

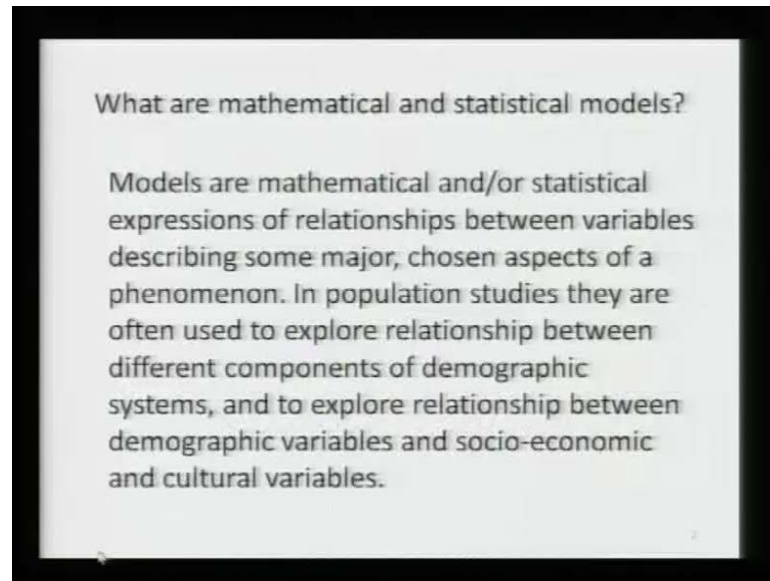
Population and Society
Prof. A. K. Sharma
Department of Humanities and Social Sciences
Indian Institute of Technology, Kanpur

Lecture No. # 12
Demographic Models

Well, this is our second lecture on demographic models. Previous lecture I said that there are some questions in population studies in particular in sociology in general which can better be answered by building certain mathematical statistical models and which cannot be answered simply on the basis of primary or secondary data. One example I gave was the example of estimating how many people have ever walked on this planet earth. Now this is a question which you cannot answer on the basis of survey. And I said that roughly it can be estimated that around hundred twenty or hundred thirty billion people have ever been born on this planet earth of which nearly seven billion are present.

Similarly, there was a question on, that if there are cultural variations in stopping rules that in different cultures you find people want different number of sons and daughters. Then can we calculate what will be the sex ratio of population under different stopping rules. That **that** is again a question which cannot be answered on the basis of survey, but if you know little bit of mathematics we can answer that question easily. Today we will talk more about demographic models.

(Refer Slide Time: 01:44)



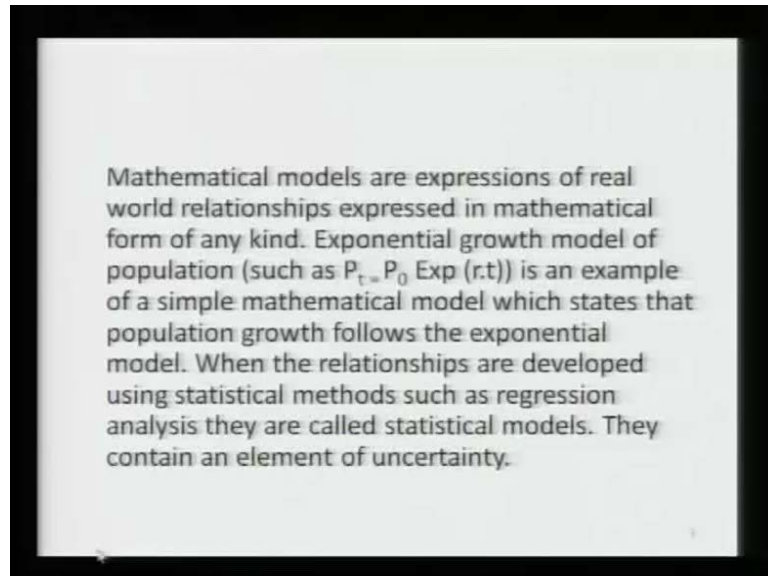
To recapitulate what are mathematical and statistical models mathematical and statistical models are expressions of relationships between variables describing some major, chosen aspects of a phenomenon. In population studies they are often used to explore relationship between different components of demographic systems, and to explore relationship between demographic variables and socio-economic and cultural variables.

You know the way I define model I said that model is simple representation of reality now you cannot represent reality in totality. So, you choose a specific aspect of reality representation of that specific aspect of reality with the help of language diagram mathematical equation statistical equations is what modeling means so, mathematical and statistical models are mathematical equations or statistical equations or distributions in terms of which it chosen is specific aspect of reality is be represented.

If you choose to represent only one simple aspect of reality you have a simple model, and if you want to represent various aspects of reality then you have a complex model, and more you move from hypothetical simple case to a more realistic representation of reality then you are moving from simple to complex model. So, in case of demography for example, if you write in a mathematical form a relationship between birth and death rate suppose it can be shown that birth rate is a function of death rate then you have mathematical model. If you write a mathematical relationship between birth rate and socio-economic variables then you have a an economic demographic model, and this can

be done by using mathematical equations or statistical distributions in terms of probabilities.

(Refer Slide Time: 04:00)



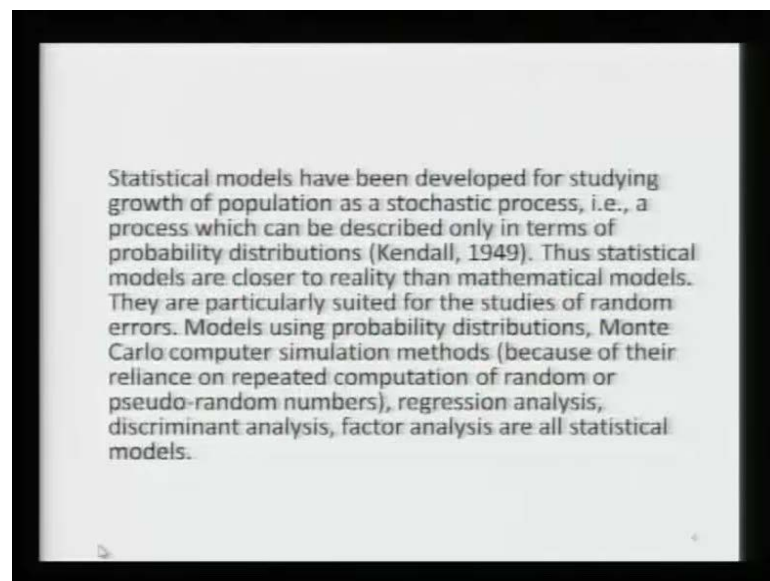
Mathematical models are expressions of real world relationships expressed in mathematical form of any kind exponential growth model of population which is usually written as $P_t = P_0 e^{r.t}$ where P_t is population of a country or a region at time t P_0 is population in the base year r is the rate of growth and t is the time between 0 and t . This is an example of a simple mathematical model which states that population growth follows the exponential model there are so many mathematical they will be actually innumerable curves which can be used for depicting growth of a population simple case is we are interested in simple algebraic equations with help of this workable feasible useful estimates of population can be made.

And in case of discrete distributions geometric in case of continuous exponential are two simple models quite often population growth is represented with the help of simple exponential model $p_t = P_0 e^{r.t}$ this is a mathematical model it is mathematical it is not statistical with the help of this model we cannot tell like if you know the population of India in 2001, and you want to estimate population of India in 2010 you will have one simple single estimate of that, P_0 will be population of the base year means 2001 t will be the time lapse between first April 2001 and the date for which you want the estimate r is the rate of growth, which can be obtained from previous

decadal growth rates or from sample registration scheme or you take a more thoughtful decision a more considered view of growth rate of the country during this time you may say that the decadal growth rate of the previous decade has been reduced by 30 percent or 40 percent or that you can say that you will take average of decadal growth rate and natural growth rate obtain from sample registration scheme this will be a point estimate.

