Some demographic methods applied to urban and rural populations of Pakistan

by

FAISAL SHAHZAD

Thesis for the degree of Master of Science in Statistics (Data Analysis)

at the Department of Mathematics, University of Bergen

Thesis date: 1st June 2011

To

My Beloved Mom (Late) and Dad

Acknowledgements:-

All the praises and thanks are for Almighty ALLAH who bestowed me with the potential and ability to make it possible.

First of all I would like to express my sincere thanks and deep gratitude to my supervisor Dag Bjarne Tjøstheim, Professor, Department of Mathematics, University of Bergen Norway for his guidance, invaluable suggestions, discerning supervision and encouragement throughout the research work. I am pretty grateful to my teachers especially to Trygve Nilsen and Ivar Heuch for their support and assistance at the department. Further, I am also thankful to the administrative staff at the department and particularly to Kristine Lysnes for her continuous assistance.

Then I feel that I do not own this dissertation solely; honestly it is also the effort of my family as they morally supported me in a hard time of thesis. Especially my thanks are due to my Mom (Late) and Dad for their never-ending prayers, too much love and emotional strength. Finally, I must not forget my brother Rafiq Shahzad, who made my dream achievable.

I also acknowledge my all Pakistani fellows especially to Tahir Farooq for giving me company and guidance during my stay in Bergen.

Last but not least, my special thanks goes to Jamal Nasir for his valuable comments and help during this work.

Contents

Page #

Summary of Findings: -

In this thesis, first of all I have tried to describe what is demography and different ways to collect demographic data. Then, I have applied some of the demographic techniques to the population of Pakistan. Here are my findings:

- First of all, I have considered the infant mortality in Pakistan and applied the test of hypothesis along with 2 x 2 table to show that there is a difference of facilities/services given by the government to the urban and rural area's population and find out the results of z and chi-square tests with p-value. The results indicate that there is really a huge difference of policies between urban and rural areas of Pakistan and I have found the p-value 0.00001 which show our hypothesis is highly significant. I have noticed that since only the 35% of the population is residing in the urban areas but still urban areas are under consideration all the times while the rest 65% areas having the less attention by the government institutions.

- Secondly by using the data given by Federal Bureau of Statistics, Pakistan I have set up different life tables for the total population, urban and rural population and for the male and females population of Pakistan. The results show that the life expectancy at birth in urban area (68.7 years) is 6% higher than the rural areas (64.3 years). Similarly, the probability of dying at the first age interval is also 10% smaller in the urban area then the rural one (i.e. 0.06444 & 0.07197 for urban and rural respectively). Moreover, the female life expectancy at birth (68.4 years) is found to be 7% higher than the male life expectancy (64.3 years).

- Third, I have applied decomposition technique introduced Kitawaga (1955) to see how much of the difference between death rates in urban and rural population is attributable to differences in their age distributions. The results shows that the original difference between the urban and rural population is -0.00210 (by equation 7.2) while the contribution of age compositional differences and contributions of age specific rate differences are -0.00052807 & -0.00157492 respectively (by equation 7.3). Further, the proportion of difference attributable to differences in age composition is found to be 25% whereas the proportion of difference attributable to differences in rate schedules is 75% which shows that both parts are contributing in the same direction to the difference.

- Lastly I have tried to make a population forecasting for Pakistan. For this, a few methods has been discussed and have made a forecast by using the compound rate of growth method and cohort component method. According to the first method, it shows that there might be 294.96 million population in the year 2032 (equation 8.13) whereas the second method states that it might be 258.09 million in the year 2031 (equation 8.14). It seems reasonable to say that the estimates found by the cohort component method are more reliable than the any other method as the cohort component is now the only method on which demographers are relying much.

Tables

1. Introduction to Demography:-

[1][2]Demography is the social science that is defined as the scientific study of the size, composition and the distribution of the human populations and their changes resulting from fertility, mortality and migration. Demography is concerned with how large or small populations are, and how populations are composed according to age, sex, marital status, race and some other characteristics. Similarly, how the populations are distributed in a physical space (e.g. in urban and in rural regions).

Demography is also interested in the changes over time in size, composition and the distribution of human populations, and how these result from processes of fertility, mortality and the migration. The fertility, mortality and migration are all three basic processes of a demographic development. When-ever the populations change in size, composition, or distribution; the changes depend solely on one or more of the three demographic processes.

In other words, Demography is the study of many of the most important events in our lives and we are very much involved in these events as well. There are only two times in our life when we will have an almost 100 percent change of being identified by name and listed in our local newspaper. When we are born or when we die. These are events that demographers study. Similarly another extremely important event in the lives of many of us include getting married and also for some of us, getting divorced. These are the additional aspects that are studied by the demographers. Further, another important event that almost everyone is doing at least once in life is to move from one residence to another. So, demographers also study the change in residence. In short, it is not at all an overstatement to say that demographers study when we are born and when we die, as well as many of the really important events occurring in our lives in between.

Moreover, we can also say that the demography is “A statistical study of human population”. It is a very general science and can be applied to any kind of dynamic human population. It encompasses the study of the size, structure and the distribution of these populations and

---

[1] Population and Society, An Introduction to Demography by Dudley L. Poston and Leon F. Bouvier
[2] Demography: Measuring and Modeling Population Processes by Samuel Preston

time-based changes in them in response to birth, migration and death. The term demographics is often used erroneously for the demography but refers rather to the selected population characteristics as used in government or opinion research or marketing or demographic profiles used in such researches.

Demography requires the study of specific information that can be gathered from vital statistics records or population censuses and the people who study and records these information's are called demographers. Further, the demographers must know both how to obtain the information and how to interpret it scientifically.

Statistical concepts that are essential to the demography include birth and death rates, life expectancy and the infant mortality rates. These concepts can further be split for studying the expectancy of each gender. A census can provides these kinds of information and also the vital statistics. In some of the studies, demography of an area is extended to take into account income, education, family unit structure, housing, ethnicity, race and the religion.

Advertising relies on the demography, because good service providers need specific information to reach the maximum number of potential customers in their target. As well as, education also relies on the demography to help gathering information, to provide necessary governmental and/or local assistance. A large scale demographic example is the collection of data for an entire country and such data might be used to determine a need for world assistance due to famine, disease, infant mortality, or other serious issues.

It can also be said that demography is an interesting science used to create statistics. Sociology which is the study of social behavior and the society is an example of an independent area of study in which demography is often used. Similarly, economics is also a specific range of study that employs the science of demography. Anyone can review the basic information about the demography of any specific country by reviewing its most recent censuses.

1.1 Population:-

The population concept can be defined in many different ways like:

- The people who inhabit a territory or state.
- A group of organisms of the same species inhabiting a given sector.

- The entire aggregation of items from which samples can be drawn (statistics).
- The number of inhabitants (either total number or the number of a particular class or race) in a given area (city or country etc.).

To a statistician, “population” refers to a collection of items, for example balls in an urn. Demographers use the term “population” in a similar way to denote the collection of persons alive at a specified point in time who meets certain criteria. Thus they may refer to the population of China on July 1, 2010 or to the population of United States on July 1, 2010 etc. Moreover, demographers also use this term “population” to refer to a different kind of collectivity, one that persists through time even though its members are continuously changing through attrition and accession. Thus, the population of China may refer to the aggregate of persons who have been alive in the area and then collectivity persists even though a virtually complete turnover of its participants occurs at least once each century.

The demographic study focuses on this enduring collectivity. It is particularly addressed to study changes in its total size, the growth rates and its composition. But while the emphasis is on understanding the aggregate processes, demography is concerned with the implications of those processes for individuals. Thus, the frequent concern of demography is to trace out the consequences of changes at the individual level for the behavior of aggregate processes.

The area that is used to define a population is such that inter-breeding is possible between any pair within the area and it is also possible with cross-breeding with individuals of other areas.

A human population control is the practice of curtailing increase in population, usually by reducing the birth rates. An example of mandated population control is People’s Republic of China’s one-child policy. Having more than one child results in heavy fines that are imposed based on the income of the family and other factors. As a result of this policy, this has led to allegations that practices like forced sterilization and infanticide are used and the sex ratio at birth becomes 114 boys to 100 girls that may be evidence that the latter is often sex-selective. It can be helpful to distinguish between fertility control as individual decision making and population control as a state level or governmental policy of regulation population growth rate. But the fertility control may occur when individual or couples take steps to decrease or to regulate the timing of their own child-bearing.

Moreover, the term “population” is also used to refer to a set of potential measurements or values, including not only cases actually observed but those that are potential observable.

The World’s population is increasing very fast. Although many countries are trying to reduce their birth rate and some of the countries have taken serious steps but still it’s increasing day by day. An estimated population of the world is 6.5 billion (according to UN it is 6.8 billion).Asia is the largest continent not only by population but also by area. Its population consists of 60% of the total world’s population.

1.2 Demographic Data:-

Demographic data are the characteristics of a human population as used in government, marketing or opinion research or the demographic profile used in such researches. The demographics include gender, race, income, age, disabilities, mobility (in terms of travel time to work or number of vehicles available), home ownership, educational attainment, employment status and even location.

The basic sources of demographic data are the “National census”, “registers” and the “surveys”. The registration of population events are usually compiled and published annually or monthly but they are gathered continuously.

National census is the most important factor in the demographics. A census is the same as taking snapshot of a population at one point in time, say once every ten years and in this snapshot getting a picture of the size of the population its characteristics and its distribution. In the national census everyone in the population is enumerated and all the demographic events (births, deaths or migration) that occur in the population are supposed to be registered. Moreover, the census is the total process of collecting, compiling and publishing demographic, social and economic data pertaining at a specified time for all the persons in a country or territory. So, then the principal objective of a census is to obtain data about the size, composition and distribution of the population. A census may also contain the information on economically active and inactive parts of the populations, such as the data on the industrial and occupational composition of the marketing population and economic (salary and income) data.

On the other hand, registers may be thought of as a continuous compilation of major population events, such as births, deaths, marriages, divorces and sometimes migrations as well. When any birth or death occurs, it is registered with the government registration system thus occur continuously. Both census and registers are intended to cover the entire population.

The third important factor of demographic data is the surveys. By definition, the surveys are administered to only a fraction of the population. Yet they often gather the data on many of the items included in censuses and registers (i.e. birth, death or migration) plus the additional material of interest.

1.3 Registration systems:-

Registration systems pertain to the populations demographic events (births, deaths and in some place migrations) and measure them as they occur, while censuses provide a cross-sectional (one point in time) portrayal of the size, composition and distribution of the population. Moreover, censuses are static, registers are dynamic and continues activity. Although many countries are maintaining the registers of births, deaths, marriages, divorces and abortions and some of the countries also maintain migration registers, but principally registers apply to births and deaths only.

A population register is a list (i.e. a register) of persons that includes the name, date of birth, address and a personal identification number. In Europe, the Netherlands and the Nordic countries are maintaining some kind of population registers and many developing countries either have them in place or are planning to implement them, while in the eastern Europe, under the communists “population registers were used for control (of people) as well as for administrative purpose and the successor regimes for the most part have not maintained them. The United States of America doesn’t maintain any kind of national population registers.

For the demographers, these population registers are of interest because they contain birth and deaths record (certificates) but not all birth and deaths registrations occur in the context of population registers. In fact, since a large number of countries do not maintain any population register, the registration of many births and deaths occur outside the population register.

The vital events that are births, deaths, marriages, divorces, fetal deaths (still births) and induced termination of pregnancies (abortions), for most of the countries in the world, are recorded in their civil registration system but these registration systems need not necessarily be the population registers and indeed, many are not. Although civil registration data are not 100% accurate and complete, in the more developed countries their quality is far better than that in the developing countries. [3]In 1996, the Carla Abou Zahr and John Cleland study shows, the civil registration systems in the developing countries are "seriously defective, it would not be correct that the data are of little value to demographers". [4]While demographers have developed special techniques for data adjustment and analysis, yields a rough notion of trends and differentials in these demographic events (Judson and Popoff 2004).

According to the researchers "vital statistics form the basis of fundamental demographic and epidemiologic measures". The data derived from civil registration systems as well as from the actual records of vital events (i.e. births, deaths, marriages, divorces, fetal deaths or abortions) is called "vital statistics"

Now if we want to know that how complete is the registration of births and deaths in our world today? Then one must differentiate developed countries from the developing ones. The UN International Children's Emergency Fund's (UNICEF) research center has estimated that there were around 50 million babies unregistered, which consists of more than two fifth of all the babies born in the year 2000 (UNICEF 2001). These unregistered children are found in the countries "where there is little awareness of the value of birth registration and where the registration network is inadequate, where there are no public campaigns, or where the cost of registration of children are prohibitive" (UNICEF, 2002:10). So, in general, most unregistered babies are born in the developing countries, largely because these countries are more likely to face political, administrative and economic hurdles in registration.

On the other hand, in some countries, gender discrimination and son preference also lead to female babies being excluded from the birth registration ([5]Hudson and den Boer, 2004). UNICEF has noted that in the year 2000, more than 70% births in the Sub-Saharan Africa

---

[3] Who counts?4, The way forward by Carla Abou Zahr and john Cleland
[4] New Directions in the Development of population estimates in the US by David A. Swanson and Jerome N. McKibben
[5] Missing women and Bare Branches gender balance and conflict by Hudson and den Boer, 2004

and 63% in Asia were unregistered. In only South Asia, there were an estimated 22.5 million unregistered births, which is the largest number among all the areas of the world.

However, this doesn't mean that all the developing nations have seriously incomplete birth registration, but many countries in the former Soviet Union have virtually universal records of births. This is because of their well-established birth registration systems, high-quality medical facilities and well-trained medical personnel's.

While, the other most important factor of vital event, is death record. The death certificates are usually filled out by funeral homes with personal information about the decedent provided by one or more of the surviving family members as well as by the physician in attendance at the death or by the coroner. In addition to deceased age which may be misreported by the family members particularly for the elderly, the certificates typically include occupation, residence place and the cause of death information.

