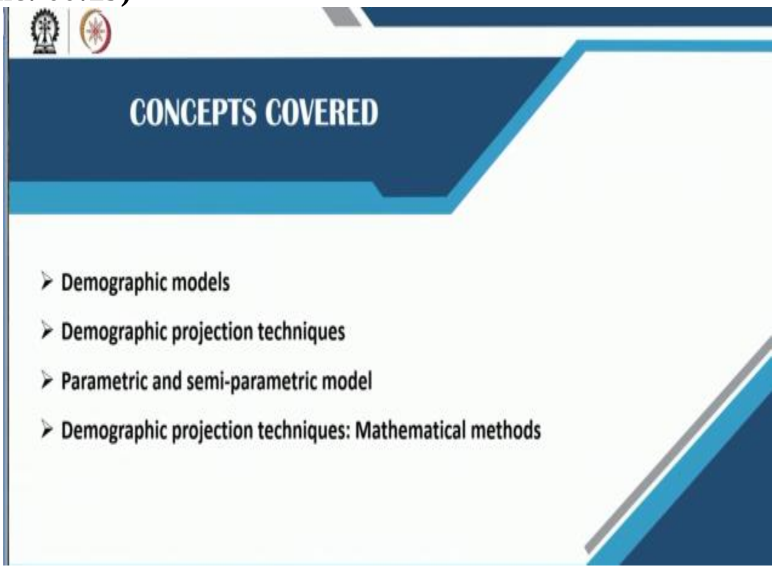
Urban Landuse and Transportation Planning
Prof. Debapartim Pandit
Department of Architecture and Regional Planning
Indian Institute of Technology - Kharagpur

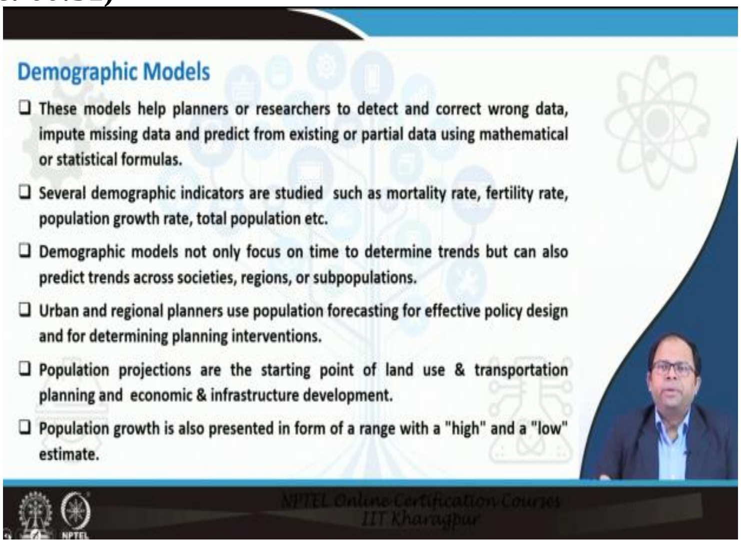
Lecture - 17
Demographic Models - 1

(Refer Slide Time: 00:29)



Welcome to lecture 17. In this lecture, Demographic models will be covered. The topic will be covered in two lectures. The different concepts that will be covered in this lecture are Demographic models, Demographic projection techniques, Parametric and Semi-parametric models, Demographic projection techniques-mathematical methods. Mathematical methods will be covered in this lecture in detail and the other methods will be covered in the subsequent lecture.

(Refer Slide Time: 00:51)



Demographic models refers to the models which tries to predict the different values of demographic indicators in a future time period. These models help planners or researchers to detect and correct wrong data, impute missing data and predict from existing or partial data using mathematical or statistical formulas. When prediction is done for future years using these models and if there is some wrong or missing data, such gaps can be filled. Prediction can be done for the entire population from the existing partial data, ie, if there is only a partial data of a population or some population subgroups.

In these models, several indicators are studied such as mortality rate, fertility rate, population growth rate, total population and so on. Demographic models not only covers the total population but also covers all the different components such as the mortality rate of the population, the fertility rate, its rate of change and so on.

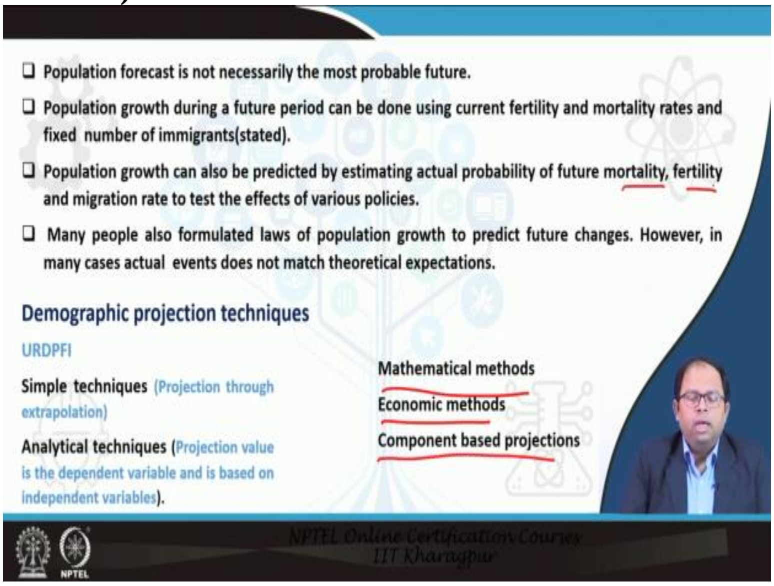
These models not only focus on time to determine trends, but can also predict trends across societies, regions, or sub populations, for instance, such models not only predicts that there is certain change in fertility of a particular population or age group such as the female population for a particular area for a time period, but also predicts the trend across several geographic areas.