But we know that nobody can predict exactly what is the size of India's population today, can we describe size of population of India today in terms of probabilities like can I say that the probability that the size of population of India is less than 1 billion is 0 between 1 billion and 1.1 billion it is 0.2 between 1.1 billion and 1.3 billion it is 0.7 like that, if I do so, and describe what is likely to be the size of India's population today in terms of probabilities and ranges of size then I have a statistical model that is the difference between mathematical and statistical model.

(Refer Slide Time: 07:16)



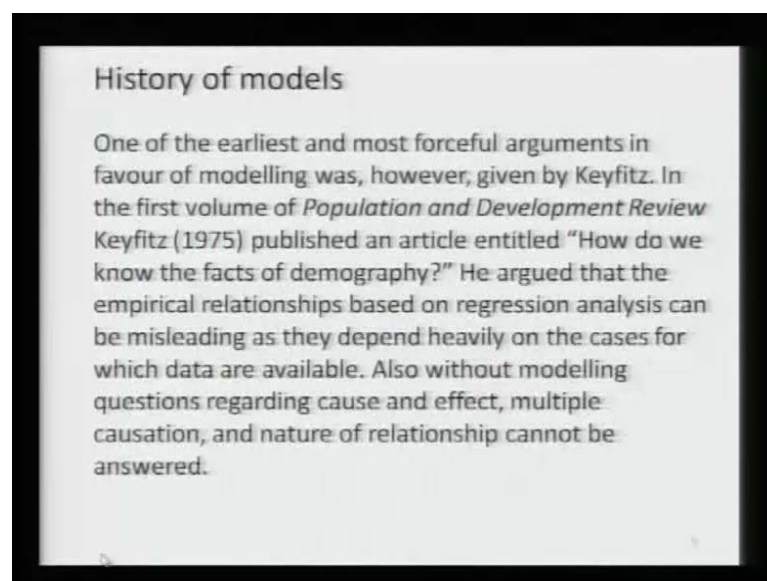
Mathematical models are deterministic and statistical models can take care of stochastic random fluctuations statistical models have been developed for studying growth of population as a stochastic process wherever we have stochastic process with time dependent process uncertain events that is a process which can be described only in terms of probability distributions. Thus statistical models are closer to reality I tell in my class always that social reality can never be described in realistic terms you there is no

social reality at least I am not familiar with any aspect of society which can be really described in deterministic terms.

All the social phenomena we encountered in our daily life are all probabilistic we can only talk in terms of chance like whether this lecture will end today at exactly 5 or 5:10 or 5:15 it is a matter of chance everything is subject to change that means, usefulness of statistical model will be much more in description of sociological variables than the usefulness of mathematical models. so, statistical models are particularly suited for the studies of random errors models using probability distributions Monte Carlo computer simulation because of their reliance on repeated computation of random or pseudo-random numbers, regression analysis discriminant analysis factor analysis they are all examples of statistical models.

I think most of you are not familiar with some of these terms Monte Carol simulation or regression analysis or factor analysis may be in a different class on statistics the you can come to know about the in Monte Carlo simulation we take a person randomly and using certain assumptions trace out all the relevant event that are possible in his life that way we go for second person third person and generate random events and on the basis of experiences of a selected number of person then aggregate statistics are obtained. And with **with** help of those statistics then we try to draw inferences about connections between different variables.

(Refer Slide Time: 09:40)



History of models: One of the earliest and most forceful arguments in favor of modeling was given by Nathan Keyfitz in the last lecture I took his name Nathan Keyfitz a well known demographer in the first volume of population and development review everybody is familiar with the journal population and development review one of the most read journals dealing with sociological aspects of population trends. In the very first issue of that journal in 1975 Keyfitz published an article with the title “How do we know the facts of demography?” He argued that the empirical relationships based on regression analysis can be misleading as they depend heavily on the cases for which data are available. Also without modeling questions regarding cause and effect, multiple causation, and nature of relationship could not be answered you know what he is trying to say if you have a regression analysis a relationship between two variables because one of them is called dependent variables another independent variable.

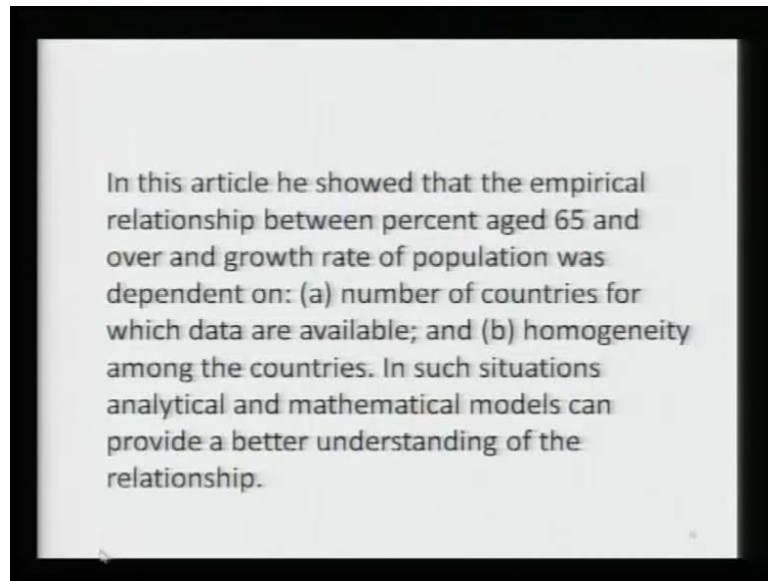
Like dependent variable may be growth rate of population and independent variable may be birth rate and you run a regression analysis between them regress growth rate on birth rate using a equation that growth rate equal to a plus b into birth rate plus some error term and on the basis of empirical data from a number of countries estimate the values of a and b then you have found a relationship between growth rate and birth rate that may be called empirical relationship. The problem with empirical relationship is that what the values of a and b in your regression analysis would be depends on how many countries you have included which countries you have included regression analysis on growth rate and birth rate will give you one set of values of a and b if you run this regression for developing countries and another set of values if you run this regression for developed countries.

If you have a mixture then you are going to have another set of a and b then which values of a and b really describes relationships between growth rate and birth rate cannot be said on empirical basis. So, to understand if there is a real link between the two you have to depend on models so, Nathan Keyfitz argued that for this was Nathan Keyfitz example growth rate and birth rate ratio or the relationship between age structure may be proportion of population at is 60 and 65 or more and birth rate of the country.

Again you can have empirical relationships, but those empirical relationships are not universal they depend heavily on which countries you have included in the analysis whether they are socio-economically advanced or less advanced or just developing what

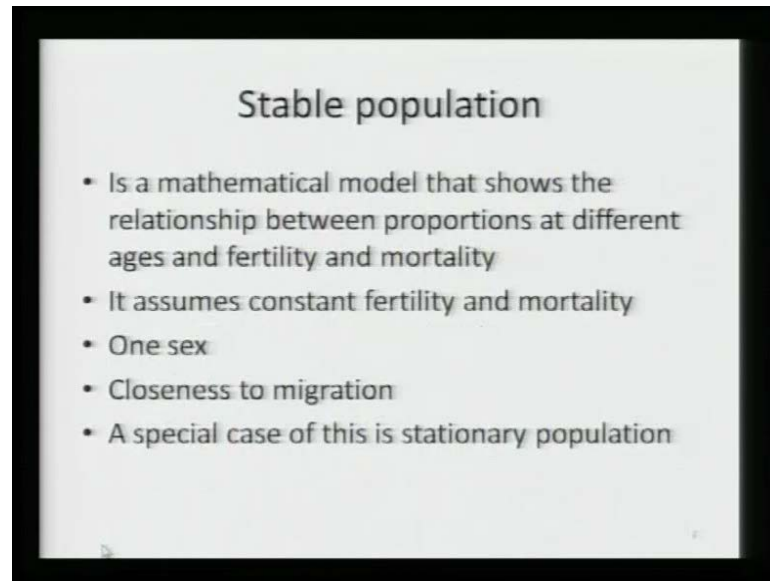
is demographic profile at what stage of transition their populations are so many factors, and therefore, we need to build models in terms of this search relationships can be studied. In this article he showed that empirical relationship between percent aged 65 and over and growth rate of population while dependent on number of countries which data are available.