1.4 Statistical Techniques:-

There are several methods for analyzing the data that are being used to find the demographic results by different researchers in different times. In most of the below mentioned methods, researchers have tried to find the [6]odds ratio or p-values to significantly accept or reject their hypothesis. If one wishes to know the association/correlation between the infant mortality and other socio economic variables such as poverty level, maternal age, education level, effect of smoking in pregnancy, maternity centers, place of origin, and migration etc. then the following methods might be the best approach to find the odds ratio or chi-square and/or p-value. The methods are as follows:

- Logistic regression (Multiple logistic regression)
- Kaplan-Meier survival test (Survival and Event History Analysis by Odd O. Aalen)
- Etiologic fractions (EF's)
- Two stage least square regression (with various explanatory variables)
- Chow's test (to compare the significance differences)
- Hazard logistic regression (Multivariate analysis)
- Generalized estimation equation (GEE)

- Chi-square test (Bio Statistics by Wayne W. Daniel)
- 2 x 2 table
- Meta-regression analysis
- Pooled adjusted and unadjusted odds ratio
- Dose response regression slopes
- Deviance goodness of fit
- Pearson goodness of fit
- Test of hypothesis for proportions (Probability and Statistical Inference by Hogg and Tans)
- Maximum likelihood estimation (MLE) (An introduction to generalized linear models by Dobson and Barnett)

Now let us have a close look on some of these methods.

1.4.1 Logistic regression:-

[7]Logistic regression (sometimes called the "logistic model or logit model") is used for the prediction of the probability of an event by fitting data to a logit function. This logistic regression is the same as generalized linear model that is used for binomial regression. It makes use of several of predictor variables that may be either numerical or categorical. In other words we can say that, this is a generalized linear model in which the outcome variables are measured on a binary scale. For example the responses may be present or absent, alive or dead. Moreover, the probability that a person can have a heart attack in a specific time can be predicted from the knowledge of the person's age, sex and body structure (health). Similarly, an infant mortality in a specific time period can be predicted from the knowledge of maternal age and education, maternity centers, poverty (income level), ruler and urban areas facilities and etc.

A categorical/binary variable has only two values such as "yes" or "no", " success" or "failure" rather than continuous. This logistic regression is used extensively in the fields of medical and social sciences as well as business marketing applications such as prediction of a customer's worth to purchase a product or cease a subscription. Since, logistic regression is the same as generalized linear model that is used for binomial regression but it is different

---

[6] Odds ratio = (P/1-P) means that odds=Success/failure
[7]An introduction to generalized linear models by Annette J. Dobson and Adrian G. Barnett, Wikipedia and different web sites.

from ordinary linear regression. The main differences of using logistic regression instead of ordinary linear regression are:

- If we use linear regression, the predicted values (estimated values) could become greater than one or less than zero, such values are theoretically inadmissible in logistic regression.
- One of the assumptions of regression is that the variance of Y is constant, independent of X (homoscedasticity). This cannot be the case with logistic/binary variables, because the variance is P.Q (Bernoulli distribution), when P=50%=0.50 of the people are 1s than variance will be (0.5) (0.5) = 0.25, its maximum value. Where p is the probability of success while q is the failure probability. As we go to more and more extreme values then the variance is decreases. But when P=0.10=10% than variance will be (0.1) (0.9) =0.09. It means, as P approaches to "1" or "0", the variance approaches to zero.
- In the significance testing, the regression coefficients rest upon the assumption that errors of prediction (Y-Y') are normally distributed. Because Y only takes the values 0 or 1. This assumption is pretty hard to justify. Therefore, the test of the regression coefficients is suspect if you use the linear regression with a binary (logistic) regression.

The logistic function is:

$$f(z) = \frac{e^z}{1+e^z}$$

Where Z = α + βX is the measure of total contribution of all the independent variables used in the model and is known as logit. Moreover, in the logistic regression, the dependent variable is logit, that is the natural log of odds:

$$\log(odds) = \log it(p) = \ln(\frac{p}{1-p})$$

$$\log it(p) = \alpha + \beta X$$

$$\ln(\frac{p}{1-p}) = \alpha + \beta X$$

$$\Rightarrow p = (\frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}})$$

1.4.2 Multiple Logistic Regression:-

[89]Multiple logistic regression is an extension of the logistic regression to the care where we have multiple explanatory variables. The basic idea is the same that the probability of one outcome is modeled as a function of the linear combination of several independent variables. A special case of multiple logistic regression is when the probability varies as a polynomial function of a single quantitative independent variable and this is the similar to the polynomial regression.

A case where the dependent variable can take any numerical value for a given set of explanatory variables, the multiple regression can be used. But in the case, where the dependent variable is qualitative (dichotomous, polytomous) the logistic regression will be used. In multiple regression the dependent variable is assumed to follow normal distribution but in the case of logistic regression the dependent variable follows Bernoulli distribution (if it is dichotomous), which means that it will be only 0 or 1.

Further, when there are two or more explanatory variables these may be:

- Continuous
- Discrete (nominal/ordinal)
- Both continuous and discrete (or mixed)

Multiple logistic regression model is the same as generalized linear model that has the random component/part binomially distributed (the response variable is a dichotomous variable) and the other component is linear predictor with more than 1 variable:

$$Odds = \frac{p}{1-p} = e^{\alpha+\beta_1 X_1+\beta_2 X_2+......+\beta_k X_k}$$

and a link which is the logit:

---

[8] Multiple logistic regression by Bret Larget, University of Misconsin-Madison.
[9] Multiple logistic regression for dichotomous responses by Carolyn J. Anderson, University of Illinois at Urbana Champaign.

$$\Rightarrow \log(odds) = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k$$

Similarly as before,

$$\Rightarrow p_i = \frac{e^{\alpha+\beta_1 X_1+\beta_2 X_2+......+\beta_k X_k}}{1+e^{\alpha+\beta_1 X_1+\beta_2 X_2+......+\beta_k X_k}}$$

1.4.3 Chi-Square test:-

[10]In statistics and probability theory, the chi-square distribution (also called chi-squared or χ^2-distribution) with K degrees of freedom is the distribution of a sum of square of k independent standard normal variables/random variables. It is a much used probability distribution in statistical inference, for example in hypothesis testing or in the construction of confidence intervals. The known situation in which the chi-square distribution is used are the common chi-square test for goodness of fit of an observed distribution to a theoretical one and of the independence of two different criteria of classification of qualitative data and the third use is the confidence interval estimation for a population standard deviation of a normal distribution from sample S.D.

If $Z_1 + Z_2 + ... + Z_k$ are the independent standard normal random variables, than the sum of their squares is:

$$X = \underset{\chi_1^2}{\underset{\uparrow}{Z_1^2}} + \underset{\chi_1^2}{\underset{\uparrow}{Z_2^2}} + ... + \underset{\chi_1^2}{\underset{\uparrow}{Z_k^2}}$$

$$\Rightarrow X = \sum_{i=1}^{k} Z_i^2$$

[10] Biostatistics (A foundation for analysis in the health science by Walne W. Daniel) + Web

Then X follows the chi-square distribution that has only one parameter k (a positive integer that is called the number of degrees of freedom). Where the number of Zi's shows the standard normal variables.

For the use of chi-square distribution in statistics, there is the chi-square test. A chi-square test is any statistical hypothesis test in which the sampling distribution of the test statistics is a chi-square distribution when the null hypothesis (H_0) is true, or any in which this is asymptotically true. This means that the sampling distribution (if H_0 is true) can be made to approximate a chi-square distribution as closely as desired by making the sample size (n) large enough.

The test statistic for the chi-square test is:

$$X^2 = \sum_{i=1}^{n} \frac{(o_i - e_i)^2}{e_i}$$

Where o_i is the observed frequency for the ith category of the variable of interest and e_i is an expected (theoretical) frequency asserted by the null hypothesis. Further, when the null hypothesis is true, this X^2 is approximately distributed as χ^2 with k-r degrees of freedom. Where k is the number of groups for which observed and expected frequencies are available and r is the number of restrictions or constraints imposed on the given comparison.

Moreover, the chi-square distribution is the continuous distribution that is based upon an underlying normal distribution. In the hypothesis testing, the chi-square test is also called the test for association and is used where it is obvious approximation, e.g. on testing the homogeneity of correlation coefficient. It has two similar uses, test of homogeneity of variance and of goodness of fit of observed continuous data to theoretical distribution.

2. Sources of Demographic Information:-

[11]There are many ways to collect the demographic data/information in which national census, household sample survey and registers are the most common ones. Also household surveys are among the major sources of demographic and social statistics in many countries. It has been recognized that population and housing censuses are conducted at long interval of about ten years while for most countries the administrative record system for social statistics are not well developed or is incomplete. On the other hand, household surveys, provide a convenient avenue for collection of detailed and varied socio-demographic data pertaining to conditions under which people live, well-being activities in which they engage, demographic characteristics and culture factors that influence behavior and economic and social change. However, this does not preclude the complementary use of data generated through household sample surveys with data from another sources such as national censuses or administrative records.

So, the three sources (national censuses, sample survey and the registration) of data are generally used to estimate demographic parameters.

2.1 National Censuses:-

As described (in chapter 1) a population census is the total process of collecting, compiling, evaluating and disseminating demographic, social and other data at a specified time covering all persons in a country. This national census is the major source of social statistics with its obvious advantage of providing data for small geographical units/researchers. A census is an ideal method for providing information on size, composition, and the distribution of the population as well as social-economic and demographic characteristics. In other words, we can say that the population censuses are the only feasible vehicle for obtaining small area statistics that are vital for local planning for social services such as the number and locations of schools and hospitals. So generally the population census collects information for each individual in the household and each set of living quarters.

[11] Main sources of socio-demographic statistics by Jeramiah P. Banda, UN Statistics Division

Moreover, there are two types of methods (de-facto and de-jure) for obtaining data and anyone can be adopted for census, depending on government policies. The de-facto approach counts the people where they are found at the time of census, while the de-jure approach counts people in the usual place of residence (where they formally lived either they are not available/absent at the time of enumeration).

2.2 Surveys:-

Sample surveys are the key source of data on social phenomena. They are one of the most flexible methods of data collection. Any kind of subject/issue can be investigated through these household surveys. In the household sample surveys, a part of the population is selected from which the data or information are collected and then inferences are made to the whole population. Because in the household sample surveys, there are similar workload for interviewers and a longer time period assigned to data collection and there are possibilities of veering most subject matters in greater detail than in censuses. Moreover, there is a scope for training field staff more intensively. In fact, not all the data needs of a country can be met through census taking, therefore, household sample surveys provide a mechanism for meeting the additional and emerging needs on a regular basis. The flexibility of sample surveys, therefore, makes them excellent choice for meeting data users needs for statistical information which otherwise would not be available and incomplete.

Most of countries have in place household sample survey programs, which include periodic and ad-hoc surveys and also it have been advisable that the surveys should be part of an integrated statistical data collection system of a country. Further, in the area of demographic and social statistic, intercensal household sample surveys can constitute part of this system and there are different types of sample surveys that can be applied to collect data on social and demographic statistics such as specialized surveys, multi-subject surveys, multiphase surveys and longitudinal surveys. The selection of any specific type of survey will depend on a number of factors including resources, subject matter requirements and finally the logistical considerations while a specialized survey covers single subject of issue such as nutrition or time-use. But the one greatest advantage of surveys is that they are less expensive and time consuming as compared to a census.

2.2.1 World Fertility Survey:-

[12]The world fertility survey (WFS) was one of the most important international actions in demographic data collection and analysis of the twentieth century (between 1973 and 1984). Around 66 countries of the world carried out comparable survey of population fertility and 44 of them were developing countries that participated in this world level survey and received a financial and technical assistance channeled through WFS's headquarters. Thus, an impressive geographical coverage was achieved through some of the largest countries (China, India, and Brazil etc.) declined to participate.

The idea for this international program is to resolve conflicting assessments of the fertility impact of family planning programs. The US Agency for International Development and United Nations Fund for population activities were the two dominant financial sponsors of World Fertility Survey (WFS). In this sample survey, data were gathered on reproductive behavior and related social psychological indicators from 40% of the world population. This cross-national fertility survey was the important source of fertility and related demographic information in the fields of statistics and demographic communities.

2.2.2 Demographic and Health Survey:-

[13]In 1984, the world fertility survey was followed by another coordinated international program of research, the Demographic and Health Survey (DHS) with more than 200 sample surveys carried out in 75 developing countries. DHS's are nationally representative household sample survey with a large sample size (usually around 5,000 to 30,000 households). These surveys provide data for many variables in the areas of fertility (reproductively), population, nutrition, health and etc. These surveys are typically conducted on every five years basis to permit comparisons over time. Where interim surveys are conducted between DHS rounds and have shorter questionnaires with smaller samples then the DHS surveys (2000 to 3000 households). Similar to world fertility survey (WFS), demographic and health survey (DHS) also provide the demographic information previously unknown about the countries in which they are implemented. Recently the DHS was completed in 2006 in a developing country Nepal.

2.2.3 Other Fertility Surveys:-

Every country has time to time involved in different kind of surveys to get know about the status of their country's population and fertility/mortality rates, as well as the overall progress on the other socio-economic factors. Moreover, to get know about the rural and urban population and their reproductive behavior. The number of these demographic surveys has grown gradually over the years and among these surveys, some of them are conducted by the government level, while in the developing countries most of them are conducted with the support of Non-Government Organizations (NGO's) e.g. United Nations, World Health Organization, etc.

Another aspect of conducting these surveys is to provide researchers, educators, students, policy makers and others with a data resource to examine issues related to families and fertility. Further, these surveys/projects are the continuing mission of the demographic and behavioral sciences or social sciences of the organization/department of the government to research on fertility, the family and family planning.

2.3 Registers:-

[14]The registers (also called population registers) is a mechanism for the continuous recording of selected information pertaining to each and every member of the resident population of a country or territory, making it possible to determine up-to-date information about the size and characteristics of the population at selected points in time. Because of the nature of the register, its organization as well as its operation should have a legal basis. Population registers start with a base consisting of an inventory of the inhabitants of an area and their characteristics, such as date of birth, place of birth, sex, place of residence, marital status, language and citizenship. Moreover, to assist in locating a record for a particular person or family, an identification number is provided. The population register can also contain other socio-economic data, such as education or occupation etc. But, the population registers are

---

[12] "Population and Society (An Introduction to Demography) by Dudley L. Poston, JR and Leon F. Bouvier" + Web
[13] "Population and Society (An Introduction to Demography) by Dudley L. Poston, JR and Leon F. Bouvier"
[14] United Nations Population Division (Population Registers in Demographic and social statistics).

updated by the four events i.e. births, deaths, marriages and divorces records, which is the part of civil registration system (also called vital statistics records) of the country.