Urban and regional planners use population forecasting for effective policy design and for determining planning interventions and hence predicting population becomes the most important task of demographic modeling.

Population projections are the starting point for land use and transportation planning, which is the focus of this course and also for economic and infrastructure development. So, only if projection is done for the population and the different demographic indicators, like age and sex groups, planning for a future time period is possible. When land use transportation plan is done for an urban area, it is important to firstly ascertain what is going to be the future population, however, the future population has to be determined not only for the total population, but also for each individual population subgroup. The population subgroup, as discussed in the previous lecture, could be a 2 dimensional one, 3 dimensional one and so on. It is important to consider sub categories, for example, a population group having a certain income category, belonging to a particular gender and a particular age group can be considered as a small population subgroup.

Population growth is presented in form of a range with high and a low estimate that means, since the population is projected and there are so many unknowns, it is better to have a range of estimate with a high and a low value instead of having a single value.

(Refer Slide Time: 04:26)



Population forecast is not necessarily the most probable future. In most cases, when population forecast is discussed upon, assumption is done for the current fertility and the mortality rates of the population along with fixed number of immigrants coming in to the particular urban area, which is exogenously determined based on certain estimates or previous historical data and so on. So, when prediction for a future time period is done using older or current fertility rates and mortality rates, forecasting may not get perfectly done, however, it could be used for future determination of policies, and for the future activities of that particular area. It is important to note that it is not exactly determining the probability of a likely population figure for a particular time period.

Population growth can also be predicted by estimating actual probability of future mortality, fertility and migration rate to test the effects of various policies. So, at times, detailed studies to determine the future mortality rates, future fertility rates based on other indicators or other relationships that these indicators have, such as with age, income levels etc. can be conducted. Then the final population prediction can be done from the prediction of future values for this particular indicators. Thus, a much better estimate could be obtained however it is difficult.

Many people have formulated laws of population growth to predict future changes, for instance, they find that mortality is related with age such that when the age increases, there is a higher chance of mortality. Such laws usually matches to some extent with the population of certain country, but it can be found that when that same equation is tried to fit for another country, it doesn't represent properly or it may happen that it may not fit for all age groups. In many cases, it does not match the theoretical expectations. Now, as the different kind of demographic models is known, the next step is to develop these projection techniques for different demographic indicators.

Demographic projection techniques

URDPFI

URDPFI has listed 2 techniques, for projecting population in particular.

Simple techniques - These are techniques where projection is done through extrapolation

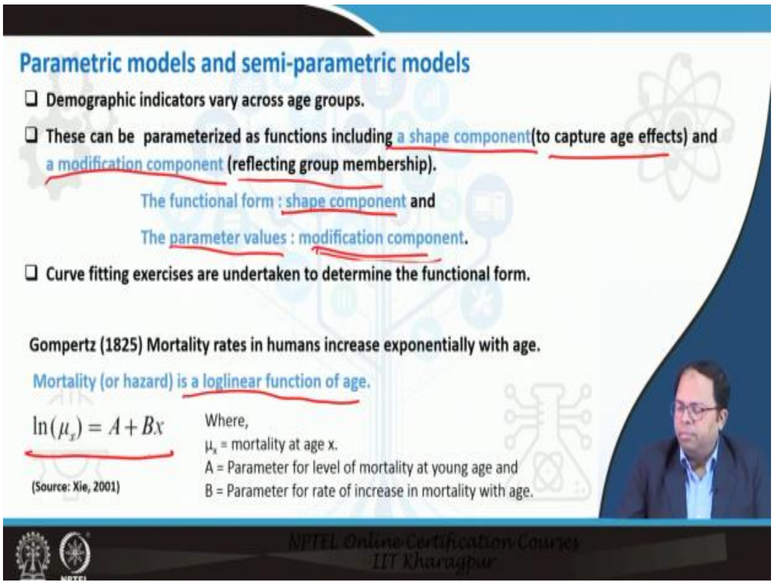
Analytical techniques – There are analytical techniques where projection values is the dependent variable and is based on some independent variables that is based on certain relationship

Different projection techniques, particularly population projection techniques which are listed from other literature sources are listed as

a. Mathematical methods
b. Economic methods
c. Component based projection methods

These are the 3 primary ways we can project population but along population projection, and there are so many other parameters like mortality, fertility which needs to be projected.

(Refer Slide Time: 08:07)



Parametric models or semi-parametric models

As it is known that demographic indicators vary across different age groups. Parametric models can be used to determine how demographic indicators such as mortality, fertility varies between the different age groups. A relationship can be formed using certain few parameters like age, gender or income group etc. such that these parameters are related to fertility or mortality. Also, it is better to have the number of parameters as less as possible.