(Refer Slide Time: 13:17)



Homogeneity among the countries whether all the countries are similar or different in such situations analytical and mathematical models can provide a better understanding of the relationship this is what he argued.

(Refer Slide Time: 13:27)



One of such model in demography one of the most used most talked about model in demographic studies is the model of stable population this stable population model is a mathematical model that shows the relationships between proportions of populations at different ages and fertility and mortality. In developing this model it is assumed all models are based on certain assumptions building this model it is assumed that fertility and mortality rates are constant.

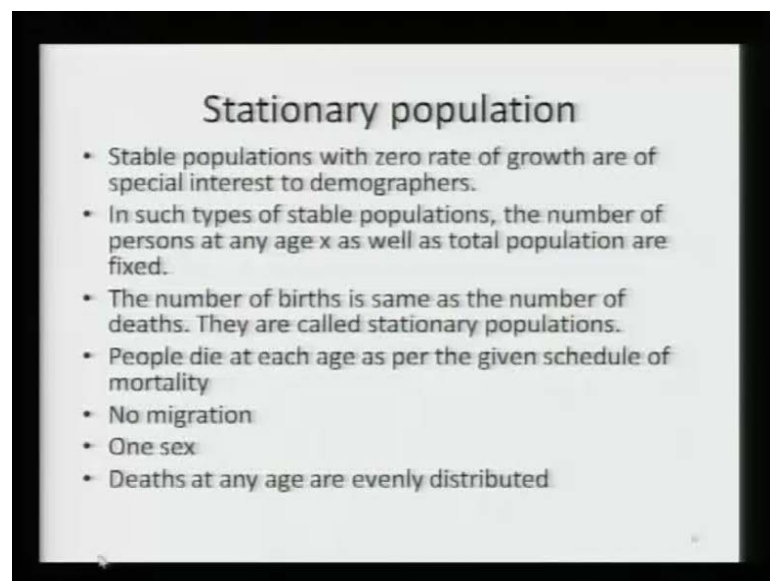
Actually long back it was observed by a demographer that if you take a population, and subject it to a given scheduled of fertility and mortality rates then after a certain period of time its age distribution becomes stationary kind of there is no further change in the age distribution of population that age distribution of population was taken to be the limiting case of the stable population with given schedules of fertility and mortality. Other assumptions are that the whole population consist of one sex, and I will say how is possible no population consist of one sex, but these are models, and as I said still population is one such model.

Somebody observed long back that if you take a population with any type of age distribution subjected to a constant schedule of fertility and mortality for sufficiently long period of time then you find that after sometimes age distribution of this population becomes stationary means percentage of populations at different ages becomes constant. That age that limiting stationary age distribution can be called stable age distribution,

and that stable age distribution depends totally on schedules of fertility and mortality, and ultimately on birth and death rate only.

Not on the previous age distribution with which the population started this is stable **stable** population model, now there are some other assumptions all models are based on certain assumptions, and one assumption is that the population consist of one sex only you will ask how it is possible? No human population can consist of one sex yes, well it is an assumption. All models are representations of reality with respect to one or two selective aspects. So, here we are assuming that the whole population consists of one sex. Another assumption that the population is close to migration; that means, there is neither any in migration nor any out migration or in case of international migration you can say there is no immigration no emigration and special case of this is stationary population.

(Refer Slide Time: 16:32)

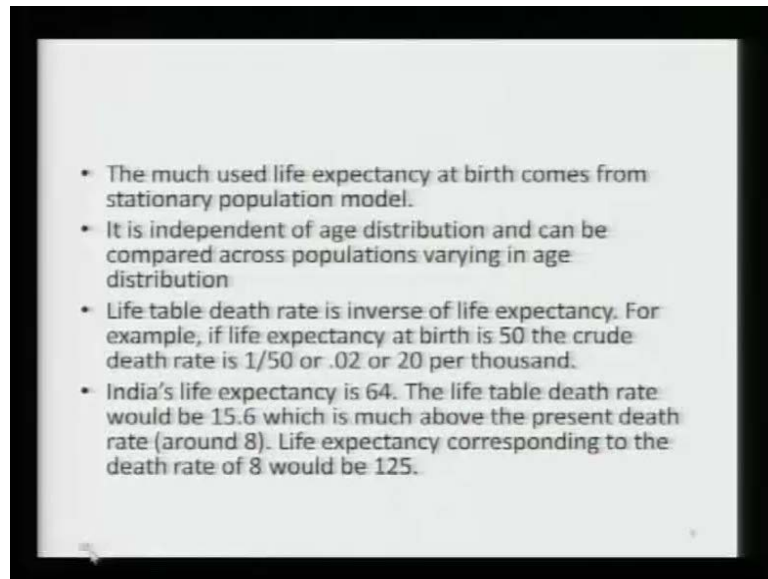


Stationary stable populations with zero rate of growth are called stationary population and they are of special interest to demographers in such types of stable populations the number of persons at any age x as well as total populations are fixed. The number of births is same as the number of deaths and they are called stationary populations people die at each age as per the given schedule of mortality there is no migration the whole population consist of one sex and death at any age are evenly distributed, those of you who have ever seen a life table for some countries somewhere recently world health

organization has prepared life tables for most countries of the world for which reliable data exist and you can see that life tables.

It is interesting to see what life table shows that life table age distribution is the stationary model of population.

(Refer Slide Time: 17:37)



The much used life expectancy at birth comes from stationary population model and that is one reason why for comparing health status of different countries we use life expectancy and not death rate. Because death rate depends on the age distribution of population life expectancy does not depend on the age distribution of population the age distribution of population is itself generated by the schedule of mortality or age specific death rates.

So, life expectancy is independent of age distribution and can be compared across populations varying in age distribution from this kind of age distribution or stationary populations also you can calculate death rate. And though death rates can be compared as much as life expectancies for different countries there is a simple connection that life table death rate is just the inverse of life expectancy. So, for example, if the expectancy at birth is 50 the crude death rate for that country would be 1 upon 50 or 0.02 or 20 per 1000 population

If the life expectancy is 100 then crude birth rate would be one upon 100 or 0.01 or 10 per 1000, you see today most countries have crude death rate round 7 or 8, but their life expectancy varies from as low as 40 to as high as 80. The reason is that this crude death rate being dependent on the age distribution cannot be compared, but this life expectancy from 40 to 80 can be compared. And you can actually calculate crude death rates corresponding to these life expectancies by taking reverse of them .

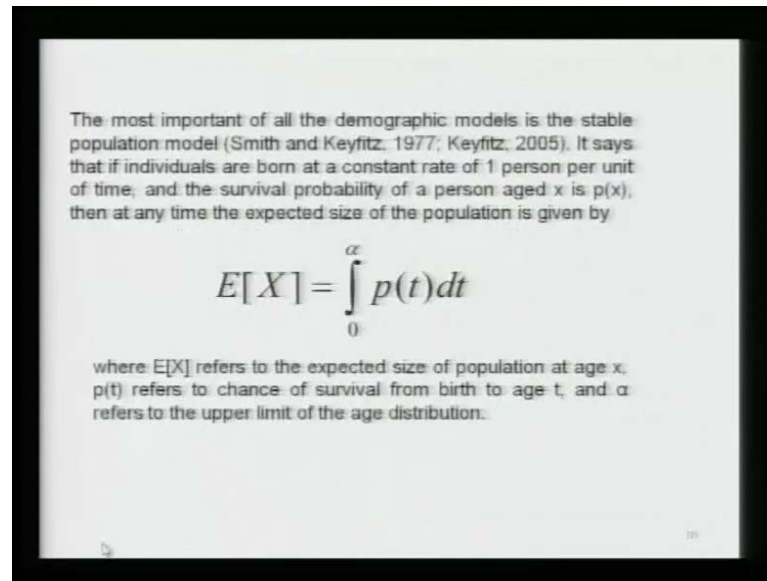
So, for life expectancy of 40 crude death rates would be one upon 40 or 25 per 1000 for life expectancy of 80 it will be one upon 80 India's life expectancy is 64. So, the life table death rate would be 15.6 means one upon 64 **one upon 64** multiplied by 1000 if we express crude death rates in per 1000 terms and that comes out to be 15.6. So, this 15.6 is the life table death rate actual death **death** rate of India is around 8 see the difference actual death rate of India is 8 life table or stationary death rate would be 15.6. The difference is that this crude death rate of India around 8 shows the experience of two things the present age distribution of population which is quite young and the age specific death rates operating at different ages.