2.4 Sources of demographic data in Pakistan:-

[15][16][17][18]Pakistan, officially called the Islamic Republic of Pakistan is a country in South Asia. It has a 1046-kilometers coastline along the Arabian Sea and the Gulf of Oman. It has common borders with Afghanistan in north-west, Iran in the west, India in the east and China in the north. It has four provinces named Balochistan, Khyber Pakhtunkhwa, Punjab, and Sindh. Moreover, it also includes two centrally administrated areas, i.e. Federally Administered Tribal Area (FATA) and Azad Jammu and Kashmir (AJK), one territory Gilgit-Baltistan and one capital territory Islamabad.

If we look at the demographic situation of Pakistan, then according to the official population clock of Pakistan, an estimated population in January 2011 is 175.06 million, which is the 6th most populous country in the world with having 2.49% share in the world's population. Pakistan is a developing country in the world and around 20% of the population lives below the international poverty line of US$ 1.25 a day. The life expectancy at birth is 62 years for male while 63 years for female ([19]PDS 2007). The expenditure on health is 2% of the GDP. The mortality rate under age 5 is 97 per 1000 live births, whereas infant mortality rate is 75 per 1000 live births and the maternal mortality stands at 254 cases per 100,000 women.

Statistics Division is the central department of Government of Pakistan that frames policies for the development of statistical services in the country. It provides solid statistical base to national and international planners, researchers, students, policy makers and other data users in various socio-economic sectors. It has three attached departments.

- Federal Bureau of Statistics-FBS
- Population Census Organization-PCO
- Agriculture Census Organization-ACO

---

[15] WHO Pakistan pages
[16] United Nations Pakistan pages
[17] Statistics Division, Government of Pakistan
[18] National Institute of Population Studies, Pakistan
[19] Pakistan demographic survey 2007

2.4.1 National Census:-

As described earlier, a census is the procedure of systematically acquiring and recording information about the inhabitants of a given population in a country or territory and it is a regularly occurring and official count of a particular population after a specific time period.

Similarly, the census of Pakistan is conducted decennially across the entire country. After the political independence in 1947, the first census was conducted in 1951 and it was decreed that the census have to be carried out once in every 10 years. Than, the second national census was conducted in the year 1961, whereas the third census was conducted in 1972, one year due because of war with India. The next country wide population estimate was carried out in 1981, while the 5th national census was conducted in 1998 (7 years late), due to the political issues. Pakistan's 6th national census was scheduled in October 2008, but it too has been delayed in 2010 due to political issues and is now planned in March/April 2011.

In the first census 1951, Pakistan's population was 34 million, while in 1998 census it becomes 132.35 million and according to Pakistan Census Organization (PCO) estimates, in January 2011 it has crossed 175 million figures.

The following table gives us some overview of Pakistan's population by the population census records:

Table 2.1:

Population Censuses of Pakistan						
			Census Year			***(000)***
Province/Region	↓ **1951**	↓ **1961**	↓ **1972**	↓ **1981**	↓ **1998**	**2010-11 estimates**
Pakistan	33,740	42,880	65,309	84,254	132,352	175.06
Punjab	20,541	25,464	37,607	47,292	73,621	million
Sindh	6,048	8,367	14,156	19,029	30,440	
Khyb.P.Khaw.	4,557	5,731	8,389	11,061	17,744	
Balochistan	1,167	1,353	2,429	4,332	6,566	
FATA	1,332	1,847	2,491	2,199	3,176	
Islamabad	96	118	238	340	805	

FATA: Federally Administrative Tribal Areas, Islamabad: Capital Territory
Source: Pakistan Census Organization (PCO), WHO Report 2010

2.4.2 Pakistan Demographic Survey (PDS):-

[20]Population census and the civil registration system (birth and death registration system) are the traditional sources of vital statistics. As several other developing countries in the world, the civil registration system in Pakistan appears to be highly deficient and inadequate to provide reliable birth and death statistics. The inability of decennial censuses and in the absence of efficient civil registration system, to provide birth and death statistics during the intercensal period, several demographic surveys have been undertaken by the Federal Bureau of Statistics, Pakistan in the country since early sixties either independently or in collaboration with other organizations. Pakistan Demographic Survey (PDS) is the latest series of demographic surveys that was launched in 1984.

The main objectives of this Pakistan Demographic Survey is to collect birth and deaths statistics in order to arrive at various measures of fertility and mortality for Pakistan and its

[20] Federal Bureau of Statistics Pakistan web and Pakistan Demographic Survey.

four provinces, separately for urban and rural areas, as well as to estimate current rate of natural increase/decrease of population at national and provincial levels. Further, in most of the PDS surveys, the coverage of population was done by the [21]de jure approach.

2.4.3 National Database and Registration Authority (NADRA):-

[22]National Database and Registration Authority (NADRA) is a Federal department of Government of Pakistan that is responsible for issuing the national identification cards to the citizens of Pakistan with having more than 11,000 technical and management personnel and over 400 domestic offices with 5 international offices. NADRA is one of the largest organizations in the country. National Database and Registration Authority (NADRA) was established as National Database Organization (NDO) an attached department under the Ministry of Interior in 1998. On March 2000, National Database Organization (NDO) and Directorate General of Registration (DGR) merged to form NADRA, an independent corporate body with requisite autonomy to operate independently and facilitate good governance. In the same year 2000, NADRA launched the Multi-Biometric National Identity Card project developed in conformance with international security documentation issuance practices.

In other words, one can say that NADRA set out on the journey of civil registration of all Pakistanis in 2000 and in a short span of time NADRA created a state of the art centralized Data Warehouse, Network infrastructure and Interactive Data Acquisition System to issue secure Computerized National Identity Cards (CNIC). The introduction of this new fool-proof comprehensive and highly sophisticated computerized system helps in reducing the identity theft. Now, NADRA's National Data Warehouse is one of the largest centralized data base in the world that hosts the data of over 96 million citizens and runs various Transaction Processing.

Finally, by providing solutions for identification to citizens, e-governance and secure documents that deliver multi-pronged goals of mitigating identity theft, safe-guarding the interest of public, NADRA becomes the largest biometric citizen database in the world.

[21] A person who usually live in that area whether he is present or absent at the time of enumeration
[22] National Database and Registration Authority (NADRA) + web

2.4.4 Civil Registration Management System (CRMS):-

[23]The Civil Registration Management System (CRMS) has been developed by National Database and Registration Authority (NADRA) for the registration of four vital events. i. e. births, deaths, marriages and divorces. The purpose of this CRMS project is to automate all the local governments in a country and provide computerized registration and certificate issuance of four vital events. The CRMS system provides up to date status reporting facilities for selected vital events, while providing business decision models for strategic decision making. NADRA has successfully developed and implemented CRMS in Pakistan at the grass-roots level, an online web based connectivity for provisioning of monitoring facilities to designed Government officials.

[23] National Database and Registration Authority (NADRA), Pakistan

3. Fertility:-

Fertility is the natural capability of giving birth or refers to the actual production of children, which in the strictest senses a biological process. There are three main concepts of fertility. Fertility is the actual production of male and female births and refers to a natural behavior, the second one is the reproduction that is also an actual production but refers to the production of female births only. While the third concept of fertility is fecundity, that refers to the potential or the biological capability of producing live births.

3.1 Reproductive measures:-

[24]As a measure “fertility rate” is the number of children born per family/couple, person or population. Whereas crude birth rate (CBR) is the cross-sectional (i.e. period) measure and refers to the number of births occurring in a population in a year per 1000 people.

$$\text{CBR} = \frac{number\ of\ births}{midyear\ population} * 1000$$

Moreover, the general fertility rate (GFR) is another cross-sectional measure of fertility which is superior to CBR because it restricts the denominator to women of childbearing age 15-49 years.

$$GFR = \frac{number\ of\ births}{midyear\ population\ (F{:}\ 15-49)} * 1000$$

If one wishes to calculate GFR but have available data of CBR only, than an estimated value (for women 15-44 year) is given by the formula:

$$GFR = CBR * 4.5$$

While the total fertility rate (TFR) (also called fertility rate, period fertility rate or total period fertility rate) of a population is calculated as the sum of age-specific fertility rate

[24] Fertility transitions in developing countries: progress or stagnation by John Bongaarts

(ASFR) of women in five-year age interval on total multiplied by five and this rate estimates the number of children a cohort of 1,000 women would bear if they all went through their childbearing age years exposed to the same age-specific fertility rate in effect for a particular time.

$$TFR = (\Sigma ASFR) * 5$$

According to the surveys, the world total fertility rate is 2.5 (2009) whereas the world crude birth rate stands at 20.3 per 1000 population for the year 2005-10. In the past 2-3 decades, fertility has declined at a rapid pace in most of the developing countries and the total fertility rate of the developing countries dropped from 6.0 births per women in 1960s to 2.9 births in 2000-2005 (UN 2007).

[25]Pakistan as the 6th most populous country in the world has shown a slow decline in fertility despite limited efforts has been made for improvement and stands at 3.28 births per women. However, the fertility decline in Pakistan has engendered much interest since it has been suggested that its fertility transition is controversial in certain important respects, while demographically the fertility behavior is changing over time, not only in the urban population but also the rural area's population. Numerous fertility surveys give us an idea about that the Total Fertility Rate in Pakistan in different decades.

---

[25] Levels of recent and targeted determinants of fertility in Pakistan by Jamal A. Nasir

Table 3.1:

Estimates of Total Fertility Rate in Pakistan		
Decade	***Source***	***TFR***
1980s	Pakistan Contraceptive Prevalence Survey (1984-85)	6.0
	Pakistan Demographic Survey (1984-88)	6.9
	Pakistan Demographic and Health Survey (1990-91)	5.4
1990s	Pakistan Demographic Survey (1992)	5.8
	Pakistan Contraceptive Prevalence Survey (1994-95)	5.6
	Pakistan Fertility and Family Planning Survey (1996-97)	5.3
2000s	Pakistan Reproductive Health and Family Planning Survey (2000-01)	4.8
	Pakistan Demographic Survey (2001)	4.1
	Pakistan Demographic Survey (2003)	3.9
	Pakistan Demographic Survey (2005)	3.8
	Pakistan Demographic Survey (2006)	3.7
	Pakistan Demographic Survey (2007)	3.7
2010	Estimated	3.28

Source: Federal Bureau of Statistics Pakistan, World Fact book by CIA, Country profile by WHO

3.2 Age-specific fertility:-

[26]Age-specific fertility rate refers to the number of births to women in a particular age category in a particular year compared to the number of women in that age category. In other words, age-specific fertility rate is defined as the number of births by age of mother per 1000 females in the same age group. Age-specific fertility rate is, a way to measure fertility trends by different age groups and usually expressed as births per women or births per 1000 women in the age category.

The following table shows the age-specific fertility rate (per 1000 women) by urban and rural areas in Pakistan.

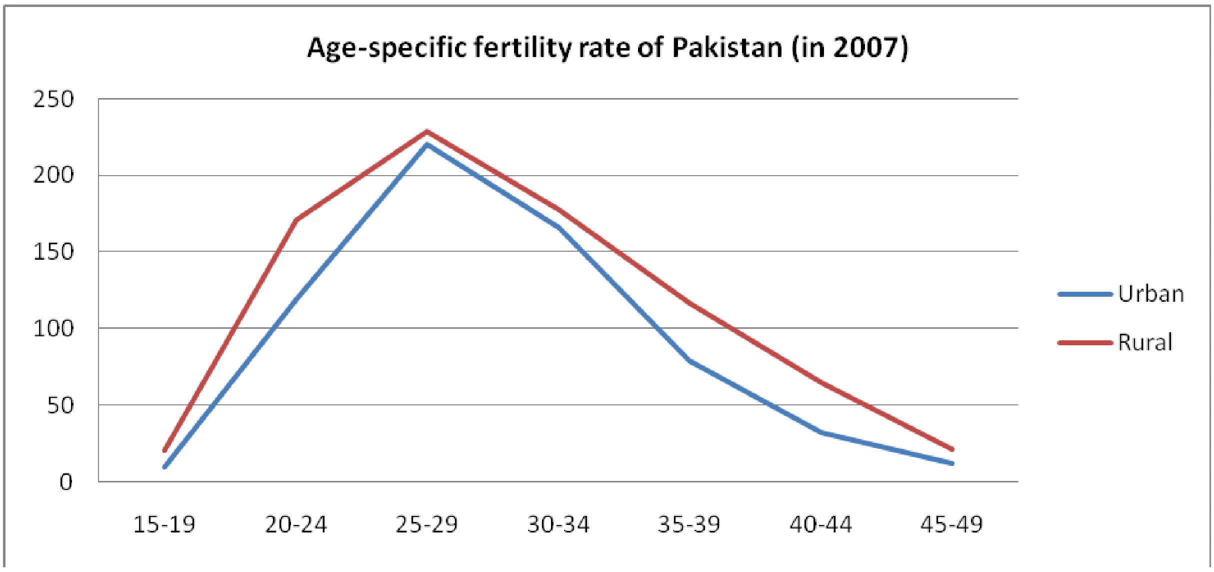
[26] Pakistan demographic survey 2007 + Google search