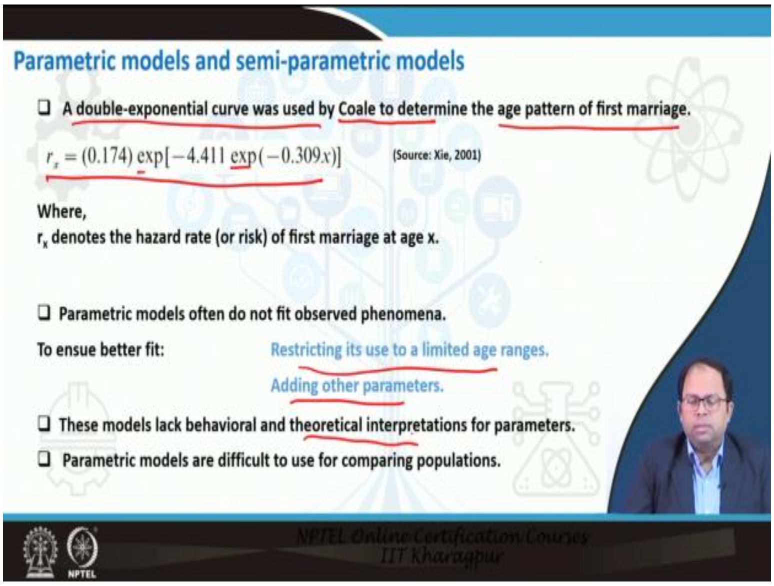
The function which relates these with the actual demographic indicator is called a *shape component*. This component captures the effect of increase or decrease in a parameter, for example, this function predicts how the different age groups are related with a particular demographic indicator. Also, there is a *modification component* which reflects the group's membership, ie, as the population belongs to a particular age group, there has to be certain considerations in the equation. So, the functional form of the equation relates to the shape component and shows how growth happens and the parameter values are related with the modification component. Now, this could be done with the curve fitting exercise with the existing data. For the existing data, we try to fit a particular curve; it could be an exponential curve or a linear relationship or any other function which shows the best relationship and fit.

Gompertz in 1825, said that mortality rates in humans increase exponentially with age and a relationship is framed as:

$$\ln(\mu_x) = A + Bx$$

When a parametric model is made, it can be said that mortality or hazard is a log linear function of age. Here, μ_x is the mortality at age x, A is the parameter for level of mortality at young age and B is the parameter for rate of increase in mortality with age. And here, the log value of μ_x which is mortality at age x is equated to (A + Bx). The parameter values can be modified as per the age group. For instance, mortality is less for young people and B value is considered changing for particular age group and A value could be fixed.

(Refer Slide Time: 11:20)



Similarly, for determining the age pattern of first marriage, Coale suggested a double exponential curve function, and is given as:

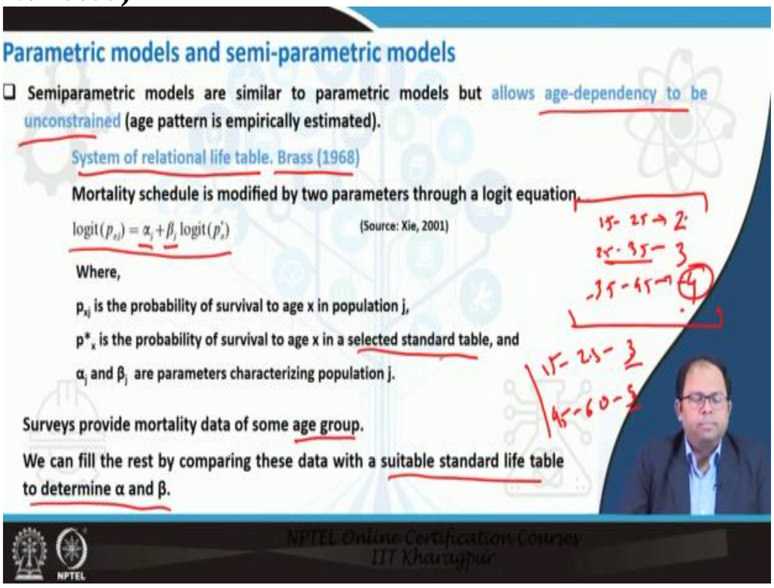
$$r_x = (0.174)\exp[-4.411\exp(-0.309x)]$$

Here, 2 exponentials are taken. $\mathbf{r_x}$ denotes the hazard rate or risk of first marriage at age x. So, the likely chance of a girl marrying at a particular age or the risk of a girl marrying at a particular age can be determined.

This is a parametric function, trying to find that at what age a girl is likely to marry, because that will lead to maybe childbirth and so on. This is how we can estimate the different indicators and how it is related with age groups and it can be observed that when a population transitions from one age group or a group of persons transitions from age group to another, how the fertility rates or how the nuptial rates will change. Both of these equations, as observed, may or may not fit to a particular population or it may not fit with different age groups. So, to ensure better fit, the use of these equations can be restricted to certain limited number of ages or some other parameters can be added to the equation to make it more flexible.

These models also lack behavioral and theoretical interpretations for different parameters because the best curve fit is what is found to say that this is how the population fertility rate will change or likelihood of a person marrying at a certain age will change, but actually, there are no theoretical interpretations explaining the relationship. Parameter models are also difficult when 2 populations are compared because the parameter values A and B values will change as clear from the previous equation and that will make it difficult to compare between the 2 populations.

(Refer Slide Time: 13:39)



In semi-parametric models, such issues are taken care of to a certain extent. Semi-parametric models are mostly similar to parametric models, but this allows age dependency to be unconstrained, that is, age pattern is empirically estimated, that means, changes in values of those parameters are estimated from some empirical data. Brass, in 1968, created relational life tables where, for instance, for different age groups, he has given certain mortality values; for age group 15 to 25, there is a certain value, for age group 25 to 35, there is a certain value and so on. So, corresponding mortality rate will increase such as 2, 3 or 4 and so on. Mortality schedule is modified by 2 parameters through a logit equation as given by Brass.