While this life table death rates 15.6 would be the death rate of India which will result from the present age specific death rates and age distribution of life table population or stationary population which is itself generated from age specific death rates.

So, life table **life table** death rate is a function of only one thing age specific death rates set of age specific death rates or schedule of mortality while actual crude death rate is a function of two things schedule of mortality and the present age distribution.

Life table is therefore, a better thing to compare health status of different countries or achievements in the field of health, family welfare, gender, empowerment of women, etc. Life table life expectancy corresponding to the death rate of 8 which is usually the death rate of different countries today would be 100 and 25 see the big gap

(Refer Slide Time: 21:45)



The most important of all the demographic models is the stable population model (Smith and Keyfitz, 1977; Keyfitz, 2005). It says that if individuals are born at a constant rate of 1 person per unit of time, and the survival probability of a person aged x is $p(x)$, then at any time the expected size of the population is given by

$$E[X] = \int_0^{\alpha} p(t) dt$$

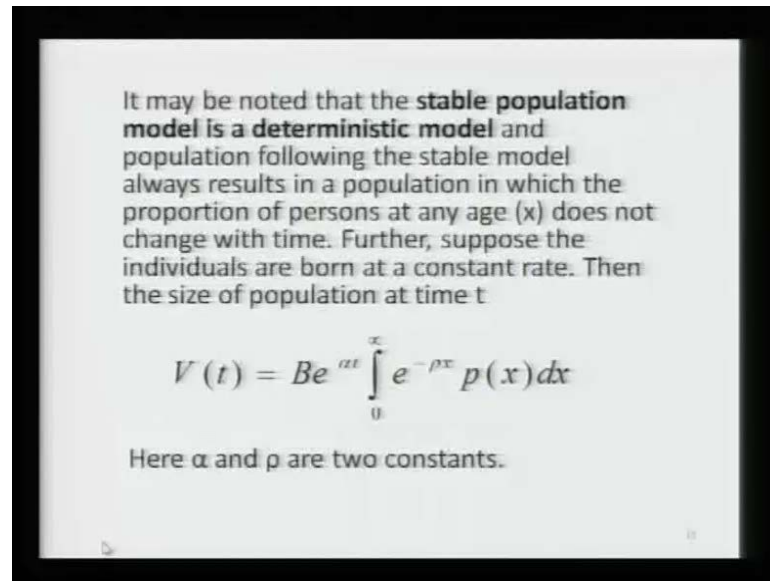
where $E[X]$ refers to the expected size of population at age x , $p(t)$ refers to chance of survival from birth to age t , and α refers to the upper limit of the age distribution.

The most important of all the demographic models is therefore, the stable population model, it says that if individuals are born at a you can derive this mathematically that individuals are born at a constant rate of 1 percent per unit of time it may be year, it may be month, day or something. So, 1 percent per unit per year and the survival probability of a person aged x is $p(x)$ then at any time the expected size of population is given by expected value of x equal to integral 0 to α $p(t) dt$, where expected value of x refers to the expected size of population at age x $p(t)$ refers to chance of survival from birth to age t and α refers to upper limit of the age distribution.

Empirically you can find out what is the α beyond which nobody lives in case of India say α may be taken to be 1985 or 90 for some development can be taken as 95 or 100. So, this is an upper limit upper possible limit there is a difference between life expectancy and longevity. Life expectancy refers to expected years one is expected to live a child who is born today what will be his average age of death or at what age is that child expected to die that is life expectancy. So, a child born today in India is expected to live for 64 years our age specific death rates are such.

Longevity means that this is the age beyond which nobody survives. So, life expectancy is 64 longevity may be 85 or 90 beyond 85 or 90 nobody survives in India something it is an assumption.

(Refer Slide Time: 23:49)



It may be noted that the **stable population model is a deterministic model** and population following the stable model always results in a population in which the proportion of persons at any age (x) does not change with time. Further, suppose the individuals are born at a constant rate. Then the size of population at time t

$$V(t) = Be^{\alpha t} \int_0^{\infty} e^{-\rho x} p(x) dx$$

Here α and ρ are two constants.

It may be noted that stable population model is a deterministic model we are not talking of probabilities, we are talking of a fixed relationship and population following a stable model always results in a population in which the proportion of persons at any age x does not change with time. Further suppose the individuals are born at a constant rate **rate** other than one per unit of time then the size of population at time t would be V equal to $B e^{\alpha t} \int_0^{\infty} e^{-\rho x} p(x) dx$. Now a new term ρ has appeared in this equation and $p(x) dx$.

Here α and ρ are two constants in any model there are certain constants they **they** are called parameters of the models. So, there are two constants α and ρ in the earlier case in this case there was **there was** only one parameter here I wrote only one constant say α , but now you have two things you have α you have ρ and it is written 0 to infinity.

(Refer Slide Time: 25:08)

This population grows or declines at rate: ρ .

Its age distribution, i.e., proportion surviving to age x is proportional to:

$$e^{-\rho x} p(x)$$

Using the property of stable populations that age distributions of two stable populations never cross each other census growth rate and proportion of population up to a certain age (normally 35 years) were used to estimate birth and death rates for those populations which lacked reliable and complete data on them.

13

So, you have two constants alpha and these are these constants are called parameters of the model, this population grows or declines at rate rho. So, now, you have a concept of growth rate of population also its age stationary population has zero growth rate, but stable population can grow can decline and in limiting case when the growth rate of stable population is zero it is called stationary population.

When stable population with rate of growth zero are called stationary population that is the difference that a stable population is more realistic that way because the stable population can expand it can increase in size, it can decrease in size it can remain constant and that limiting case of stable population in which growth rate of population is zero is called stationary population its age distribution means age distribution of stable population that if proportion surviving to age x is proportional to $e^{-\rho x} p(x)$.

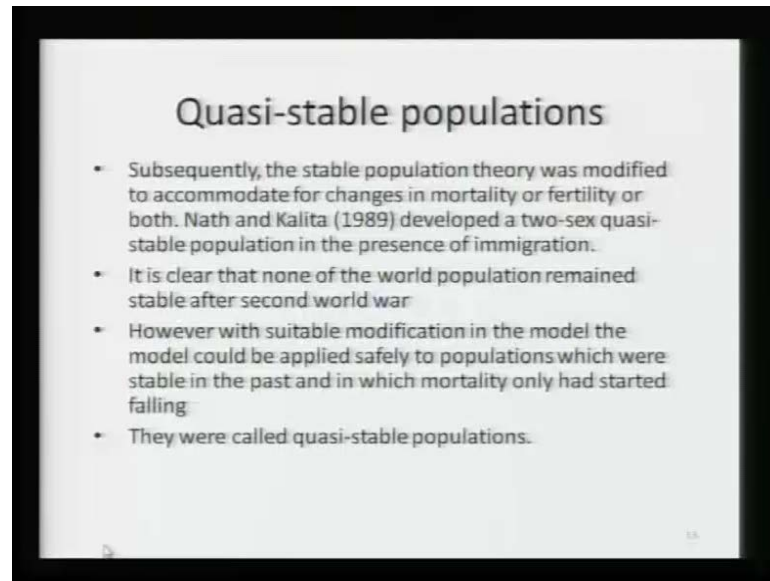
Now, using the property of stable population that age distribution of two stable population never cross each other census growth rate and proportions of population up to a certain age normally 35 years were used to estimate birth and death rates for those populations which lacked growth rate of population is zero is called stationary population its age distribution means age distribution of stable population that if proportion surviving to age x is proportional to $e^{-\rho x} p(x)$.