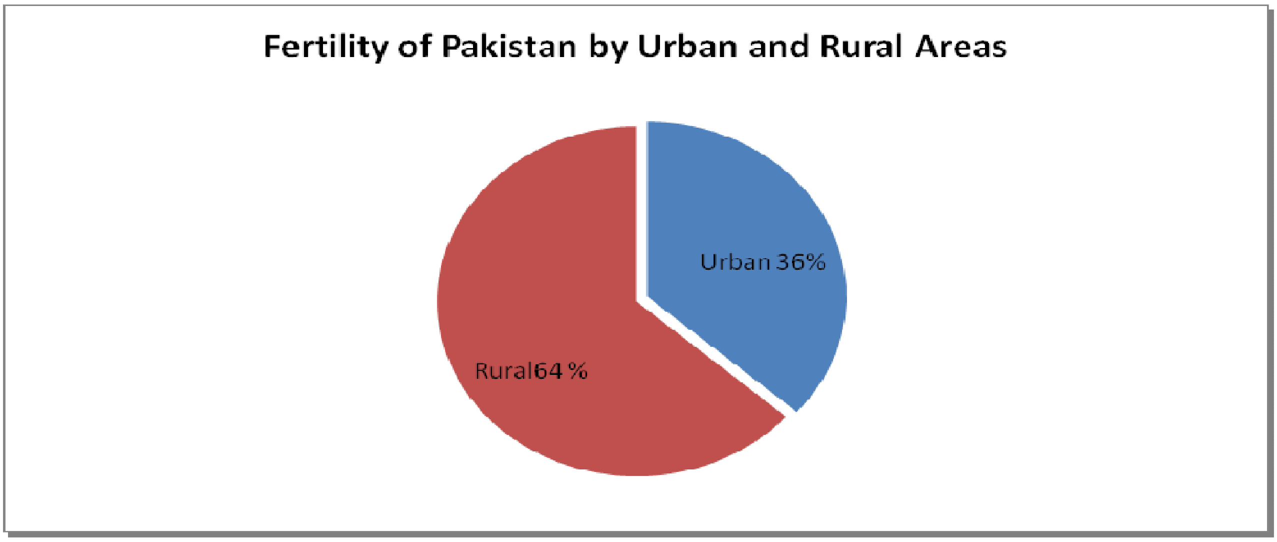
Table 3.2:

Age-specific fertility rate by Urban and Rural areas of Pakistan in 2007			
Age Group	**All Areas**	**Urban**	**Rural**
15-19	16.1	9.2	20.5
20-24	150.0	119.0	170.6
25-29	225.3	219.9	228.4
30-34	173.3	166.1	177.2
35-39	102.6	79.2	116.4
40-44	52.6	32.3	64.5
45-49	18.0	12.4	21.6

Source: Federal Bureau of statistics, Pakistan,

Here we can see that the age group 25-29 is the most fertile time whereas the age 45-49 is the least and minimum fertile period for the Pakistani women. Further, the figures indicate that the urban fertility is quite lower than the rural one, as the awareness about family planning is less in the rural areas of Pakistan. Moreover, rural areas population consist of around 64% and urban is on 36%, so it obvious to have higher fertility rate in rural areas.




Fertility of Pakistan by Urban and Rural Areas
Urban 36%
Rural64 %

4. Migration:-

[27]According to the demography (from chapter 1), we know that there are three factors that change the population (either by adding or subtracting members), fertility, mortality and migration. So, the third one, migration factor plays also an important role in the population change. There are two types of migration; internal and external.

4.1 Internal and External Migration:-

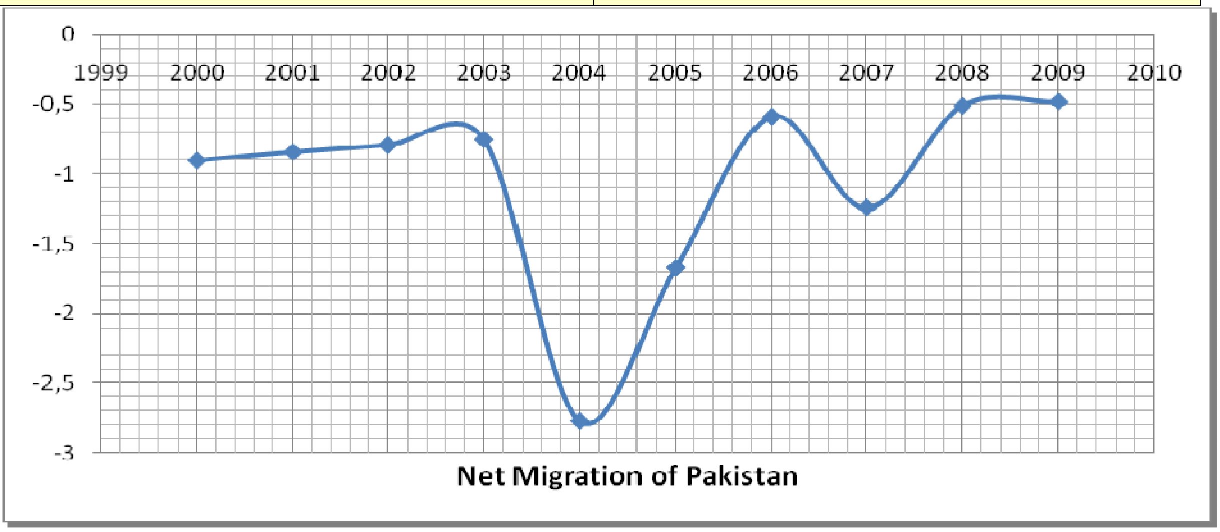
Internal migration is the change of temporary or permanent residence within a country involving a geographical move. Any person whose residential move involves the crossing of a political boundary means moving from one country to another country is called external migration or migrant. Migration is a physical movement by human from one place to another, sometimes over long distances or in large groups. This movement of population has continued under the form of both voluntary migration within one area, region or country and involuntary migration (which includes the slave trade, trafficking in human beings etc).

According to the world migration report 2010 (by the International Organization of Migration), in 2010 the number of international migrants was estimated at 214 million and if this number continues to grow at the same pace as during the last 2 decades, it could be 405 million by 2050. [28]This entry includes the figures for the difference between the number of persons entering and leaving the country during a year per 1000 people (based on the mid year population). Mostly the developing countries have the external migration and this can be in the form of education, job, visit, temporary or permanent residence. An excess of person leaving the country is referred to as net emigration (e.g. -2.26 migrants per 1000 people) and an excess of persons entering the country is referred as a net immigration (e.g. +1.46 migrants per 1000 people). So the net migration rate indicates that the contribution of migration to the overall population change. High levels of migration can cause problems such as increasing unemployment and potential ethnic strife while the high level of emigration also cause problems as the reduction in labor force and perhaps in certain key factors.

If we consider Pakistan, like many other developing countries, there is a lack of data arrangements about the migration and I have not found any relevant data from Pakistan. But according to the CIA fact book, the following table gives us the overview of Pakistan's net migration in the last few years.

Table 4.1:

Net migration rate of Pakistan	
Year	**Net Migration Rate**
2000	-0.90
2001	-0.84
2002	-0.79
2003	-0.75
2004	-2.77
2005	-1.67
2006	-0.59
2007	-1.24
2008	-0.51
2009	-0.48



27 Population and society, by Dudley L. Poston. JR.and Leon F. Bouvier + Internet
28 CIA, World fact book

5. Mortality:-

[29]Mortality is the condition of being mortal or susceptible to death. Mortality is another main concept of demography, as described earlier (in ch.1).

5.1 Infant Mortality:-

A few, if any human experiences are more tragic or emotionally devastating than the death of an infant or child. Infant mortality receives special consideration from the demographers. The infant is a child of less than one year of age. The infant mortality rate is an important indicator of nation's socio economic welfare, culture factors, status of hygiene and availability & utilization of medical services. Tremendous decline in infant mortality took place since 1990s but despite this reduction, the infant mortality rate (IMR) is still high especially in less developed countries.

The infant mortality rate (IMR) expressed as a rate per 1,000 live births, of a child born in a specified year dying before reaching the age of one. According to the Millennium Development Goals (MDG's) indicators collected by United Nations (UN), IMR in 2007 was 47 at world level whereas this rate was 5 for developed countries and 51 for developing countries (Millennium Development Goals Indicators 2009). Therefore, reducing health inequalities is not just about saving lives but also about improving life chances. Infant mortality is a sensitive measure of overall health of a population, providing an important measure of the well being of infant, children and pregnant women.

According to the figures (United Nations, World Health Organization and UNICEF reports), 4 million [30]neonatal deaths occur every year with 99% of neonatal deaths occur in low and middle-income countries and 38% of deaths in children < 5 year occur in the first month of life,. In low income countries, one out of 10 children dies before the age of 5, while in the wealthier nations, this number is only one out of 143. Today nearly 11 million children under the age of 5 die in our world every year, while over 1200 every hour most from easily preventable or treatable causes. Moreover, 60 to 80 percent of neonatal deaths occur in

[29] United Nations, WHO and UNICEF reports
[30] Neo-natal is 0-6 days of live

babies born with low birth weight, and approximately 30% of children are underweight from Africa and over 50% from south Asia. Further, an estimated 585,000 women who die each year in childbirth.

Under-five mortality also remains very high in Southern Asia, about one in 14 children died before age of five (2008 records) and the progress is too slow to meet the Millennium Development Goals 2015th target. Each death is a devastating loss for the family and we need to do all we can to prevent these tragic losses. Over the past few years, some of the key building blocks for infant mortality targets have been put in place by the Southern Asian Governments. The Governments make tackling health inequalities as a priority.

5.2 Infant mortality in Pakistan: -

Pakistan, a developing country, where a lot of things are lacking including quality healthcare and greater unawareness of health issue, etc. Even small steps like focusing on and promoting hygiene and sanitation and also stressing special breast-feeding can lead to improved health but these are put on the backburner. Being a member country of Millennium Development Goals (MDG’s by UN), MDG-4 stipulates that Pakistan must reduce its under-five child mortality rate by two thirds between 1990 and 2015, i.e. from 130 deaths per 1,000 children in 1990 to 85 per 1,000 in 2015 and according to Pakistan demographic and health survey, this rate was 94 per 1000 children in 2006-07. Further, MDG-4 also states that Pakistan’s infant mortality rate (IMR) should be reduced from 100 per 1,000 children in 1990 to 40 per 1,000 in 2015, and 2006-07 records states that IMR was 78 per 1000 children (PDHS 2006-07). According to 2009-10 estimates, Pakistan’s IMR is standing at 70.5. The following table gives us information about the Pakistan’s position in the infant mortality along with the world trends.

Table 5.1:

World Trends in Infant Mortality from 1960s to 1999s				
	High income countries	**Middle income countries**	**Low income countries**	**Pakistan**
1960	35	120	167	149.7*
1970	26	94	140	135.7
1980	13	69	117	116.9
1990	8	46	97	100.7
1995	7	39	89	94.3
1999	6	31	77	84.8**

*year 1965, **year 2000
Source: infant and child mortality in Andhra Pradesh by edoardo Masset and Howard White

Even Pakistan is considered as the low to middle income country, but in the health care sector, it is considered far behind from the middle and high income countries. Further, the majority of Pakistan's population is living in the rural areas and it is considered as the rural country, where only the 35% of the population is residing in the urban areas and this ratio is 65% for the rural and the rural areas are the one who have worst situation in Infant mortality. The following table presents the infant mortality rate by rural and urban areas in the recent years. It states that the in the year 1989-94(average) infant mortality rate was 102.8 whereas after 20 year it becomes 70.5 (2009), which shows 31% decrease in 20 years. [31]But being a member of Millennium Development Goal (MDG) countries, MDG 4 states that Pakistan should make efforts to reduce its infant mortality rate by 78 in 2006-07 to 40 by 2015 (Pakistan demographic and health survey, talking the MDGs by USAID, Pakistan).

[31] MDG 4: Reduce child mortality, Millennium Development Goals by United Nations

Table 5.2:

Infant Mortality Rate in Pakistan			
Surveys	**All Areas**	**Urban**	**Rural**
1989-94 average	102.8		
1995-97 average	88.2		
1999-01 average	79.5		
2000	79.8	69.3	86.8
2001	77.0	68.9	80.9
2003	76.2	67.2	80.6
2005	76.7	67.1	81.2
2006	76.1	66.4	80.7
2007	75.2	66.5	79.4
2009 estimates	70.5		

Source: Different Pakistan Demographic Surveys (PDS), Google estimates

According to the Pakistan demographic survey, in the year 2000 there was 86.8 infant mortality rate in the rural area whereas this was only 69.3 in the urban, which shows 20% higher mortality in the rural areas. Similarly, in the year 2007 these figures becomes 79.4 and 66.5 in rural and urban areas respectively, still 16% higher. [32]In fact there are several biological factors that present their role in infant mortality like the multiple births, mother's age, length of preceding birth interval and the sex of child in rural and urban areas. The higher mortality of children from the multiple births especially during the neonatal period is common in the developing countries and the complications at birth and low birth weight are considered among the most important determinants of higher risk of death of twins ([33]Sullivan 1994). The researches show that there is a u-shaped relationship between mortality and mother's age at birth which means that the infant mortality tends to be higher for children born to young mothers and lower to older age mothers and this kind of effect can be found to be large in the developing countries like Pakistan, where the women are giving birth to children at a very young age especially in the rural areas. Generally, the children born

[32] Infant and child mortality in Andhra Pradesh by Edoardo Masset and Howard White

[33] Low birth weight in Tennessee 1994-2004 by Sullivan Epidemiology and Evaluation department

after a very short interval to the previous, presents higher mortality rates. According to the FBS Pakistan, there were 82,299 infant mortality cases found in the urban areas while this amount was 205,893 for the rural areas.

When discussing the infant mortality situation in rural and urban areas, one should not forget the other factors like environmental, behavioral and socio-economic factors that play their role. Acute respiratory infections and diarrhea are the most common and important causes of infant and child mortality in Pakistan and worldwide (IIPS, 2000). Water, air, food and fingers are the principal ways through which respiratory and intestinal diseases are transmitted. According to a research for the analysis of child survival in developing countries says that the physical environment to which children are exposed is likely to have an influence on mortality (Mosley and Chen, 1984). Further according to UNICEF (1999), 73% of death of children under five is a consequence of low cost treatable diseases.

The socio-economic factor includes the mother's education and economic well being. The correlation between mother's education and infant mortality is well documented in a large number of researches and for various countries. Children of illiterate mother have a much higher probability of dying at all ages then the educated mother (Sullivan 1994). Moreover, education can also be highly correlated with the income, but as an explanatory variable. With the overview of all these factors that can be a cause of infant mortality, let's have a look on age specific mortality situation in Pakistan and then we may have an idea to compare these urban and rural areas.

