal time point of time and the optimal effort is here. If you are reducing up to say this point you can reduce theoretically and you see effort is increasing tremendously. But still if you will again go on reducing there will be some point and beyond it you are reducing the time you cannot complete your project you may not be successful.

So, what is that point, up to what point you can go on compressing the schedule and if you come further schedule that then you may not succeed. So, that point Boehm has studied many applications and he has concluded this. Boehm observed that there is a limit beyond which the schedule of a software project cannot be reduced by buying any more personnel or equipment.

Because some of the what because some of the activities are sequential they are not parallel, so how many what personnel you may hire. So, since these things activity has sequential, so they cannot complete the extra persons they will just sit idle. So, what is this limit? This limit occurs roughly at 75 percent of the nominal time estimate. So, this limit it will occur at 75 percent of the nominal time estimate beyond this if you will schedule, so your project may not succeed.

(Refer Slide Time: 27:59)

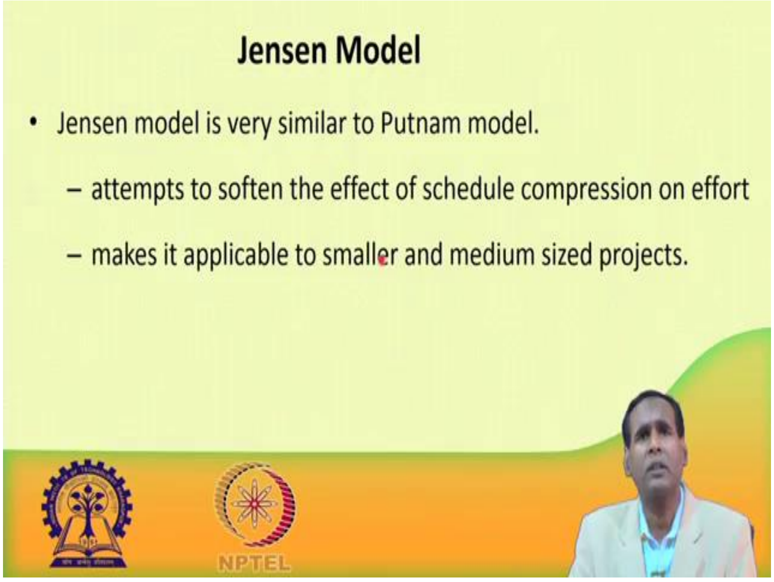


If a project manager accepts a customer demand to compress the development time by more than 25 percentages. See here I have already told you the limit is up to 75 percentage. So, there; that means, if a project manager accept a customer demand to compress the development time by more than 75 25 percent, because we have already seen that it is 75 percent the limit.

If a project manager he accepts a customer demand to compress the development time more than 25 percent, then it is very unlikely he will succeed you may not complete the project you will he may fail. Because every project has only a limited amount of parallel activities where you can employ more people most staff for those areas, but there are several sequential activities.

These sequential activities cannot be speeded up by hiring any number of additional engineers and many of the engineers they will have to sit idle. So, if will compress the development time by more than 25 percent, then certainly your project may fail it may not succeed you may not get success, because many engineers they will just sit idle because for them no work. Because many activities are sequential they are not parallel they are sequential, and any number of what the sequential activities cannot be speeded up by hiring any number of what personnel, so this engineers will sit idle.

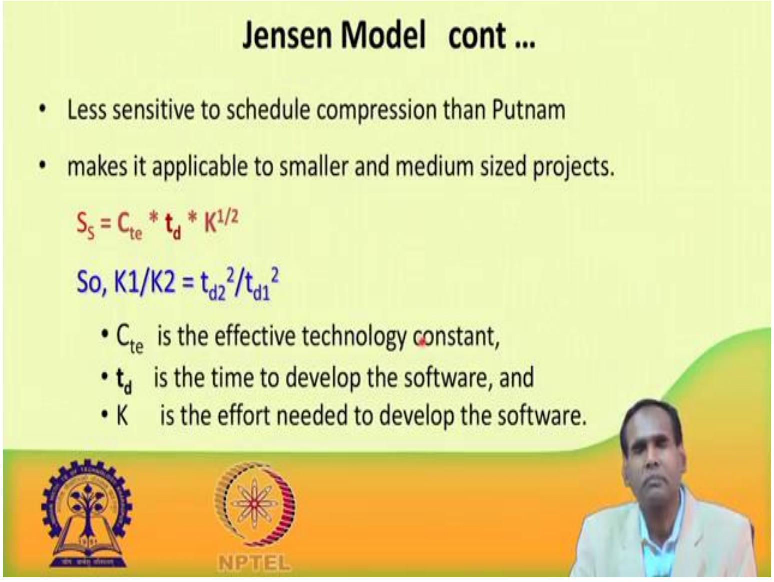
(Refer Slide Time: 29:30)



We have seen that one of the major drawback of what this Putnam model is that it does not work for what it does not work well among for the medium and small size systems. So, in order to overcome that Jensen came he has proposed a model. This is known as Jensen's model, Jensen's model is very similar to Putnam's model, he attempts to soften the effect of scheduled compression on effort.