Now, using the property of stable population that age distribution of two stable population never cross each other census growth rate and proportions of population up to a certain age, normally 35 years were used to estimate birth and death rates for those populations which lacked reliable and complete data on them. I know that is for students of sociology who do not have background of mathematics it may be difficult to follow all these formulae there are only three lectures anyway devoted to modeling and you have to bear with me.

What you can understand from this modeling or from this lecture is that one that models are about relationships, either mathematical or statistical between certain aspects of demographic reality. This these aspects may relate to processes or this aspects may relate to composition or structure they may relate one aspect of composition to one aspect of processes. Like here one aspect of composition means age distributions is being related to growth rate of population and once you have such kind of relationships here I mention one example of how these models were used in 50 and 60 for less developed countries we did not have reliable data on birth and death rates.

But for most countries we had some figure of growth rate, we also had some idea of age distribution of these populations and using stable population theory or stable population model and these two things either growth rate or proportion of population below 35 or above 35 or both age distribution and growth rate demographer estimated birth and death rates for various countries which were formed to be quite acceptable by the planners and demographers.

(Refer Slide Time: 28:44)



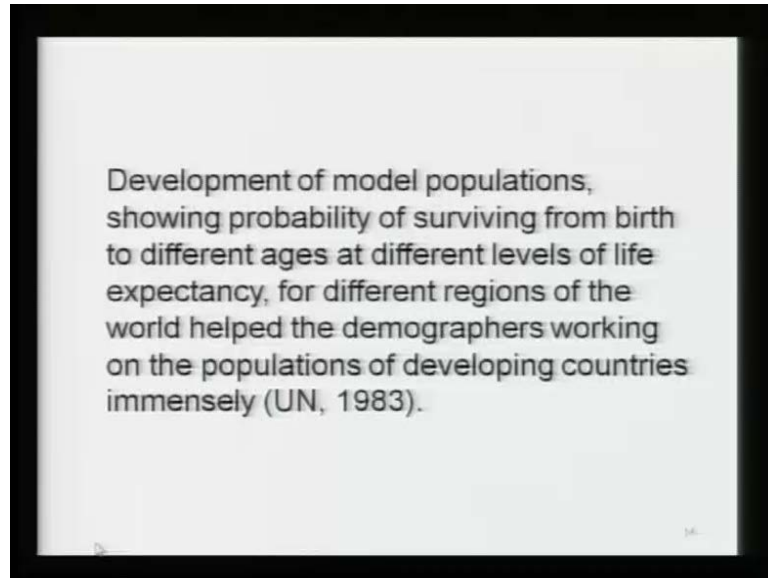
Problem in this is as we move from simple to complex I said that when assumptions are simple or you are showing depicting only one aspect of reality you have simple model when you have more things to represent then you have complex model. Now, in stable population it is assumed that schedules of fertility and mortality are constant and growth rates are 0, that model could easily be used for a large number of populations up to 1950s. In less developed countries there were many populations which were not growing at all and in which birth and death rates were constant. So, stable population model was of substantial help.

Gradually in 1960s, 1970s and for many countries in 1950 itself their death rate started falling. So, this assumption of constancy of fertility or and mortality schedules was violated now you had a situation in which fertility schedule was fixed, but mortality was declining or mortality was improving life expectancy was going up and crude death rates were declining. So, the assumption of stable population model were violated in that situation a relatively more complex model which was called quasi stable population wise developed in which one aspect of population processes fertility is constant and mortality changing.

So, for application to less developed countries for population 1970s and 1980s quasi stable population model became more useful and it was more widely used than the stable population model that could be used for populations up to 1940 or 1950s. Now, it is

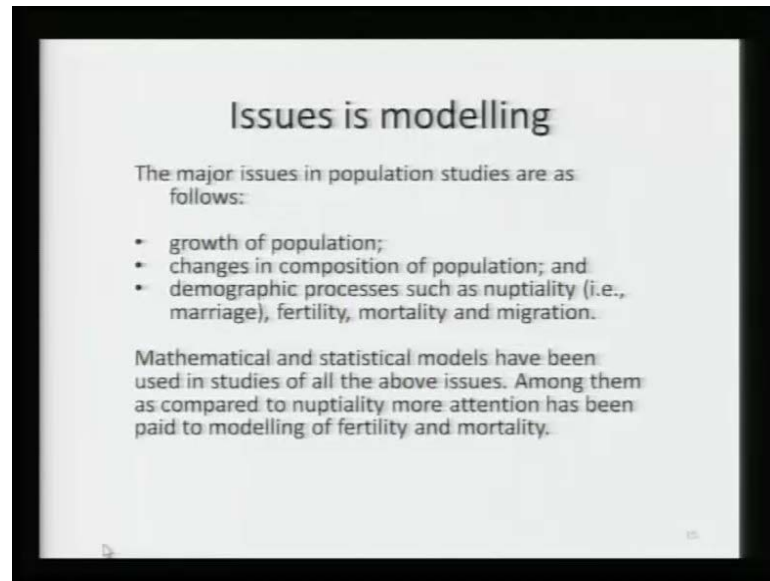
cleared that none of the world population remain stable after Second World War, but with suitable modification in a mortality component the model could safely be applied to these populations.

(Refer Slide Time: 30:55)



The stable population model with this modifications was called quasi stable population model, development of model populations showing probability of surviving from birth to different ages at different levels of life expectancy for different regions of the world helped the demographers working on the populations of developing countries immensely for estimation of rates for population projections for planning purposes and for variety of other purposes.

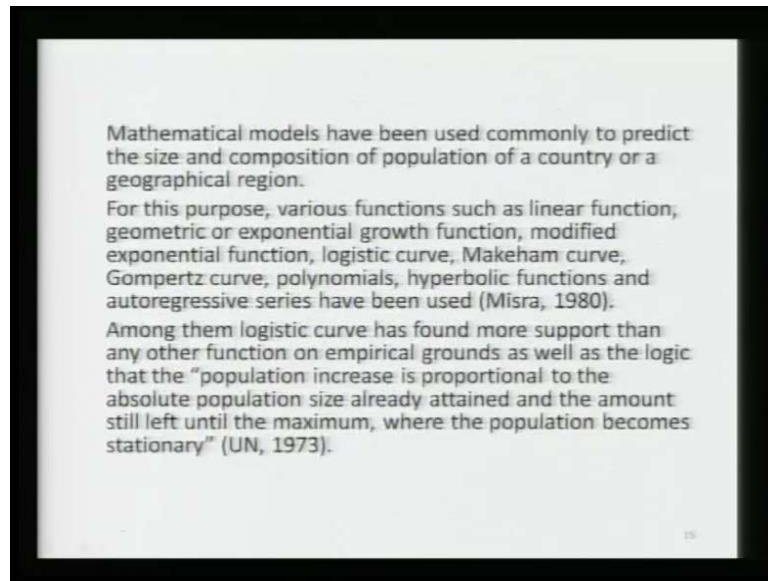
(Refer Slide Time: 31:25)



This shows the importance of modeling, now the issues of modeling in demography or the major issues in population studies are as follows. Growth of populations can we describe growth of population in general in terms of a mathematical model or growth curve. Changes in composition of populations age structures, marital status, occupational, economic and so on. Among them description of age composition of population is more sort after than demographic processes such as nuptiality and marriages fertility mortality and migration.

You have less models for migration more models for fertility and mortality, mathematical and statistical models have been used in studies of all these issues among them as compared to nuptiality more attention has been paid to modeling of fertility and mortality.