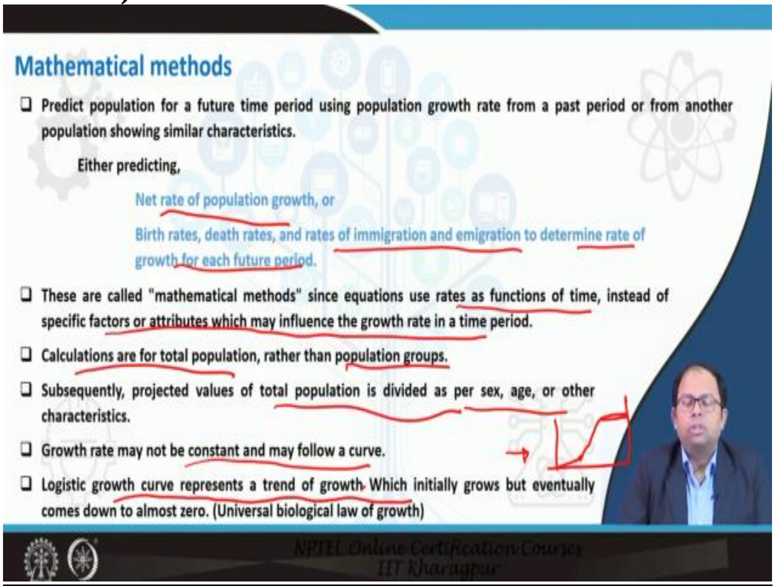
$$\text{logit}(p_{xj}) = \alpha_j + \beta_j \text{ logit}(p^*_x)$$

So, here from the equation it can be observed that there are 2 parameters such as alpha and beta. These are the parameters characterizing this particular population group or subgroup or whichever age group they are in and p_{xj} is the probability of survival to age x in population group j and p^*_x is the probability of survival of age x in a selected standard table. So, this is where this rate is being considered.

It means that, if we survey a particular population or rather a sample as the entire population cannot be surveyed and we have got the mortality values for 2 age groups, such as 3 for the 15 to 25 age group and 5 for the 45 to 60 age group, these values can be compared within themselves, and also with the value from the standard table.

Alpha and beta values gets adjusted based on this empirical data and then it is computed for all the other age groups which are missing from the data. Surveys provide mortality data of some age group and the rest can be filled by comparing this data with a suitable standard life table to determine alpha and beta values which is already given. So, comparison is done and based on the relationship, we values for the age groups which are not provided in the data set can be filled. Thus, the semi-parametric model is adjustable compared to the previous models.

(Refer Slide Time: 16:43)



Population projection techniques - Mathematical models

There are broadly three groups of population projection techniques; mathematical methods, economic methods and the component methods. In mathematical methods, population is predicted for a future time period using population growth rate from a past period or from another population group/area showing similar characteristics.

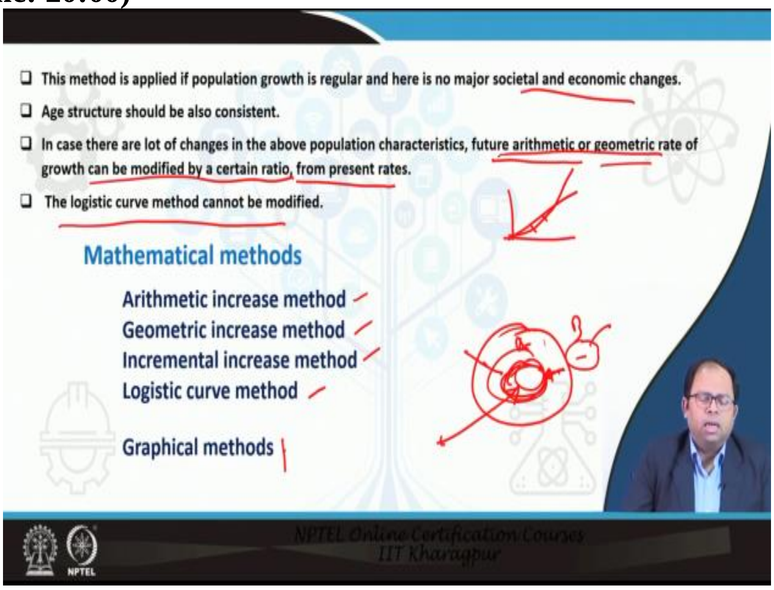
Either the *net rate of population growth* or *birth rates, death rates and rates of immigration and emigration to determine the net rate of growth* for each future period could be predicted.

These are called mathematical models, because the equations use rates as functions of time, instead of specific factors or attributes which may influence the growth rate in a time period. So, unlike the parametric models, these consider only the previous population growth rates and not the relationship with other components.

Calculations are also done for the total population, and the different population age groups are not considered. Once the total population is determined after projection it is divided as per sex, age or other characteristics to get the values for each population component.

Growth rate may not be constant and may follow a curve. Thus, instead of considering the same growth rate as previous years, it can be considered that, the growth rate follows a particular relationship or a particular curve. The logistic growth rate represents one such trend of growth or growth curve and is very popular. Logistic curve represents a pattern where initially the growth rate is low, and then there is an increase which is finally followed by stabilization of the rate to almost become 0 showing no further increase. This follows the universal biological law of growth. Many things in nature actually show this kind of a growth pattern. Hence, in many cities or many urban areas where there is normal growth and there is no special event occurrence, logistic growth curve can be followed and it gives very accurate results.