5.3 Age-specific Mortality:-

The United Nations records indicate that the life expectancy at birth is 62.7 and 63.1 for male and female respectively in Pakistan. [3435]In Pakistan, approximately one million deaths occur every year due to natural or different causes whereas 36% death are the under age 5. Moreover, the mortality is higher in males than the females in each age group. According to the Pakistan demographic survey (PDS-2007) the age groups 00-04 and 70-74 having the higher mortality. The figures indicate that the rural areas have more than double deaths compared to urban one. We can also see that the age group 00-04 is the most crucial period

---

[34] Pakistan demographic survey 2007

for surviving whereas the age group 30-34 is the least one in both areas of Pakistan, especially in the rural areas. Moreover, facts represents that the rural areas have not only more deaths but also the higher infant mortality ratio. The following chart illustrates the age specific mortality in Pakistan.

Table 5.3:

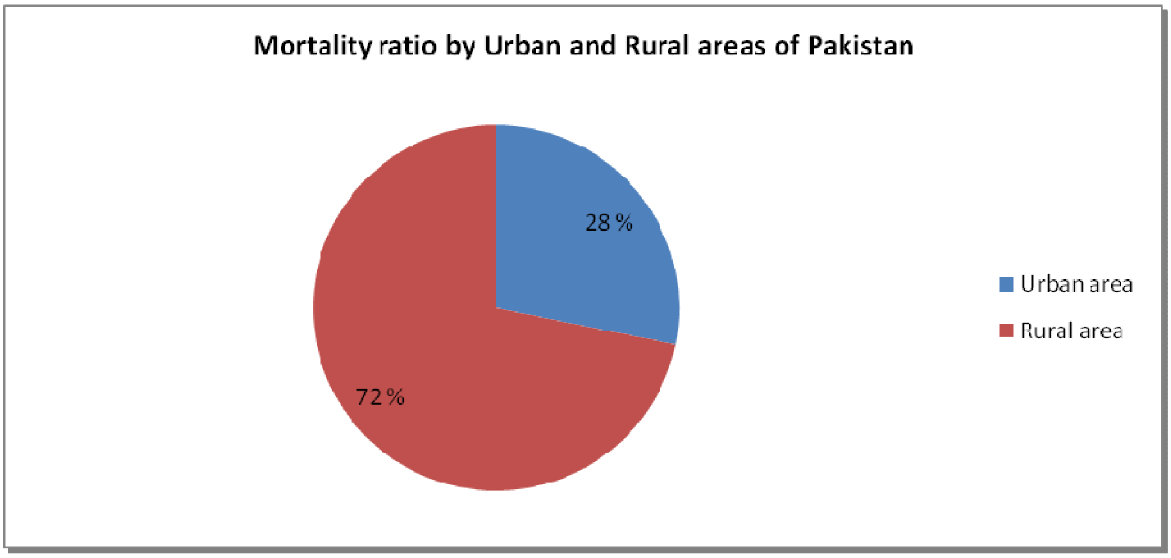
Age-specific Mortality by Urban and Rural Areas in Pakistan

Age at death	All areas	Urban area	Rural area
All ages	1019533	287516	732017
00-04	365729	99064	266665
05-09	33029	5458	27571
10-14	20510	5825	14685
15-19	27417	5551	21867
20-24	20192	7205	12987
25-29	23990	5700	18290
30-34	11624	4571	7053
35-39	24880	6877	18003
40-44	30926	9089	21837
45-49	27537	6219	21318
50-54	43405	16172	27234
55-59	51031	15100	35931
60-64	68531	20888	47643
65-69	58286	18410	39876
70-74	82325	24391	57934
75-79	37357	8951	28405
80-84	34079	10352	23748
85 & above	58685	17713	40972

Source: Pakistan Demographic Surveys (PDS 2007) by federal bureau of statistics, Pakistan

[35] World mortality report 2005 by United Nations

Table 5.3 indicates that in each age group the number of dying people is higher in the rural areas then the urban one. Which clearly point out that there is a difference of health care and the above mentioned factors/facilities in rural than the urban. The following pie chart may also helpful to be aware of the overall mortality ratio in rural and urban areas, while due to the larger population in the rural areas that is why having the higher mortality. i.e. 72% in the rural areas of Pakistan while this share stands at 28% for the urban population. This hypothesis is confirmed by a test of difference of distributions or the proportions of two different populations that have certain characteristic.



In our case, $\hat{p}_1$ and $\hat{p}_2$ denotes the number of infant deaths in the urban & rural areas divided by the total number of births in the urban & rural areas. Moreover, I shall treat this data like a 2 x 2 contingency table to find out the chi square and p-value. A probability value (p-value) will guide us about the significance of the test.

5.4 Determinants of infant mortality:

The following table 5.4 shows the data about the number of infant mortality case, the number of surviving infants and the total number of births in both areas of Pakistan (PDS-2007, PDS-2001).

Table 5.4:

Number of surviving infants and infant deaths in Pakistan			
2007	**Urban**	**Rural**	**Total**
# of surviving infants	1155754	2387027	3542781
# of infant deaths	82299	205893	288192
	1238053	**2592920**	**3830973**
2001	**Urban**	**Rural**	**Total**
# of surviving infants	1110005	2323080	3433085
# of infant deaths	82185	204424	286609
	1192190	**2527504**	**3719694**

The data presents that there was 3,830,973 total births in the year 2007 from which 2,592,920 births in the rural population that stands at 68% of the total births. This scenario was the same in 2001 as well where the total births were 3,719,694 whereas the rural population's births were 2,527,504, i.e. same 68% birth from the total births.

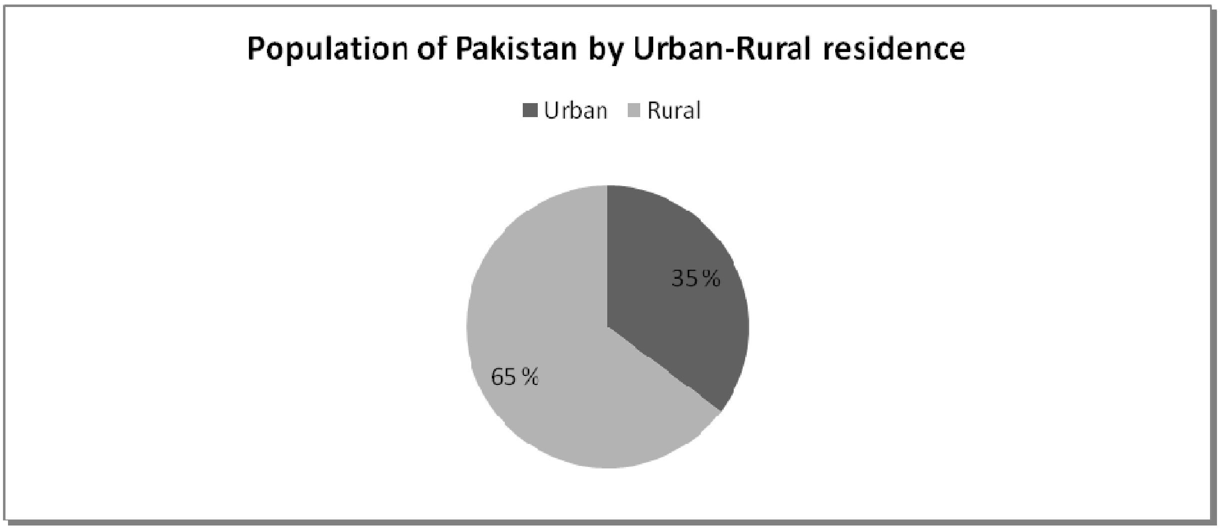
Now, the next table illustrates the results of the test of hypothesis for two proportions/population of rural and urban and a 2 x 2 contingency table on our null hypothesis: that there is a difference of facilities/services given by the state to urban and rural areas.

Table 5.5:

Test of hypothesis for two proportions and 2 x 2 contingency table				
	Z-value	**Tabulated value**	**Chi-square**	**p-value**
Infant mortality comparison in urban and rural areas in the year 2007	-46.21	-1.645	2014	0.0000001
Infant mortality comparison in urban and rural areas in the year 2001	-41.59	-1.645	1625	0.0000001

According to the test of hypothesis for two populations by using table 5.6, the z-calculated value shows a very big difference from the tabulated value, which clearly indicates that our hypothesis is accepted that there is a difference of facilities in the rural and urban areas that's why there is a difference in mortality rates as well specially in infant mortality. Further, the results of 2 x 2 table, the chi-square and p-value in both of the years 2007 and 2001 indicates that our hypotheses is extremely statistically significant and also states that there is a difference of policies/facilities in the rural areas then the urban one. A p-value of less than 0.05 shows the test is statistically significant but in our case, it's much less then 0.05, which presents a highly significant test.

The data shows that in the year of 2001, there was a huge difference in the rural areas from the urban one and then after 6 years period (in 2007) it remains the same. This presents that there wasn't much difference happens in a six year period. So the need is really to upgrade the rural areas as much as possible as almost 65% of population is residing there. This is the responsibility of the government to start effective rural policies at the earliest according to the above mentioned factors especially the health care sector and not only start but really to implement them.



Finally the following table 5.6 will give us some calculations of the above mentioned test. The remaining calculations can be found in the similar way.

Table 5.6:

Calculations of table 5.5	
$H_0 : p_1 > p_2$, $H_1 : p_1 < p_2$, $Sigfic.level = 95\%$	$Z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$
Where: $\hat{p} = \sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$\hat{p}_1 = \dfrac{82299}{1238053} = 0.06647$, $\hat{p}_2 = \dfrac{205893}{2592920} = 0.0794$ (for 2007) Number of infant deaths in urban/rural areas divided by the total number of births in the urban/rural
$n_1 = 1238053$ & $n_2 = 2592920$ *for 2007*	$\hat{p}_1 = \dfrac{82185}{1192190} = 0.0689$, $\hat{p}_2 = \dfrac{204424}{2527504} = 0.0809$ (for 2001) Number of infant deaths in urban/rural areas divided by the total number of births in the urban/rural
$n_1 = 1192192$ & $n_2 = 2527504$ *for 2001*	$\Rightarrow Z = \dfrac{0.06647 - 0.0794}{\sqrt{\left(\dfrac{0.06647(1-0.06647)}{1238053} + \dfrac{0.0794(1-0.0794)}{2592920}\right)}} = -46.21$
Source: Test of hypothesis for proportions from the book Probability and Statistical Inference by Hogg and Tans 2x2 contingency table for the Chi-square and p-value	

Our null hypothesis states that there is a huge difference of policies between rural and urban population's areas whereas the alternative is the opposite. On the 95% significance level, a test of hypothesis for two proportions has been applied on it to check whether our hypothesis is true or false. Then the results show that there is really a huge difference of policies between the two different areas populations and the need is to look after the both areas at the same level by the state.

6. Life Table:-

[36][37]A life table, mortality table or actuarial table is a statistical model for measuring the mortality (or any other type of 'exit' experiences) of a population, controlling for age distributions is called a life table. A life table includes:

- the probability of surviving any particular age
- remaining life expectancy for people at different age
- the proportion of the original birth cohort still alive
- Estimates of a cohort's longevity characteristics.

The life table is one of the important devices used in statistics especially in the fields of demography and it is a table that displays various pieces of information about the dying out of a birth cohort. There are different types of life tables, like:

- Current/Period vs. Generation/Cohort
- Complete vs. Abridged
- Single vs. Multiple decrement
- Increment/Decrement tables

Life tables are usually constructed separately for men and women because of their different mortality rates. The first column of the life table is invariably "age" and the remaining columns tabulate age-related functions pertaining to mortality, such as the number of survivors to various ages, deaths in a particular age interval, probabilities of death in various age intervals, age-specific death rates and so on. Life table is one way of summarizing a cohort's mortality experience, other ways, for example, are in the form of mathematical functions, or in the geographical form.

The most frequent used column of a life table is the expectation of life at age x or life expectancy at age x, is denoted by "e_x^0" and described as the average number of

[36] Demography: measuring and modeling population processes by Samuel H. Preston.
[37] University of North Carolina at Chapel Hill report on life table

additional years that a survivor to age x will live beyond that age, while another important column is

" ${}_{n}a_{x}$ " that refers to the average number of person-year lived in the interval x to x + n by those dying in the interval. Finally some functions of the life table (l_x, T_x, e^0x) refers to the single (exact) age, while other functions (${}_{n}d_{x}, {}_{n}p_{x}, {}_{n}q_{x}, {}_{n}m_{x}, {}_{n}a_{x}$) refers to the age interval that begin with exact age x and extended for exactly n years.

The main functions that are used to construct a life table with their mathematical expressions are as follows:

Table 6.1:

Mathematical formulas for life table

X	Exact age	
l_x	No. of people left alive at age x	: $l_{x+n} = l_x -_n d_x, l_0 base$
${}_{n}d_{x}$	No. of people dying between ages x and x + n	: ${}_{n}d_{x} = l_x - l_{x+n}$ **OR** ${}_{n}d_{x} = l_x *_n q_x$
${}_{n}q_{x}$	Probability of dying between ages x and x + n	: ${}_{n}q_{x} =_n d_x / l_x$ **OR** ${}_{n}q_{x} = \frac{n *_n m_x}{1+(n-_n a_x) *_n m_x}$
${}_{n}p_{x}$	Probability of surviving from age x to x + n	: ${}_{n}p_{x} = 1 -_n q_x$
${}_{n}L_{x}$	Person-years lived between age x and x + n	: $L_x = n * l_{x+n} +_n a_x *_n d_x$
T_x	Persons-years lived above age x	: $T_x = T_{x-1} - L_{x-1}$, $T_0 = \sum_{a=x}^{k} L_a$
e_x^0	Expectation of life at age x	: $e_x^0 = T_x / l_x$

${}_n m_x$	Death rate in the cohort between ages x and x + n	:	${}_n m_x = {}_n d_x / {}_n L_x$
${}_n a_x$	Average person-years lived in the interval by those dying in the interval	:	${}_n a_x = \frac{(-n/24)_n d_{x-n} + (n/2)_n d_x + (n/24)_n d_{x+n}}{{}_n d_x}$

Since, a cohort life table uses a specific (real) population from a defined region with conventional age groupings with death case data from death certificates from that region during the period specified, then the construction of a life table for a cohort poses little difficulty but because cohort data might be outdated, or incomplete. So, actuaries and demographers have developed another way to construct life table called “period life table”. A period life table (sometimes called current life table) uses a hypothetical population with conventional age groupings with the death case data from census records for a specified period of time, usually a year.