(Refer Slide Time: 32:20)



Mathematical models have been used commonly to predict the size and composition of population of a country or geographical region like population of India, population of different states of India for this purpose various functions such as linear function, geometric, exponential growth model, modified exponential function, logistic curve, Makeham curve, Gompertz curve, polynomials, hyperbolic functions and auto regressive series have been used.

I have decided more time to devote to these things this is something which interests me very much, but in this basic course on population and society we are more to discuss social aspect of population and not the models, but those who are some of you may be interested in modeling, may look for literature on all these things these are growth curves among them logistic curve. I will just talk about logistic curve logistic curve has been found more useful it has found more support than any other function, linear, geometric, exponential, Gompertz or polynomial as well as the logic **logic** behind logistic curve that population increase is proportional to the absolute population size already attained and the amount is still left until the maximum where the population becomes stationary is achieved.

Now, the idea is if you take exponential model then the population is starts from 0 and it very soon it increases to very high number and ultimately moves towards infinity, the common sense say that no population can ever become infinity. Similarly, if you are to

predict proportion population living in urban area you know that proportion cannot exceed 1 or 100 percent, but if you use polynomials linear models and regress percentage urban on socio-economic variables.

Then some of your predictions may be in accurate or unacceptable and they can go beyond 100 percent or the figure of one ratio of one, in such cases also logistics model which found to be more useful.

(Refer Slide Time: 34:46)

Among all the growth models existing in literature logistic model is most widely used model of population growth. The model and its assumptions are discussed below.

Assuming that r is the maximum rate of increase of the population,

$$\frac{K - N}{K}$$

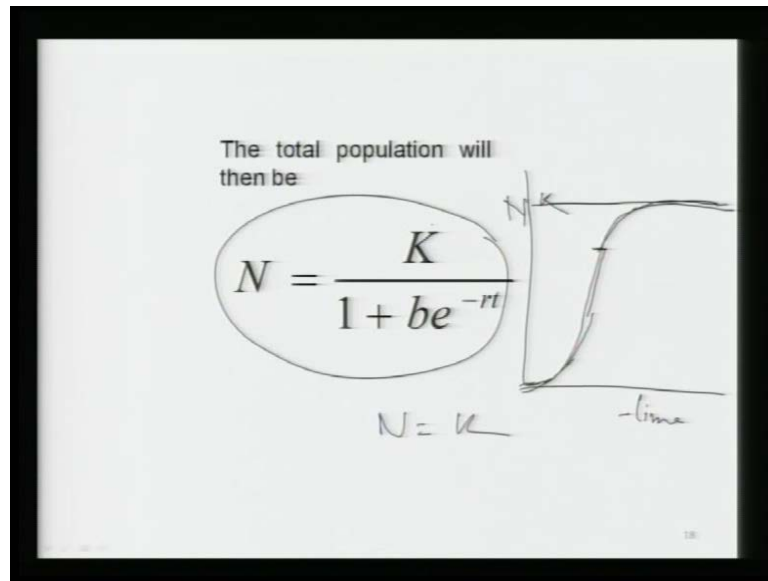
the fraction by which the actual population (N) remains below the maximum (K), then the increase of population per unit of time is

$$\frac{dN}{dT} = rN \frac{K - N}{K}$$

So, now let us look at the logistic model some aspects of it and what ultimately it gives us assuming that r is the maximum rate of increase of the population what does maximum rate means **maximum rate means** that population raises at a slow rate then its rate increases. But once a certain number is attained the rate of growth of population starts falling and let that be K minus N upon K the fraction the fractions by which the actual population N is the actual population means population at a given point of time below the maximum K population can go upto size K and N is the size at a given point of time.

Then the population per unit of time change in population per unit in time dN , N for population T for time dN by dT is rate of growth a into N into K minus N upon K suppose you can make this assumption.

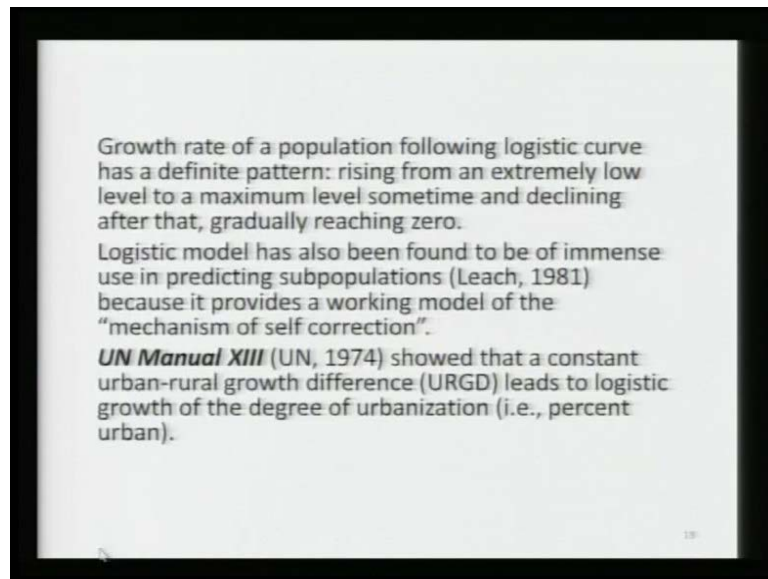
(Refer Slide Time: 35:58)



The rate of change dN by dT depends on rate of growth on actual population and on how far the actual population is from the limiting case K , it can be shown mathematically that under this assumption the curve that you get or the number in terms of **in terms of** time K etc follows the following logistic model. N equal to K upon $1 + b e$ raise power minus $r t$. So, it is an interesting if I want to show this then this model would look like this t time on x axis and on y axis gradually we start with this small population and as time passes population is increasing.

But increasing at a slower **slower** rate of growth in the beginning with time the rate of growth of population is increasing, it may increase further a times come when rate of growth start declining or the population is growing. So, population is growing rate of growth is falling and ultimately at some point of time when the size of population reaches K when N equal to K when the size of population reaches K then the population is stabilize.

(Refer Slide Time: 37:19)



Now, this kind of you know this growth curve is very well depicted with the help of this model which is called logistic growth model and this logistic growth model has been found to be of immense help in forecasting of population of different countries. Growth rate of a population following logistic curve has a definite pattern rising from an extremely low level to a maximum level sometime and declining after that gradually reaching 0.

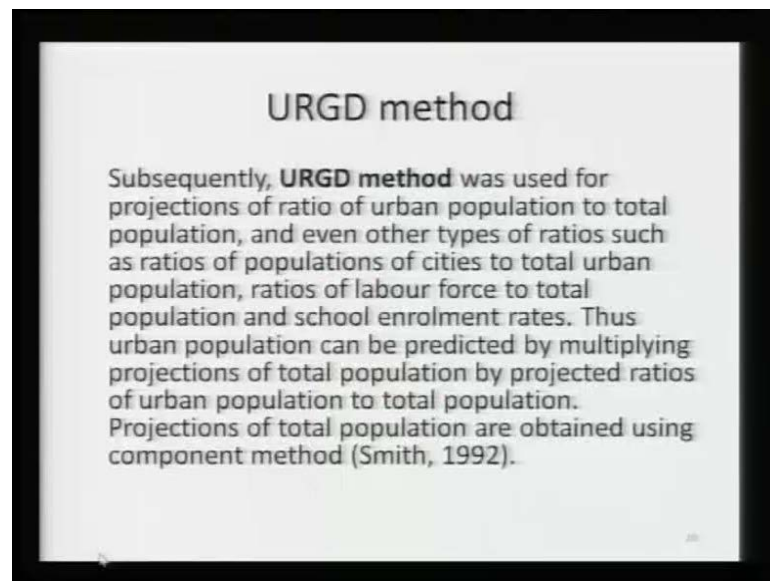
Logistic model has also been found to be of immense use in predicting subpopulations means urban and rural populations of different occupation proportion married by age populations in different industries like primary, secondary, tertiary as time passes proportion of population engaged in primary sector decreases proportion engaged in manufacturing activity increases for those proportions also logistic curve can be used for predictions. UN Manual 8 shows that a constant urban rural growth difference leads to logistic growth of the degree of urbanization means if you can assume that both urban and rural populations are growing exponentially, but the difference in growth rates of urban and rural areas is constant means urban rural growth difference URGD is constant then the proportion urban or percentage urban follow the logistic model.