(Refer Slide Time: 20:00)

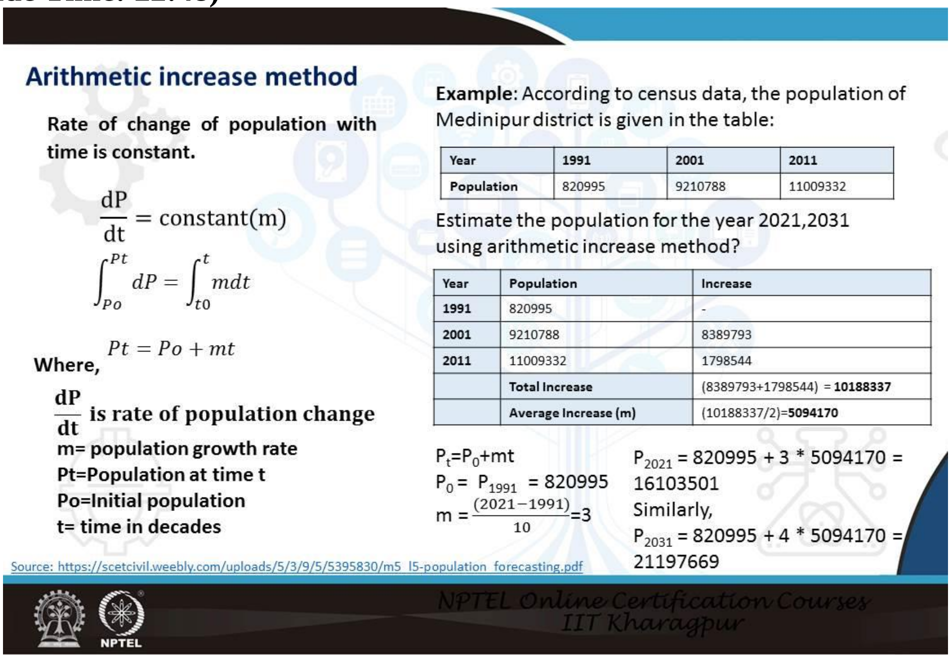


This method is applied if population growth is regular and there is no major societal and economic change. Age structure (that is for logistical curves) should also be consistent, i.e., the age structure does not undergo radical changes.

In case there are a lot of changes expected in the above population characteristics in the years to come, future arithmetic or geometric rate of growth can be modified by a certain ratio from present rates. Thus, instead of using logistic curve, a linear increase rate or a geometric increase rate can be used. The geometric rate that is estimated for previous years can be increased by a certain ratio for the future years. Logistic curve method does not allow this modification. There are several sub methods, for the mathematical methods such as Arithmetic increase method, Geometric increase method, Incremental increase method, Logistic curve method and Graphical methods.

Graphical methods are not exactly arithmetic methods. It is done manually using graph paper. It can be said that, it is similar to these kind of methods, where the previous growth is just extended. The growth in a particular urban area can also be modeled by considering the population from other urban areas. Such modelling actually assumes that the urban area under consideration (A) is following similar economic and social growth pattern as another urban area (B). This is done when data is not available for A whereas, there is data for B. Similarly, the growth rate for a smaller area can be adopted from that of a larger urban area when there is data for the larger area. So, this could be called a mathematical method, however this could also be called an economical method because, an assumption is made such that the same economic growth pattern of the larger area can be adopted for the smaller area which is contained within the larger area.

(Refer Slide Time: 22:48)



Arithmetic increase method

In arithmetic increase method, the rate of change of population with time is constant.

$$\frac{dP}{dt} = constant(m)$$

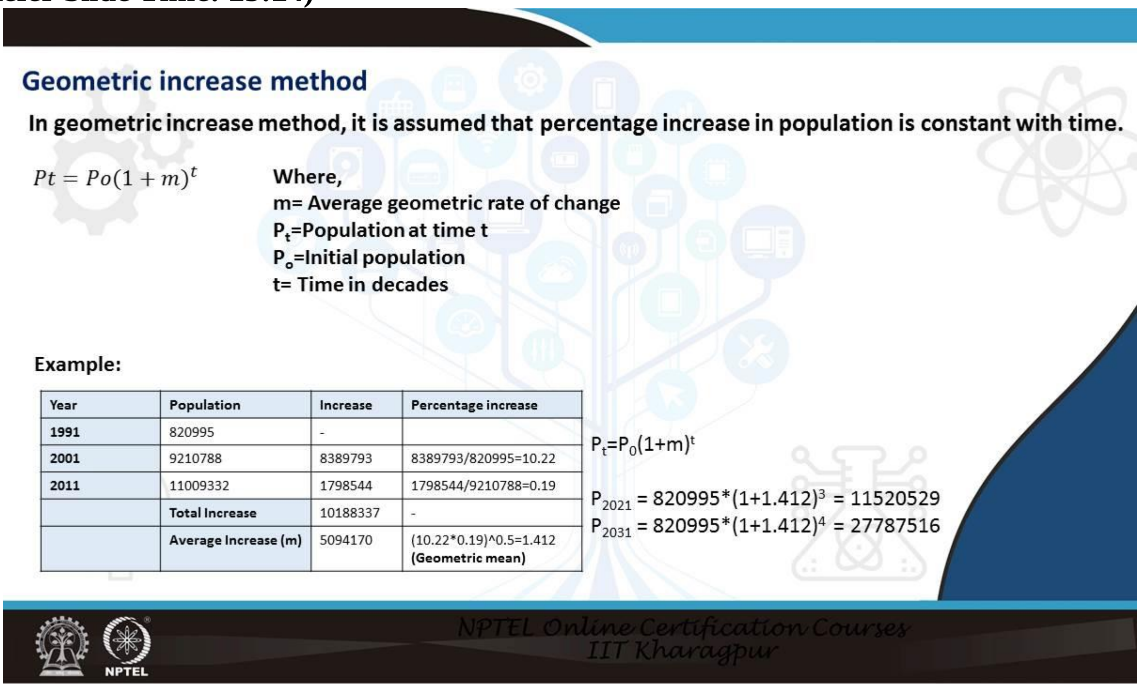
$$\int_{Po}^{Pt} \mathrm{dP} = \int_{t0}^{t} \mathrm{mdt}$$

$$\mathrm{Pt} = \mathrm{Po} + \mathrm{mt}$$

So, dP/dt is the rate of population change and m is the population growth rate, P_t is the population at time t, P_0 is the initial population, and t is time in decades. Time could be even in years and it doesn't matter much.