To construct a life table for Pakistan, I have used the number of population and the number of deaths in each age group. Moreover, I have started life table with a cohort of l_0=100,000. So according to my own constructed life tables for the rural and urban, male and female and the overall population, the urban and rural populations let us know that the life expectancy at birth in urban area (68.7 years) is 6% higher than the rural areas (64.3 years). Similarly the probability of dying in the earliest age interval is also 10% smaller in the urban than the rural areas population (i.e.0.06444 & 0.07197 for the urban and rural populations respectively). Further, my calculations states that life expectancy at birth for the male is 63.7 shorter than the female 68.4 years, 7% higher in the females, as well as the probability of dying for male in the first age interval is 25% higher than the female (i.e.0.078 & 0.062 respectively).

A life table of Pakistan’s population for the overall areas, urban and rural wise and male and females wise are given below:

Table 6.2: Period Life Table of Pakistan

Exact Age	Mid-year Pop. in age interval x to x + n	Deaths between age x to x + n during the year	Age specific death rate	Average person-years lived in the interval by those dying in the interval	Probability of dying between ages x to x+n	Prob. of surviving from age x to x + n	No. of person left alive at age x	No. of person dying between x to x+n	Person-years lived between ages x to x+n	Person-years lived above age x	Expect--ation of life at age x
X	**nNx**	**nDx**	**nmx**	**nax**	**nqx**	**npx**	**lx**	**ndx**	**nLx**	**Tx**	**ex**
0	19540467	365729	0.01871649	0.0952	0.069767	0.93023	100000	6977	372757	6581599	65.8160
4	22554631	33029	0.0014644	1.5983	0.00728571	0.99271	93023	678	462811	6208841	66.7450
9	20255889	20510	0.00101255	2.521646	0.00504994	0.99495	92346	466	460562	5746030	62.2231
14	17275679	27417	0.00158703	2.560213	0.00790379	0.99210	91879	726	457581	5285468	57.5263
19	13558584	20192	0.00148924	2.583213	0.00741859	0.99258	91153	676	454075	4827888	52.9646
24	10833092	23990	0.00221451	2.487119	0.01101159	0.98899	90477	996	449893	4373813	48.3418
29	8432325	11624	0.0013785	2.607613	0.00686885	0.99313	89481	615	445866	3923920	43.8522
34	8352417	24880	0.00297878	2.715714	0.0147838	0.98522	88866	1314	441045	3478054	39.1382
39	6777652	30926	0.00456294	2.557294	0.02255737	0.97744	87552	1975	432823	3037009	34.6880
44	6276492	27537	0.00438732	2.712641	0.02169862	0.97830	85577	1857	423244	2604186	30.4309
49	4586117	43405	0.00946443	2.698737	0.04622836	0.95377	83720	3870	408926	2180942	26.0504
54	3544175	51031	0.01439856	2.662545	0.06949134	0.93051	79850	5549	385378	1772017	22.1918
59	2933669	68531	0.02336017	2.583102	0.110356	0.88964	74301	8200	351007	1386639	18.6624
64	2038506	58286	0.02859251	2.639832	0.13342514	0.86657	66102	8820	308459	1035632	15.6673
69	1464156	82325	0.05622693	2.529011	0.24648669	0.75351	57282	14119	251112	727173	12.6946
74	654088	37357	0.05711311	2.434808	0.24988612	0.75011	43163	10786	188849	476062	11.0295
79	428280	34079	0.07957178	2.710329	0.33184511	0.66815	32377	10744	156657	287213	8.8709
85 +	354168	58685	0.1656982	6.035069	1.00000	0.00000	21633	21633	130556	130556	6.0351

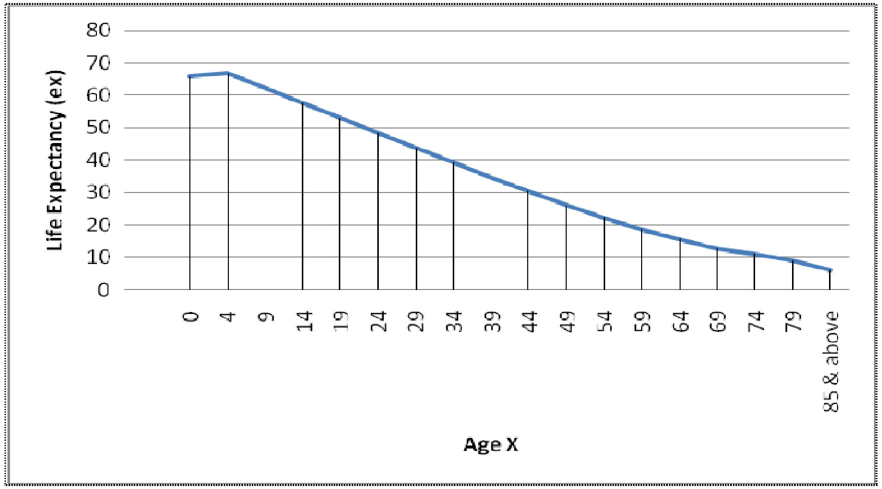
(For $_n a_x$ column, first I have used n/2 and computed the $_n d_x$ values and then computed the original $_n a_x$ values by using $_n d_x$ column. The formula of $_n a_x$ computation is written above in 5th column).

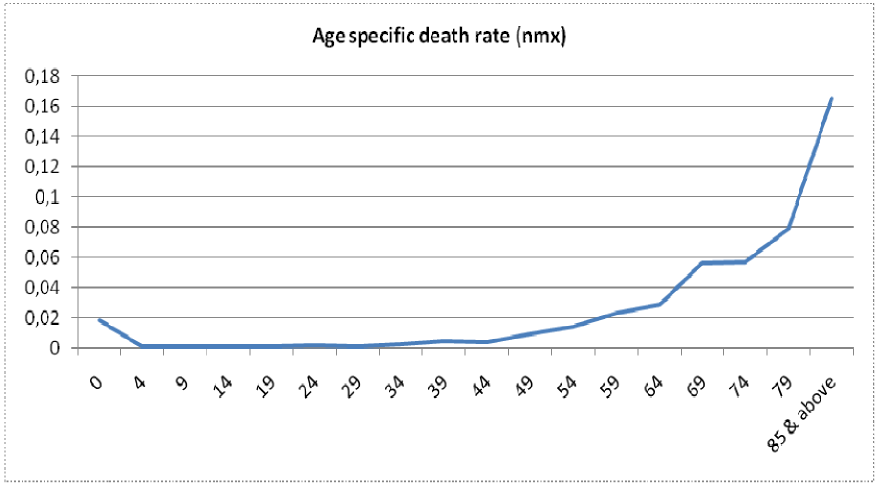
Formulas that are used to find these calculations are as under:-

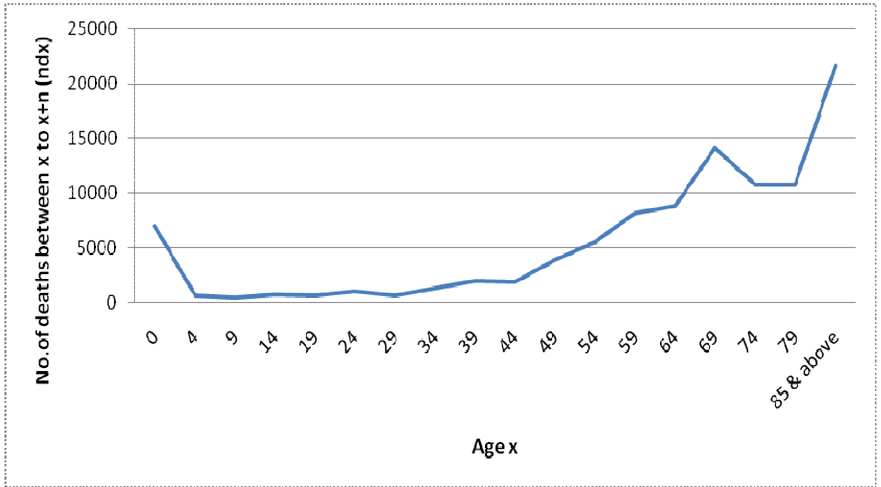
$l_{x+n} = l_x -_n d_x, l_0 base$	$_n d_x = l_x - l_{x+n}$ OR $_n d_x = l_x *_n q_x$	$_n q_x = \dfrac{n*_n m_x}{1+(n-_n a_x)*_n m_x}$	$_n p_x = 1-_n q_x$	$L_x = n*l_{x+n} +_n a_x *_n d_x$
$T_x = T_{x-1} - L_{x-1}, \boldsymbol{T_o = \sum_{a=x}^{k} L_a}$	$e_x^0 = T_x / l_x$	$_n m_x =_n d_x /_n L_x$	$_n a_x = \dfrac{(-n/24)_n d_{x-n} + (n/2)_n d_x + (n/24)_n d_{x+n}}{_n d_x}$	
See table 6.1				

Calculation of $_n a_x$ values below age 5:

For $_1 a_0$: since $_1 m_0$ = 0.01871649 < 0.107 then	0.45+2.684 * $_1 m_0$ = 0.45+2.684(0.01871649) = 0.0952
For $_4 a_1$: since $_1 m_0$ = 0.01871649 < 0.107 then	1.651-2.816 * $_1 m_0$ = 1.651-2.816(0.01871649) = 1.5983
See table 6.7	


Life Expectancy (ex)
80
70
60
50
40
30
20
10
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
Age X


Age specific death rate (nmx)
0,18
0,16
0,14
0,12
0,1
0,08
0,06
0,04
0,02
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above


No.of deaths between x to x+n (ndx)
25000
20000
15000
10000
5000
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
Age x

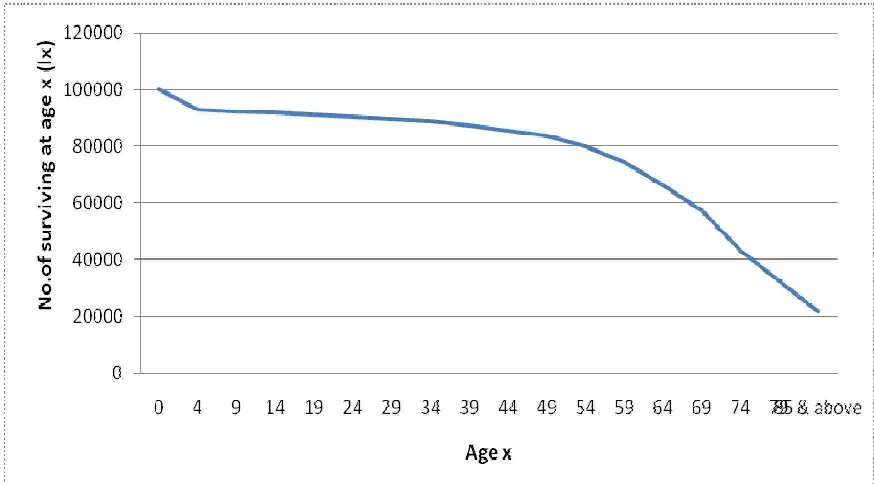
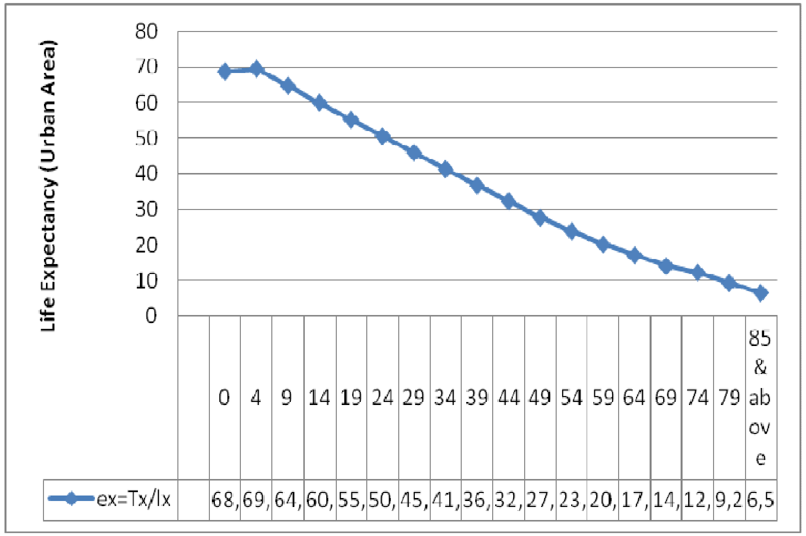

No.of surviving at age x (lx)
120000
100000
80000
60000
40000
20000
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
Age x

Table 6.3: Period Life Table of Pakistan (Urban Area)

Exact Age	Mid-year Pop. in age interval x to x + n	Deaths between age x to x + n during the year	Age specific death rate	Average person-years lived in the interval by those dying in the interval	Probability of dying between ages x to x+n	Prob. of surviving from age x to x + n	No. of person left alive at age x	No. dying between x to x+n	Person-years lived between ages x to x+n	Person-years lived above age x	Expect--ation of life at age x
X	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx	ex
0	5761626	99064	0.017193757	0.09115	0.064443896	0.93556	100000	6444	374810	6875622	68.7562
4	6759356	5458	0.000807473	1.60258	0.004026321	0.99597	93556	377	466498	6500813	69.4861
9	6854584	5825	0.000849796	2.505743281	0.004239974	0.99576	93179	395	464907	6034314	64.7605
14	6630532	5551	0.000837188	2.605828723	0.004177196	0.99582	92784	388	462950	5569407	60.0256
19	5604996	7205	0.00128546	2.583453457	0.006406713	0.99359	92396	592	460501	5106457	55.2669
24	4174036	5700	0.001365585	2.525049514	0.006804693	0.99320	91804	625	457460	4645956	50.6072
29	3112553	4571	0.001468569	2.618540361	0.007315987	0.99268	91180	667	454230	4188496	45.9368
34	3081885	6877	0.002231427	2.687731417	0.011095237	0.98890	90513	1004	450052	3734265	41.2569
39	2564848	9089	0.00354368	2.513396038	0.017562807	0.98244	89508	1572	443611	3284213	36.6917
44	2458241	6219	0.002529858	2.933765748	0.012569789	0.98743	87936	1105	436918	2840602	32.3030
49	1772219	16172	0.009125283	2.687596961	0.044608747	0.95539	86831	3873	424471	2403684	27.6823
54	1325845	15100	0.011388963	2.676012542	0.055368346	0.94463	82958	4593	403305	1979213	23.8581
59	1003276	20888	0.020819794	2.606480884	0.098948741	0.90105	78364	7754	372436	1575908	20.1100
64	713604	18410	0.025798622	2.639773936	0.121177575	0.87882	70610	8556	331660	1203472	17.0439
69	499821	24391	0.04879947	2.493838228	0.217466701	0.78253	62054	13495	276533	871812	14.0493
74	244044	8951	0.036677812	2.498297408	0.167985692	0.83201	48559	8157	222403	595279	12.2588
79	129604	10332	0.079719762	2.791938884	0.332359715	0.66764	40402	13428	195414	372876	9.2292
85 +	116534	17713	0.151998558	6.579009767	1.00000	0.00000	26974	26974	177462	177462	6.5790