Which means in the beginning rising from an extremely low level to a maximum level in the beginning when urbanization is low when virtually the whole population is living in villages urbanization or percentage urban or proportion urban rises at a very slow pace. But as time passes pace of urbanization increases a time comes when pace of urbanization reaches the maximum level for which we use the term r in logistic growth

model after that growth rate of urbanization starts falling and although a percentage of urban keeps on increasing, but growth rate of urban population growth rate of urbanization starts falling.

And ultimately at some level may be 70 percent or 80 percent or 90 percent or 95 percent urbanization gets stabilized. So, level of percentage urban will not go beyond that level.

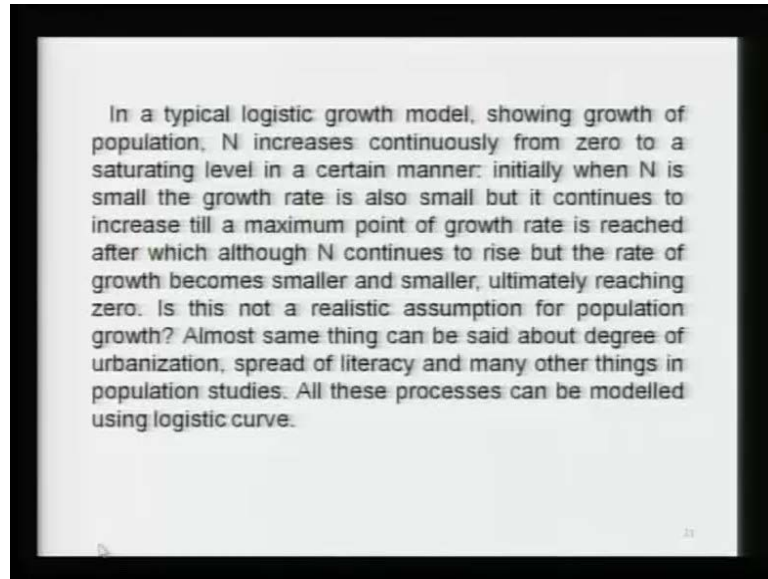
(Refer Slide Time: 39:49)



URGD method was used for projections of ratio urban population to total population and even other types of ratios such as ratios of cities to total urban population I remember when this manual 8 was published I was a student of IIPS and in India we were the first forecast population of urban areas of Maharashtra and of different cities of Maharashtra using this URGD model and this projection was found to be quite useful by government of Maharashtra. Similarly ratios of labour force total population and school enrollment rates, school enrollment rate also follow the same model.

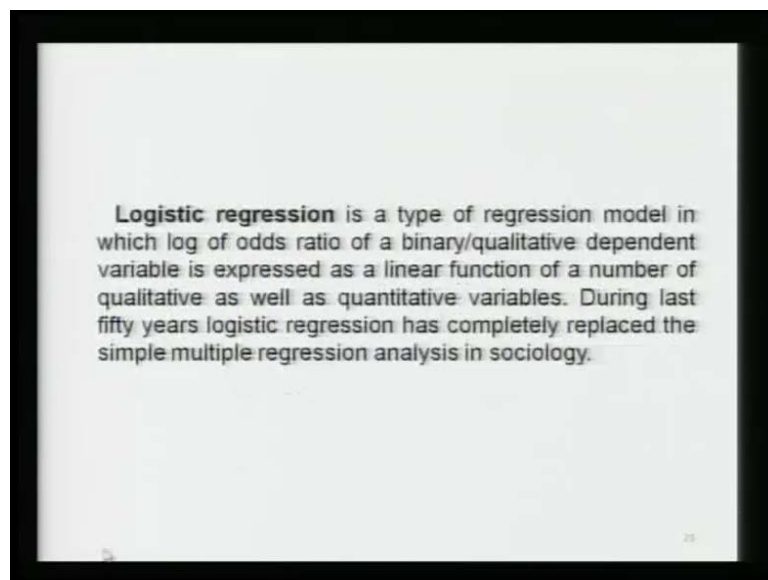
Primary school enrollment secondary school enrollment they all start with 0 there is some maximum point you can never your primary school enrollment can go from 0 to 100, secondary school enrollment rate can go from say 0 to 80. University enrollment rate will never reach 100, it may go from 0 to 40, 50 maximum say 60 and different logistic models can be used to forecast enrollment rates for different countries at different stage of development.

(Refer Slide Time: 41:09)



In a typical logistic growth model showing growth of population N increases continuously from zero to saturating level in a certain manner, initially when N is small growth rate is also small it continues to increase this is what I have already stated.

(Refer Slide Time: 41:26)

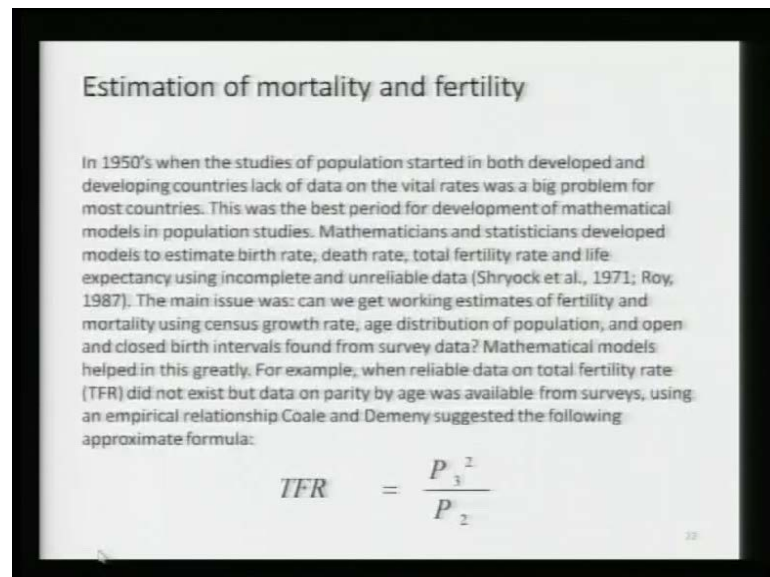


Now, logistic regression is a type of regression this is another model these days in sociology when you open a journal of sociology particularly from America, American journal of sociology or American sociological review you find that most articles are based on analysis of empirical data and most articles are using logistic regression. There

was a time say in fifties sixties when sociologists made use of multiple regression analysis as a model today more sociologists are going for logistic regression.

It is a type of regression model in which log of odds ratio of a binary qualitative dependent variable is expressed as a linear function of a number of qualitative as well as quantitative variables. So, here dependent variables can be migrant non migrants. Migration status whether somebody is migrant or non-migrant or somebody is going for sex determination test or not going for it and the independent variables are usual socio-economic variables.