In the example, the population data for Medinipur district for 1991, 2001 and 2011 is given. This data can be used to estimate the population for the year 2021 or 2031. First, increase in the population from 1991 to 2001is estimated. In the first decade, it is observed that, there is a huge rise in population from 820000 to 83 lakhs. The two decadal increase values are considered to calculate the average increase per year. Thus, this is a very simple way of projecting the population and change in growth rate is not considered. Some problems may occur since we are not considering the change in the growth rate. It can be observed from the example that, there is a huge increase during the period from 1991 to 2001 and then the increase had slowed but this is ignored when the population for the future year is estimated.

(Refer Slide Time: 25:24)



Geometric increase method

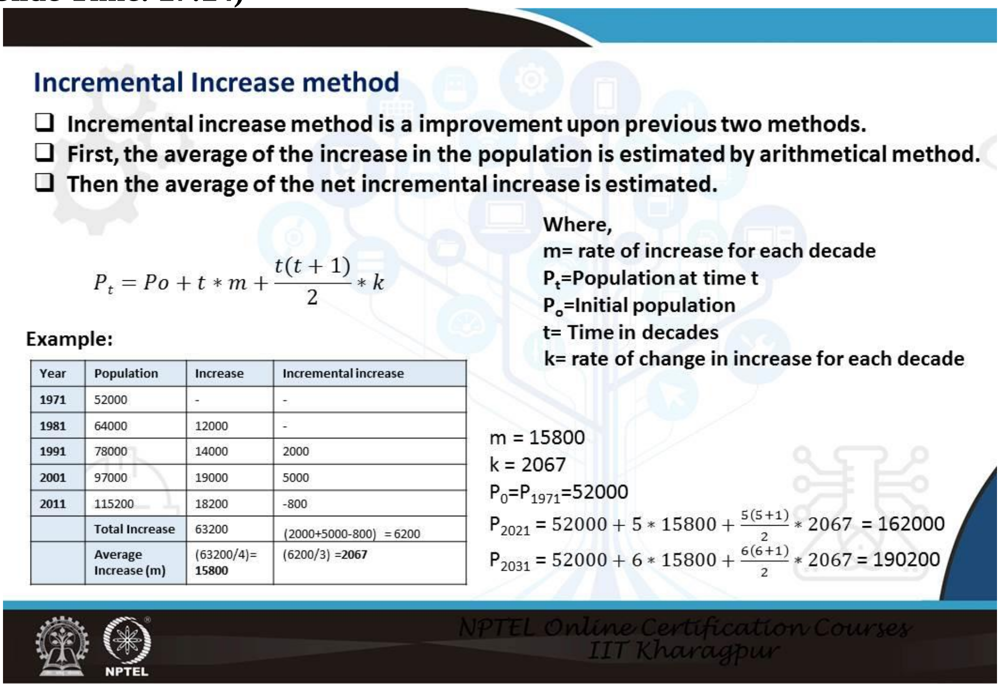
In geometric increase method it is assumed that percentage increase in population is constant with time. So, here the formula is

$$Pt = Po(1 + m)^t$$

Where, m is the average geometric rate of change, P_t is the population at time t, P_o is the initial population, and t is the time in decades.

While explaining this, the same population figures can be considered as in the previous one. Here, average increase is estimated through geometric mean which is obtained by taking square root of product of both of the percentage increase values. If 2021 population data was also included, then a power of one third would have been taken to calculate the geometric mean.

(Refer Slide Time: 27:24)



Incremental increase method

This method is an improvement upon the previous 2 methods. However, there are some issues with this method. Here, the increase of population is divided into 2 parts; the first is the average of the increase in the population and is estimated by arithmetic method.

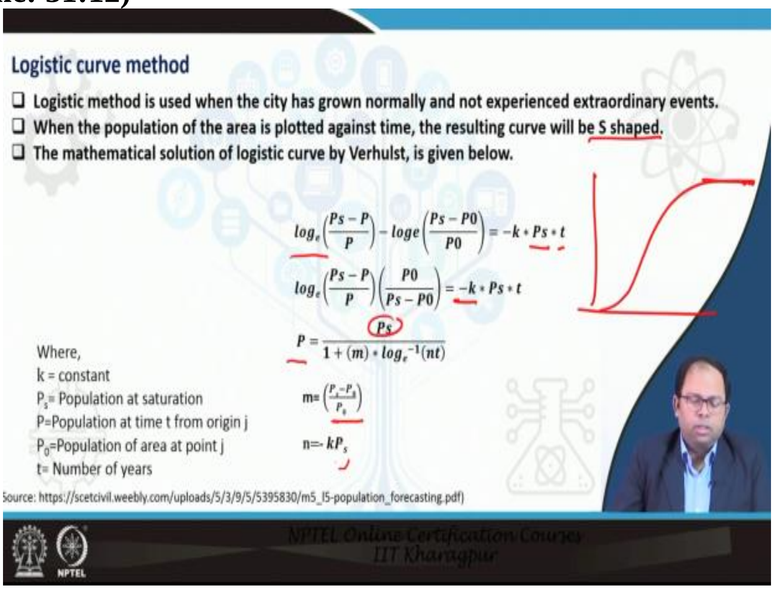
$$P_t = Po + t * m + \frac{t(t + 1)}{2} * k$$

Next,the incremental increase is also calculated along with the total increase as explained in the example problem. The arithmetic part and the incremental increase part is calculated as shown in the example to obtain the population figure for the year 2021 and 2031.