Because we have seen in case of the Putnam's model the effect of scheduled compression on the effect on scheduled compression on effort is very hard it almost its 4 times and we have already seen it is of to the power 4. So, that is why when we are reducing half the schedule almost 16 times effort is increasing ok. This makes it applicable to smaller and medium sized project, since he has tried to soften the effect of schedule companies on effort this makes it applicable to be applied in smaller and medium sized projects.

(Refer Slide Time: 30:35)



So, it is a less sensitive to schedule compression than Putnam ok, it is very much less sensitive to schedule compression than Putnam. It makes it applicable to smaller and medium sized projects, according to Jensen or according to Jensen's model the size can be specified as

$$S_S = \mathbf{C_{te} * t_d * K^{1/2}}$$

C te into td into K to the power 1 by 2.

Here the terms of the user meaning S s is the size and C te is the it is the effective technology constant, t d is the time to develop the software and K is the effort needed to develop the software and these are the formulas is there. So, you can see this formula is very much suitable for the small and medium sized program problems or projects.

From this equation similarly you can find out the what for the two values; one is the original schedule another is the revised schedule you can find out the values of K1 and K 2 to the power them then where K represent the effort. Now, you what divide K 1 by K, 2 then you can see what is the effect of schedule compression on effort.

You can see that now it is of the order of the square previously it was of the order of to the power 4, now you can see this ratio is of the order of square. So,

$$K1/K2 = t_{d2}^2/t_{d1}^2$$

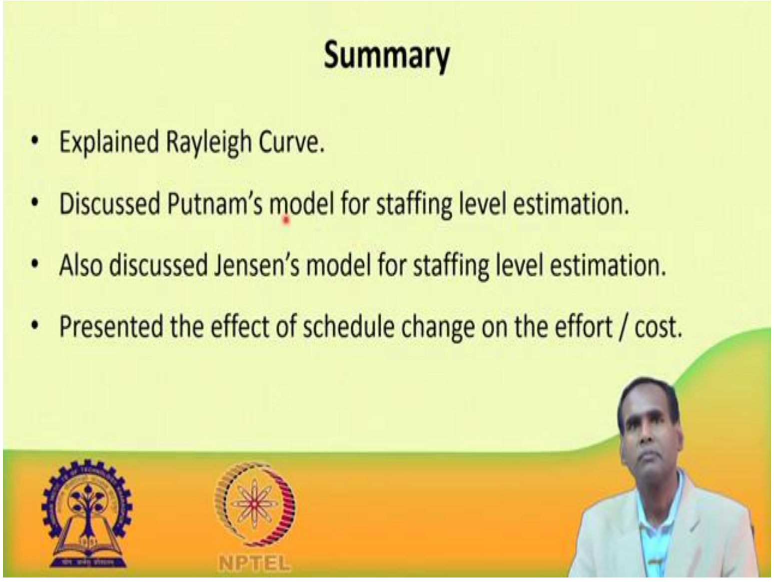
K1 by K2 is equal to t d 2 by t d 1 whole square. So, here the effort it will not increase of the order output to the power of 4 rather it will increase to the of the order of to the power 2.

(Refer Slide Time: 32:09)

So, now if you will take the same problem that we have applied for Putnam's model you can see this will work like this. If the estimated development time is 1 year then in order to develop the product in half that is 6 months. Now, the total effort and hence the total increase the total cost increase how much time 12 by 6 is 2.

So, 2 square, 2 square means 4. So, now, the effort and cost will increase 4 times whereas, in Putnam's model it was increasing by 16 times. So, this is much less in comparison to the Putnam's model, so these are the advantages of Jensen's model over Putnam's model. Now, first the if the schedule is reduced the time is reduced the effort is not increased in a if the effort is not increased rapidly. As well as these model can be easily applied to small and medium scale, small and medium scale, or small and medium size small and medium sized projects.

(Refer Slide Time: 33:12)



We have seen in this class the details of Rayleigh curve, how Rayleigh curve can be applied to what software industries. We have discussed the Putnam's model for staffing level information, estimation. We have also discussed Jensen's model for staffing level in estimation, we have presented the effect of schedule change on the effort of cost on effort and hence the cost using Putnam's model and Jensen's model

(Refer Slide Time: 33:40)

These are the references.

Thank you very much.