Life Expectancy (Urban Area)
80
70
60
50
40
30
20
10
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
ex=Tx/lx
68, 69, 64, 60, 55, 50, 45, 41, 36, 32, 27, 23, 20, 17, 14, 12, 9,2 6,5

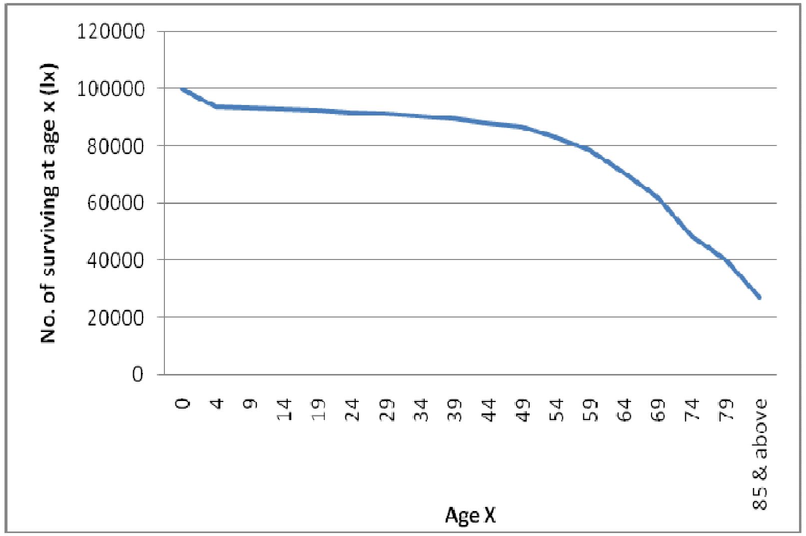
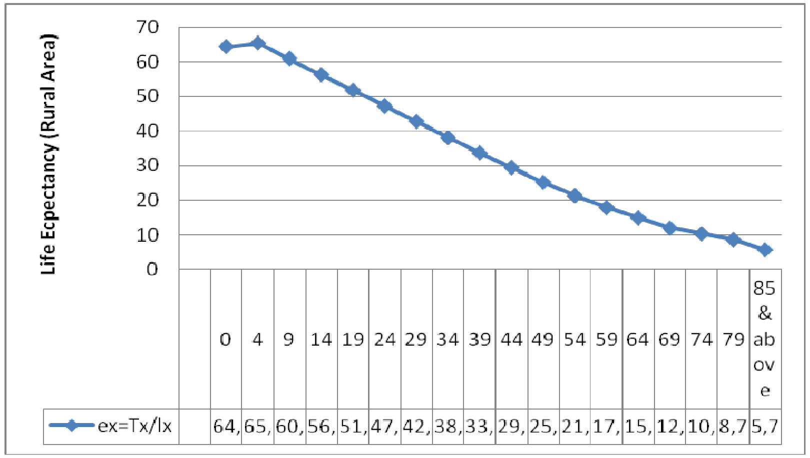

No. of surviving at age x (lx)
120000
100000
80000
60000
40000
20000
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
Age X

Table 6.4: Period Life Table of Pakistan (Rural Area)

Exact Age	Mid-year Pop. in age interval x to x + n	Deaths between age x to x + n during the year	Age specific death rate	Average person-years lived in the interval by those dying in the interval	Probability of dying between ages x to x+n	Prob. of surviving from age x to x + n	No. of person left alive at age x	No. dying between x to x+n	Person-years lived between ages x to x+n	Person-years lived above age x	Expect--ation of life at age x
X	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx	ex
0	13778841	266665	0.01935322	0.096944	0.071976062	0.92802	100000	7198	371907	6434523	64.3452
4	15795276	27571	0.00174552	1.5965	0.008676066	0.99132	92802	805	461272	6062616	65.3282
9	13401325	14685	0.00109579	2.553783018	0.005463967	0.99454	91997	503	458729	5601344	60.8860
14	10645147	21867	0.00205418	2.552070587	0.010218401	0.98978	91495	935	455136	5142615	56.2068
19	7953588	12987	0.00163285	2.582111292	0.008131048	0.99187	90560	736	450957	4687479	51.7612
24	6659055	18290	0.00274664	2.474328920	0.013639523	0.98636	89823	1225	446054	4236522	47.1651
29	5319772	7053	0.00132581	2.594414590	0.006607144	0.99339	88598	585	441527	3790468	42.7827
34	5270532	18003	0.00341578	2.727613458	0.016934311	0.98307	88013	1490	436338	3348941	38.0506
39	4212804	21837	0.00518348	2.578177854	0.025585858	0.97441	86522	2214	427077	2912603	33.6630
44	3818252	21318	0.00558318	2.648996928	0.027531630	0.97247	84309	2321	415740	2485526	29.4813
49	2813897	27234	0.00967839	2.702148862	0.047248734	0.95275	81987	3874	400253	2069786	25.2452
54	2218330	35931	0.01619732	2.654151447	0.077834810	0.92217	78114	6080	375368	1669533	21.3731
59	1930393	47643	0.02468047	2.570413734	0.116230760	0.88377	72034	8373	339237	1294165	17.9661
64	1324901	39876	0.03009734	2.638572403	0.139955959	0.86004	63661	8910	296031	954928	15.0002
69	964335	57934	0.06007663	2.544194725	0.261159245	0.73884	54751	14299	238010	658897	12.0343
74	410044	28405	0.06927305	2.41549952	0.295235591	0.70476	40453	11943	172405	420887	10.4045
79	298676	23748	0.07951091	2.656707165	0.331633366	0.66837	28510	9455	137966	248482	8.7158
85 +	237634	40972	0.17241641	5.799912135	1.00000	0.00000	19055	19055	110516	110516	5.7999


Life Ecpectancy (Rural Area)
70
60
50
40
30
20
10
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
ex=Tx/lx
64, 65, 60, 56, 51, 47, 42, 38, 33, 29, 25, 21, 17, 15, 12, 10, 8,7 5,7

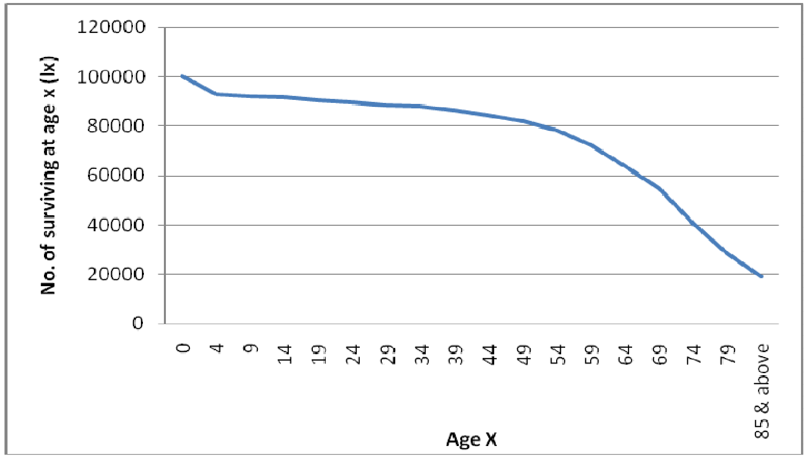

No. of surviving at age x (lx)
120000
100000
80000
60000
40000
20000
0
0 4 9 14 19 24 29 34 39 44 49 54 59 64 69 74 79 85 & above
Age X

Table 6.5: Period Life Table of Pakistan (for Male)

Exact Age	Mid-year Pop. in age interval x to x + n	Deaths between age x to x + n during the year	Age specific death rate	Average person-years lived in the interval by those dying in the interval	Probability of dying between ages x to x+n	Prob. of surviving from age x to x + n	No. of person left alive at age x	No. dying between x to x+n	Person-years lived between ages x to x+n	Person-years lived above age x	Expect--ation of life at age x
X	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx	ex
0	9783859	205840	0.021038733	0.5065	0.078393	0.92161	100000	7839	372613	6368912	63.6891
4	11710324	16652	0.001421993	1.5917	0.007075673	0.99292	92161	652	458581	5996298	65.0635
9	10636015	6871	0.000646013	2.623100864	0.003224855	0.99678	91509	295	456805	5537717	60.5158
14	9063876	16500	0.001820413	2.611721756	0.009060830	0.99094	91213	826	454001	5080912	55.7035
19	6824723	11195	0.00164036	2.579373166	0.008168301	0.99183	90387	738	450089	4626911	51.1900
24	5268436	13101	0.002486696	2.528756301	0.012356663	0.98764	89649	1108	445474	4176821	46.5910
29	3957414	8007	0.002023291	2.523699035	0.010065541	0.98993	88541	891	440477	3731347	42.1426
34	4132910	11482	0.002778188	2.741602633	0.013795125	0.98620	87650	1209	435226	3290870	37.5457
39	3496263	18802	0.005377742	2.595924683	0.026532004	0.97347	86441	2293	426469	2855644	33.0359
44	3277150	17884	0.005457181	2.752828272	0.026918654	0.97308	84147	2265	415073	2429175	28.8682
49	2429295	30870	0.01270739	2.691957205	0.061580629	0.93842	81882	5042	396804	2014102	24.5976
54	1864568	35120	0.018835462	2.599066636	0.08994206	0.91006	76840	6911	366921	1617298	21.0477
59	1637251	41470	0.025329042	2.556932202	0.119103282	0.88090	69929	8329	328821	1250377	17.8808
64	1106476	35664	0.032232059	2.609895652	0.149142381	0.85086	61600	9187	285031	921556	14.9604
69	857310	49296	0.057500787	2.533196401	0.25136913	0.74863	52413	13175	229126	636525	12.1445
74	358255	24072	0.067192363	2.44049751	0.287643242	0.71236	39238	11286	167973	407398	10.3828
79	250734	21720	0.086625667	2.640546383	0.356025886	0.64397	27951	9951	132878	239426	8.5658
85 +	202880	34274	0.168937303	5.91935578	1.00000	0.00000	18000	18000	106548	106548	5.9194

Table 6.5: Period Life Table of Pakistan (for Female)

Exact Age	Mid-year Pop. in age interval x to x + n	Deaths between age x to x + n during the year	Age specific death rate	Average person-years lived in the interval by those dying in the interval	Probability of dying between ages x to x+n	Prob. of surviving from age x to x + n	No. of person left alive at age x	No. dying between x to x+n	Person-years lived between ages x to x+n	Person-years lived above age x	Expect--ation of life at age x
X	nNx	nDx	nmx	nax	nqx	npx	lx	ndx	nLx	Tx	ex
0	9756608	159889	0,016387765	0,47196	0,0619683	0,93803	100000	6197	378137	6840156	68,4016
4	10844307	16376	0,001510101	1,62796	0,007512252	0,99249	93803	705	466640	6462019	68,8891
9	9619874	13640	0,001417898	2,470781	0,007064448	0,99294	93098	658	463848	5995379	64,3982
14	8211804	10917	0,001329428	2,484257	0,00662512	0,99337	92441	612	460673	5531531	59,8386
19	6733861	8997	0,001336083	2,593914	0,006658177	0,99334	91828	611	457613	5070858	55,2210
24	5564656	10888	0,001956635	2,442034	0,009735552	0,99026	91217	888	453865	4613245	50,5744
29	4474911	3617	0,000808284	2,802506	0,004033271	0,99597	90329	364	450734	4159380	46,0470
34	4219507	13398	0,003175252	2,684733	0,015751227	0,98425	89965	1417	446280	3708646	41,2234
39	3281389	12124	0,003694777	2,496202	0,018304804	0,98170	88548	1621	438686	3262366	36,8431
44	2999342	9652	0,003218039	2,624509	0,015961782	0,98404	86927	1388	431165	2823680	32,4835
49	2156822	12535	0,005811792	2,70892	0,028642793	0,97136	85539	2450	421571	2392516	27,9698
54	1679608	15911	0,009473044	2,793188	0,046269438	0,95373	83089	3844	405834	1970945	23,7209
59	1296418	27061	0,020873669	2,614568	0,099192087	0,90081	79245	7860	376572	1565111	19,7504
64	932030	22621	0,024270678	2,685773	0,114411291	0,88559	71384	8167	336503	1188539	16,6499
69	606846	33029	0,054427318	2,521154	0,239542455	0,76046	63217	15143	278227	852035	13,4779
74	295833	13285	0,044907093	2,419105	0,201871778	0,79813	48074	9705	216107	573808	11,9360
79	177547	12359	0,069609737	2,816663	0,29645781	0,70354	38369	11375	190403	357701	9,3226
85 +	151288	24411	0,161354503	6,197534	1,00000	0,00000	26994	26994	167298	167298	6,1975

A population whose total number and distribution by age doesn’t change with time is called a stationary population. Moreover, in the stationary population, the number of births per year remains constant and each cohort of births experiences current observed mortality rates throughout life, as well as the net migration is ‘0’ or closed to migration.