(Refer Slide Time: 29:49)

When the projections are plotted in a graph, it can be observed that, the arithmetic projection which results in a linear plot is simple and indicates that, increase happens with the same rate. In geometric projection, that rate of increase is low during the initial years and then suddenly it starts increasing. This could be applied in cases when there is transformation in the society. It can be used to represent a scenario where a lot of industrialization is happening resulting in the increase in population growth. But in most cases, it is very difficult to use this kind of projection, because it will result in overestimation. Next, we have the incremental increase method which seems in between the plot of both arithmetic progression and geometric progression. It can be observed that the rate of increase is more appropriate.

(Refer Slide Time: 31:12)



Logistic curve method

This is used when the city is grown growing normally and not experiencing any extraordinary events. So, when the population is plotted against time, the resulting curve will be S shaped as discussed earlier, and is shown in the slide. The mathematical solution to this logistic curve is given by Verhulst as below:

$$\boldsymbol{log_e\left(\frac{Ps-P}{P}\right) - loge\left(\frac{Ps-P_0}{P_0}\right) = -k * Ps * t}$$

$$\boldsymbol{log_e\left(\frac{P_s-P}{P}\right)\left(\frac{P_0}{P_s-P_0}\right) = -k * Ps * t}$$

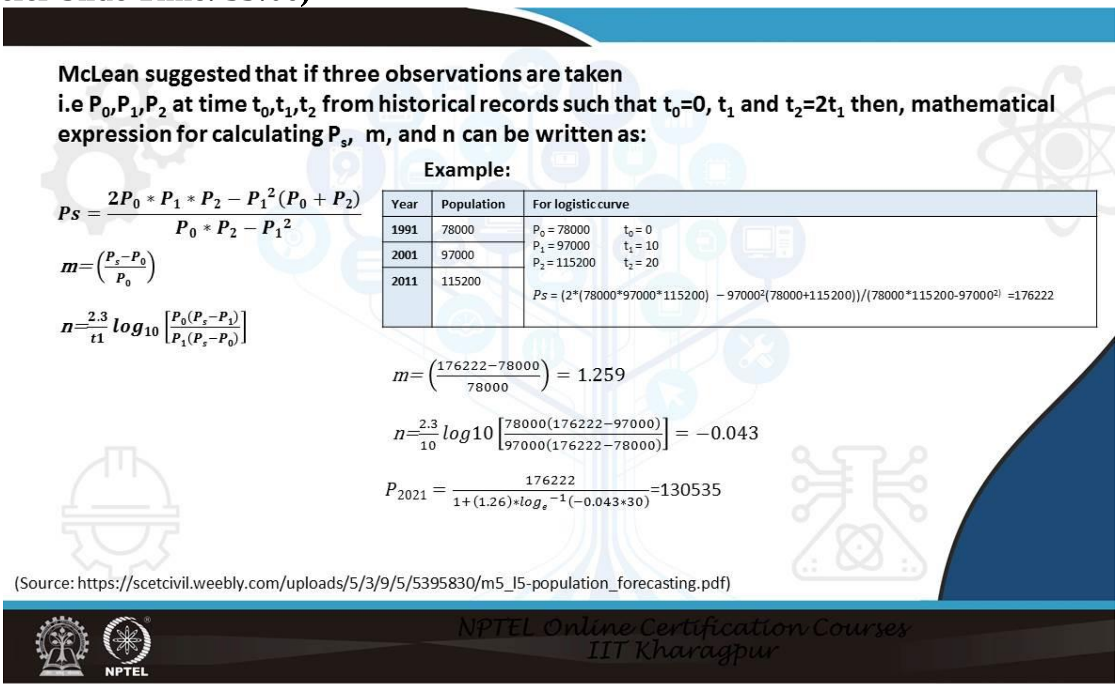
$$\boldsymbol{P = \frac{Ps}{1+(m)*log_e{}^{-1}(nt)}}$$

$$\mathrm{m} = \left(\frac{P_s - P_0}{P_0}\right)$$

$$\mathrm{n} = -kP_s$$

Where k is a constant, P_s is the population at saturation,P is the population at time t from origin j, P_0 is the population of area at point j and t is the number of years

(Refer Slide Time: 33:06)



This particular equation, was further worked upon and McLean who suggested 3 observations based on which, one can predict the population using the logistic curve method. If P_0,P_1,P_2 at time t_0,t_1,t_2 are available from historical records such that t_0=0 where t_0 is the 0^{th} year, ie the first time period and $t_2=2t_1$, ie, there is equal gaps between two time periods, then, mathematical expression for calculating P_s, m, and n can be written as:

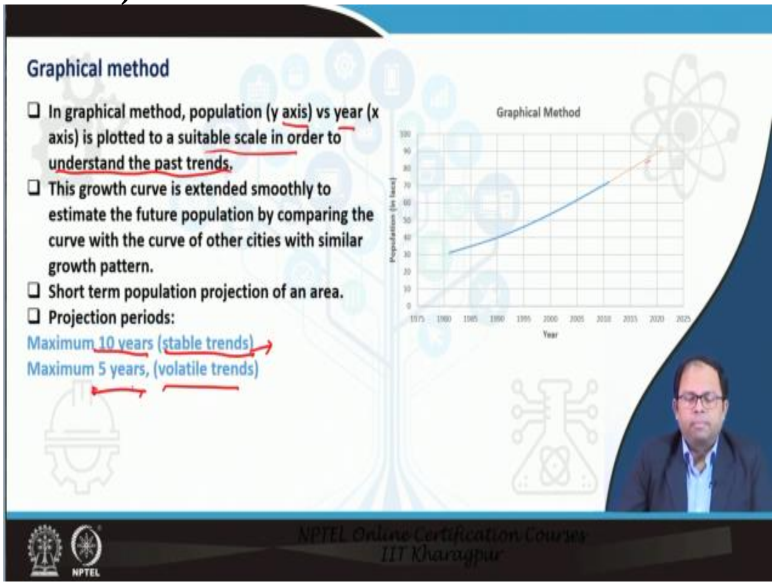
$$Ps = \frac{2P_0 * P_1 * P_2 - P_1^2(P_0 + P_2)}{P_0 * P_2 - P_1^2}$$