(Refer Slide Time: 42:41)



Estimation of mortality and fertility

In 1950's when the studies of population started in both developed and developing countries lack of data on the vital rates was a big problem for most countries. This was the best period for development of mathematical models in population studies. Mathematicians and statisticians developed models to estimate birth rate, death rate, total fertility rate and life expectancy using incomplete and unreliable data (Shryock et al., 1971; Roy, 1987). The main issue was: can we get working estimates of fertility and mortality using census growth rate, age distribution of population, and open and closed birth intervals found from survey data? Mathematical models helped in this greatly. For example, when reliable data on total fertility rate (TFR) did not exist but data on parity by age was available from surveys, using an empirical relationship Coale and Demeny suggested the following approximate formula:

$$TFR = \frac{P_3^2}{P_2}$$

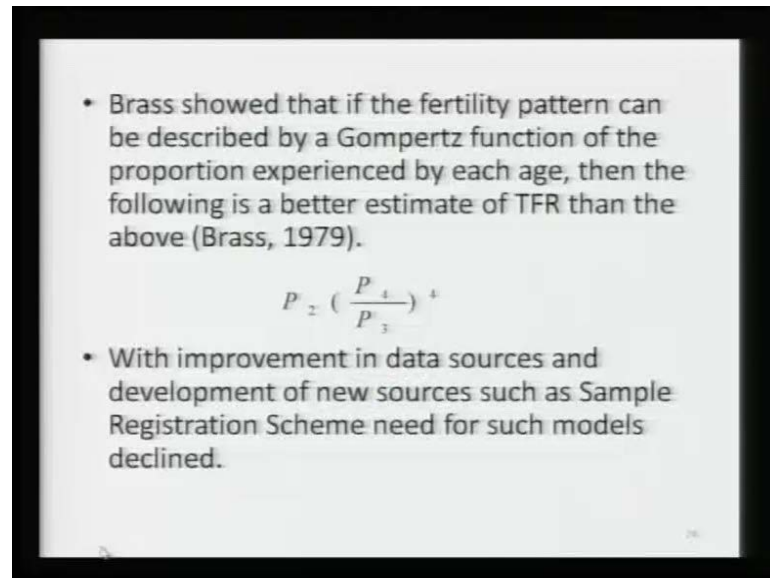
12

During last 50 years logistic regression has completely replaced the simple multiple regression analysis in sociology, then models have been used in estimation of mortality and fertility then from sciences we had some data on average parity or average number of children born from different age groups starting from 15 to 19 this **this** P 2 average parity this P 2 while use for average parity of age group 20 to 25. 15 to 20 is first second age group for reproduction is 20 to 25 and P 2 refers to average **average** parity of women in 20 to 25, P 3 refers to average parity of women in age group of 25 to 30.

Now, when we did not have reliable data on age specific fertility rates models were used this is another example of use of models that coale and demeny suggested that you can use an approximate value of **of** total fertility rate like this. Square average number of children or average parity of women in 25 to 30 and divided by average parity of women

in 20 to 25. Why did P 3 square by P 2 resemble total fertility rate. This is a model and this model is developed on the basis of certain empirical and analytical understanding of relationship between total fertility rate and parity.

(Refer Slide Time: 44:14)



- Brass showed that if the fertility pattern can be described by a Gompertz function of the proportion experienced by each age, then the following is a better estimate of TFR than the above (Brass, 1979).

$$P_2 \left(\frac{P_4}{P_3} \right)^4$$

- With improvement in data sources and development of new sources such as Sample Registration Scheme need for such models declined.

Brass another demographer showed that you get better estimate if in place of using your previous function P 3 square by P 2. You go for P 2 multiplied by P 2 multiplied by P 4 by P 3 raise to power 4 here P 2 refers to average parity in age group 20 to 25. P 3 is parity in age group 25 to 30 and P 4 is average parity in age group 30 to 35. So, depending on to which data to data in which group your sign more important more you think that 15 to 20 is more reliable or 20 to 25 is more reliable or 25 to 30 or 30 to 35.

And depending on your empirical and analytical understanding of relationship between TFR this is another estimate of TFR then TFR can be estimated from the previous equation P 3 square by P 2 or from this equation that TFR equal to P 2 multiplied by P 4 by P 3 raise to power 4. With improvement in data sources development of a new of new sources such as sample registration scheme need need for such models as declined, this is another example of application of models in mortality.

(Refer Slide Time: 45:42)

Knud (1983) compared cancer mortality between sexes, cohorts and cities by using Poisson distribution for number of deaths at a particular age and the mortality rate (defined as chance of survival to age x) as follows:

$$l_x = b x^k$$

b and k are two parameters for which maximum likelihood estimates were obtained.

Here l_x this in stationary population this tells us exactly what number of people are surviving to age x and this is expressed b a constant raise to power x raise to power k there are two parameters or constants in this model b and k this b and k are obtained these parameters are obtained from empirical data.

(Refer Slide Time: 46:15)

Gompertz model has been commonly applied for studying mortality. Using an age-dependent shape parameter, Weon (2004) used a Weibull model for mortality rate $\mu(t)$, i.e., ratio of density ($f(t)$) and survival functions ($S(t) = 1 - F(t)$) for estimating maximum longevity, as follows:

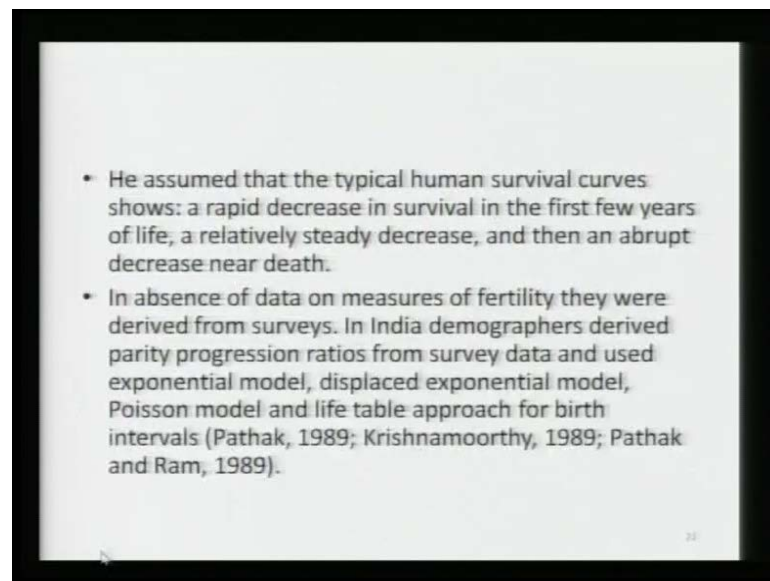
$$S(t) = \exp(-t / \alpha)^{\beta(t)}$$
$$\mu(t) = (t / \alpha)^{\beta(t)} * \left[\frac{\beta(t)}{t} + \ln(t / \alpha) * \frac{d\beta(t)}{dt} \right]$$

Similarly, Gompertz model has been commonly applied for studying mortality and here survival functions are expressed in terms of age **age** etc, at different points of time t. Those of you who come from some mathematics background may be interested in these

things and those who do not have sufficient mathematics background to understand these things may ignore just try to understand that it is possible to connect different aspects of demographic processes and composition by using mathematical and statistical model this is the only thing that I wanted to convey in this class.

The person who developed this model I should the gompertz model has been applied for studying mortality weon used a this weibull model for mortality rate.

(Refer Slide Time: 47:09)



And he assumed that typical human survival curves shows a rapid decrease in survival in the first few years of life and a relatively steady decrease and then an abrupt decrease near death. In absence of data on measures of fertility they were derived from surveys in India demographers derived parity progressions P_1, P_2, P_3, P_4 and using the formulae I have just now shown we estimated total fertility rate for India. So, like this you can it is possible everybody will not be interested in this, but what I want to stay it is possible at least some in some cases to build relationships between different components of population.

Growth rate, fertility rate, migration rate, mortality rate, marriage nuptiality or growth rate of populations as a logistic curve and it is **it is** possible to represent them with the help of growth models, it is also possible develope mathematical equations showing one or more of these things on the left hand side of equations means a dependent variables

and socio-economic certain demographic variables themselves on the right hand side as independent variables.

If you are using equations derived analytically using mathematical formulae, mathematical theories then you have mathematical model and if you are expressing behavior of peoples in terms of probabilities or you are making use of probability density functions, like some of you make familiar with binomial probability function Poisson function or a normal distribution. If you are using such distributions to describe behavior of something like number of accidents at some cross road per day what is the number of accidents and if you show number of accidents using Poisson **Poisson** probability distribution $e^{-\lambda} \frac{\lambda^x}{x!}$ you are using statistical or stochastic model.

We will talk about this model once more in the next lecture I will talk about migration model I will talk about the whole process involved in building models and I will also raise a few issues of modeling before demographers in our times.