$$m = \left(\frac{P_s - P_0}{P_0}\right)$$

$$n = \frac{2.3}{t1} log_{10}\left[\frac{P_0(P_s - P_1)}{P_1(P_s - P_0)}\right]$$

So, using the above mathematical expression Ps, m value and n value can be calculated as done in the example. Similarly, substituting the m value and n value in the equation for P, P for the year 2021 can be estimated.

(Refer Slide Time: 35:41)



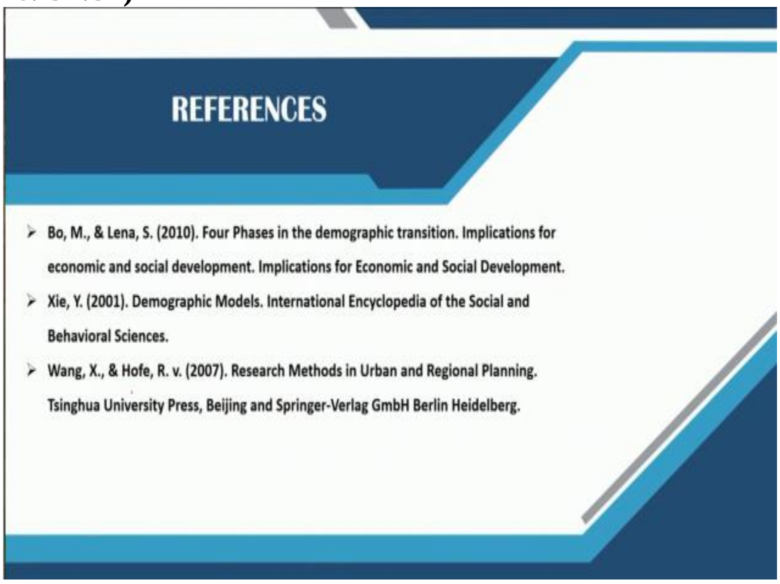
Graphical method

All the methods other than the graphical method are based on certain assumptions. For example, in the Arithmetic increase method, it is assumed that, the growth rate will remain same. In geometric growth rate method, population growth is assumed to accelerate and then for the incremental method and the logistic curve method, we are assuming that there is going to be a saturation value.

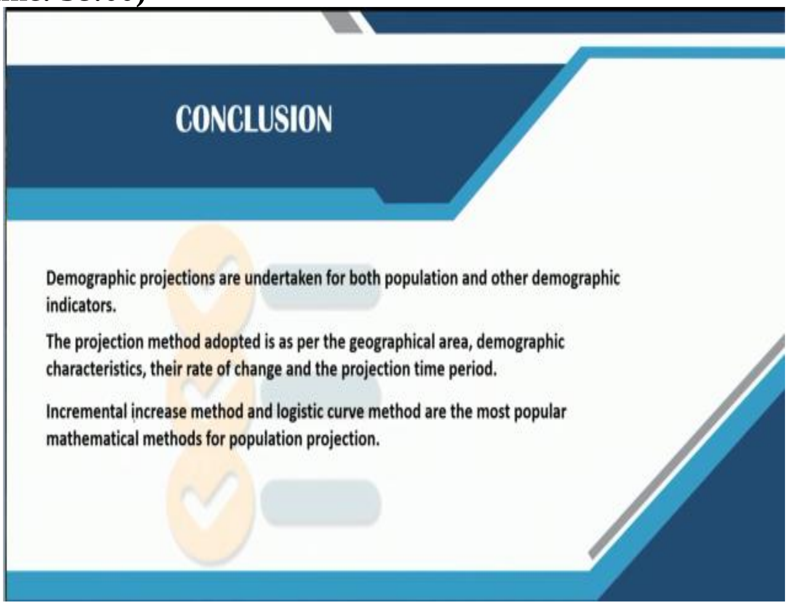
A graphical method can also be followed where the existing population data for different time periods is plotted first. Here, population is plotted in the y axis and year or the time is plotted in the x axis using a suitable scale. This actually helps us to understand the past trend in a visual way.

In the example, it can be observed that the plot is nonlinear and is a curve. This growth graph can be extended smoothly to estimate the future population by comparing the curve with the curve of other cities with similar growth pattern. Usually this could be done for short term population projection and not for very long term projection. This is because, without any particular methodology, a lot of assumptions gets considered based on the person who is actually doing the extension. So, this could be done for projection periods of maximum 10 years if the different demographic trends are stable. However, for volatile trends i.e., when the population is going through a lot of changes, then, this kind of method can be adopted only for a maximum of 5 years.

(Refer Slide Time: 37:52)



(Refer Slide Time: 38:00)



Conclusion

The demographic projections are undertaken for both population and other demographic indicators. The projection method adopted is as per the geographical area, demographic characteristics, their rate of change and the projection time period. Incremental increase method and logistic curve method are the most popular mathematical methods for population projection.

Thank you.