6.1 Life expectancy at very young age: -

A large proportion of the increase in life expectancy is attributable to decreases in mortality in childhood period and it has been seen that child expectancy estimates are very sensitive to procedures used in the very young, high mortality rate. The estimation of ${}_nq_x$ (probability of dying in age x to x + n) for very young age level, can be made by assigning deaths to the appropriate birth cohorts, but the conversion of death rate into the probability of dying it is important to recognize that the value of ${}_na_x$ is (empirically) a function of the level of mortality itself. Further, according to demographers, the lower the level of mortality, the more heavily will infant deaths be concentrated at the earliest stages of infancy and the influence of the prenatal environment becomes increasingly dominant relative to the postnatal environment.

The following table gives us the relation between the values of ${}_1a_0$ and ${}_4a_1$ and also the values of ${}_1q_0$, and the formulas presented in this table can be recommended for use in deriving ${}_na_x$ values below age 5. This method is got from the Coale and Demeny (1983) west model life tables.

Table 6.7:

[38] *Values of $_{n}a_{x}$ for use below age 5*		
	Males	**Females**
Value of $_{1}a_{0}$		
If $_{1}m_{0} \geq 0.107$	0.330	0.350
If $_{1}m_{0} < 0.107$	0.45+2.684 * $_{1}m_{0}$	0.053+2.800 * $_{1}m_{0}$
Value of $_{4}a_{1}$		
If $_{1}m_{0} \geq 0.107$	1.352	1.361
If $_{1}m_{0} < 0.107$	1.651-2.816 * $_{1}m_{0}$	1.522-1.518 * $_{1}m_{0}$

6.2 Strategies for choosing $_{n}a_{x}$:-

$_{n}a_{x}$ is the average person-years lived in the interval by those dying in the interval. There are many ways to find the set of values of $_{n}a_{x}$, like:

- By calculating it directly.
- By smoothing (graduating) the death distribution within each age interval.
- Borrowing values from another population
- The rule of thumb

  Making one of the two assumptions:

  - $_{n}a_{x}$ is half the length of the age interval (n/2) OR
  - $_{n}m_{x}$ is constant in the interval which negates the necessity of using $_{n}a_{x}$.

Usually n/2 is used for all age groups except the first one, because the mortality rate between ages 0 to 5 changes very rapidly and falling very quickly at first and then flattening out. Consequently most of the deaths occur in the early life between 0 to 5 age interval and hence $_{n}a_{x}$ is significantly less than n/2 in the first two age groups (0, 1-4).

[38] Coale and Demeny (1983) west model life table.

According to the direct method, if the data on exact age at death are available for a population (e.g. 61.19, 21.62….), then it is clearly possible to take all of the deaths during a period that fall within particular n-year wide age interval and compute ${}_n a_x$ directly and also this value would pertain to a population rather than to a cohort. But such information is rarely available and even if it is available, it is not always advisable to use it because ${}_n a_x$ values observed in a population are influenced by that population's age distribution within the n-year wide age interval.

The values of ${}_n a_x$ is not only affected by the death rate but also by the level of mortality. The higher is mortality within a particular age interval, the more will deaths be concentrated at the beginning of that interval because fewer people will survive to be at risk of death near the end of the interval. Then a method called graduating is an important method in the estimation of ${}_n a_x$ from information on the age distribution of deaths in the life table, assuming that this distribution d (a), follows a polynomial function of the second degree in the interval x-n to x+2n (Keyfitz, 1966).

$$d(a) = A + Ba + Ca^2 \qquad \mathbf{x - n} < a < x + \mathbf{2}n$$

Under this assumption:

$$ {}_n a_x = \frac{(-n/24)_n d_{x-n} + (n/2)_n d_x + (n/24)_n d_{x+n}}{{}_n d_x} \tag{6.1}$$

In the estimation produces, an estimate of n/2 is used when deaths are symmetrically distributed and the equation 6.1 requires having estimates of number of deaths (${}_n d_x$) that is usually estimated from the ${}_n m_x \rightarrow {}_n q_x$ conversion (see table 6.1) itself and thus requires ${}_n a_x$ information (where ${}_n m_x$ is the age specific death rate). To solve the problem, one must use iteration and it makes most sense to start by taking ${}_n a_x$ = n/2 to obtain a first set of ${}_n d_x$ estimates and then using these ${}_n d_x$ to obtain a new set of ${}_n a_x$ estimates with the above estimation. Thus, the new set of ${}_n a_x$ values can be reused in ${}_n m_x \rightarrow {}_n q_x$ conversion until stable estimates of ${}_n a_x$ and ${}_n d_x$ are obtained. One limitation of this method is that it does not permit the estimation of ${}_n a_x$ in the first and last age group and also it requires that all age groups should have the same width n.

Another method of ${}_{n}a_{x}$ is to borrow a set of values. If the level and the shape of the ${}_{n}m_{x}$ curve is similar to another population for which ${}_{n}a_{x}$ values have been accurately estimated then a simple and reasonable expedient method is to adopt that set of values and those borrowed values should correspond to that sex for which they are being used because ${}_{n}a_{x}$ values vary significantly between the sexes. [39]Keyfitz and Flieger (1971) provide sets of ${}_{n}a_{x}$ values for population based graduation technique above the age of 10. The values of ${}_{5}a_{x}$ for older ages tend to exceed 2.5 years, reflecting the rapid rise in mortality with age, so that deaths are concentrated towards the upper end of the age range. At the high ages, the ${}_{5}a_{x}$ values start to decline as the increasingly high mortality levels leave fewer survivors available to die at the upper end of an age range.

The rule of thumb is also an important method to find ${}_{n}a_{x}$ values, in which there are two rules that are commonly implemented in choosing ${}_{n}a_{x}$ except infancy (under age 1). Each of these rules works extremely well-leads to trivial error-when data are arrayed in one year wide age interval. The first rule is to set ${}_{n}a_{x} = \frac{n}{2}$ that means deaths are assumed to be occur on average (2.5) halfway through the interval. This assumption leads to the following equation 6.2 (when persons are dying in the interval, on average, halfway through the interval):

$$ {}_{n}q_{x} = \frac{n * {}_{n}m_{x}}{1 + (n/2) * {}_{n}m_{x}} = \frac{2 * n * {}_{n}m_{x}}{2 + n * {}_{n}m_{x}} \qquad (6.2) $$

Further, the second assumption is, the age specific death rate (${}_{n}m_{x}$) is constant in the age interval x to x+n. In this case:

$$ {}_{n}p_{x} = 1 - {}_{n}q_{x} = e^{-n * {}_{n}m_{x}} $$

Here no conversion is involving for ${}_{n}a_{x}$ is required and of course, a value of ${}_{n}a_{x}$ is implicit in this conversion formula, in particular:

[39] Population, facts and methods of demography

$$ {}_n a_x = n + \frac{1}{{}_n m_x} - \frac{n}{1 - e^{-n^* {}_n m_x}} $$

In this case, the ${}_n a_x$ is necessarily be less than n/2 (see page 49).

6.3 The Open-ended age interval: -

The methods and formulas presented so far are incapable for the open ended (or terminal) age interval. In the open ended interval 'n' is infinity. To solve this interval, return to the formula of death rate in a cohort, since:

$$ {}_n m_x = {}_n d_x / {}_n L_x \text{ (See table 1)} $$

When n = ∞ we must have:

$$ {}_\infty m_x = {}_\infty d_x / {}_\infty L_x $$

OR

$$ {}_\infty L_x = {}_\infty d_x / {}_\infty m_x $$

But the number of persons dying in a cohort above age x (whether the cohort is real or hypothetical) must be equal to the number of persons surviving to age x: i.e. ${}_\infty d_x = l_x$ so:

$$ {}_\infty L_x = l_x / {}_\infty m_x \qquad (6.4) $$

Where ${}_\infty m_x$ is observed and l_x can be calculated on the basis of mortality at all ages below x. So, the number of persons lived above age x can be calculated and used to complete life table. Of course: ${}_\infty q_x = 1.00$ and ${}_\infty p_x = 0.00$.

The Coale and Demeny (West model life table series) procedure based on the cumulative distribution and can also be used for demographic equation at older ages, where the information for these ages is aggregated into one single open ended age interval. Then, the life expectancy at age starting the open-ended age interval is by taking the reciprocal of the death rate (${}_n m_x$) in that interval. Specifically, for an open ended age interval starting at age x, e_x^0 is simply taking $1/{}_n m_x$. This formula is correct if the population in that interval was stationery, however in most populations, mortality at older ages has been declining and as a result the older population has been growing rather than stationery.

The procedures for dealing with open ended interval have become increasingly important as more people have survived to its beginning. The most commonly encountered open ended interval begins with age 85, to which age nearly half of females in recent period life table of better-off nations will survive. The analyst clearly adopt a high enough age for starting the open ended interval that only a small fraction of the population survive to that age (but where the data is available). In my calculations equation 6.4 has been used to find person year lived after 85+ age interval.

6.4 Interpretation of life tables: -

A life table summarizes the mortality experience of a population specifically a period life table and each parameter presented corresponds to a specific age or age interval e.g. the probability of surviving (${}_{n}p_{x}$) or probability of dying between x to x+n (${}_{n}q_{x}$), the age specific death rate between age x and x+n (${}_{n}M_{x}$) or the life expectancy at age x (e_{x}). The verbal interpretation has been described earlier and the graphical representation also helps us to understand the different columns of life table.

According to our different life tables (urban and rural, male and female), it can be seen that life expectancy at birth in the urban population is (69 years) higher than the rural (64 years). The life tables states that rural areas having more deaths in the early ages than the urban. It is for the reason of services given by the state to urban and rural areas. Rural areas having smaller number of health care centers and medical staff facilities than the urban, so that having more deaths in the infancy. As well as awareness on quality health care, sanitation, or food etc. in urban areas is much better than the rural. Life tables are important in making population forecasting and this is done in the chapter 8.

7. Standardization and Decomposition Techniques:-

Nearly in every population, the rate of occurrence in the demographic events varies very sharply with age. In fact, the crude rates unsuccessful to account for age variation in the underlying rate schedules. In the cases of fertility and mortality, this variation mainly reflects age difference in physiological capacity and similarly the age variation in migration rates seems to reflect primarily age differences in economic and social gains from movement. Then, a method called [40]standardization is applied to various rates in order to compare the two populations (e.g. urban and rural) mortality of any country.

Demographers mostly used the techniques of standardization to eliminate the composition effects from overall rates of some phenomena in two or more different populations. In fact, the precursors of decomposition method are the techniques of standardization and the method of standardization is itself the origins of decomposition method. Regardless of importance and wide usage in demography, standardized rates have always been considered a bit problematic and viewed as unreliable due to their dependence on an arbitrary standard. Then, this problem was impetus behind more research on the method of comparisons.

Another field of research, namely the decomposition of the difference between the crude death rates in two different population (or rural, urban populations) was developed by Kitawaga (1955).

7.1 Decomposition of differences between rates or proportions: -

An important question in demography is to find, how much is the difference between death rate in population A and B in a country is attributable to differences in their age distributions? A technique known as decomposition is given by Kitagawa (1955). We should note that there is no unique answer to the question of decomposition and there are many ways to decompose a difference, but the method given by Kitagawa has an advantage of economy and expositional cleanness. Suppose that we are interested in decomposing a difference between crude death rates (CDR) in the populations A and B.

[40] Age-standardization by Demography: Measuring and modeling population by Samuel H. Preston

Define the original differences:

$$\Delta = CDR^B - CDR^A$$

[41] $$\Delta = \sum_i C_i^B * M_i^B - \sum_i C_i^A * M_i^A \qquad \textbf{(7.1)}$$

Where: [42]CDR is:

$$CDR = \frac{D}{N} = \frac{\sum_{x=0}^{\infty} {}_nD_x}{N} = \frac{\sum_{x=0}^{\infty} ({}_nD_x / {}_nN_x) *_n N_x}{N}$$

$$\Rightarrow CDR = \sum_{x=0}^{\infty} \frac{{}_nD_x}{{}_nN_x} * \frac{{}_nN_x}{N} = \sum_{x=0}^{\infty} {}_nM_x *_n C_x$$

Where ${}_nC_x = \frac{{}_nN_x}{N}$ = the proportion of total population that belongs to the age interval x to x+n, and nDx is the number of deaths during a year.

and

$${}_nM_x = \frac{\text{Number of deaths in the age x to x + n}}{\text{Number of persons} - \text{year lived in the age x to x + n}}$$

The above CDR equation says that the crude death rate is a combination of two functions: the set of age specific death rates and the proportionate age distribution of the population. In particular, the CDR is the weighted average of age specific death rate, where the weights are supplied by a population's proportionate age distribution (strictly saying, the proportionate distribution of person-years lived). Then of course, the sum of these weights must be the unity:

$$\sum_{x=0}^{\infty} {}_nC_x = \sum_{x=0}^{\infty} \frac{{}_nN_x}{N} = \frac{N}{N} = 1.000$$

[41] Where Mi= age-specific death rate and Ci= Proportion of total population
[42] CDR (crude death rate) = Number of deaths per 1000 people in a year

According to equation (7.1), C_i^s shows the average age distribution and defined as:

$$C_i^S = \frac{C_i^A + C_i^B}{2}$$

Now back to equation (7.1):

$$\Delta = \sum_i C_i^B * M_i^B - \sum_i C_i^A * M_i^A \qquad (7.2)$$

We will now divide each of these terms into two equal parts and add and subtract some additional terms, thereby keeping the difference (delta) constant:

$$\Delta = \frac{\sum_i C_i^B * M_i^B}{2} + \frac{\sum_i C_i^B * M_i^B}{2} - \frac{\sum_i C_i^A * M_i^A}{2} - \frac{\sum_i C_i^A * M_i^A}{2}$$
$$+ \frac{\sum_i C_i^B * M_i^A}{2} - \frac{\sum_i C_i^B * M_i^A}{2} + \frac{\sum_i C_i^A * M_i^B}{2} - \frac{\sum_i C_i^A * M_i^B}{2}$$

Rearranging this equation:

$$\Delta = \frac{\sum_i C_i^B * M_i^B}{2} + \frac{\sum_i C_i^B * M_i^A}{2} - \frac{\sum_i C_i^A * M_i^B}{2} - \frac{\sum_i C_i^A * M_i^A}{2}$$

$$+ \frac{\sum_i C_i^A * M_i^B}{2} + \frac{\sum_i C_i^B * M_i^B}{2} - \frac{\sum_i C_i^A * M_i^A}{2} - \frac{\sum_i C_i^B * M_i^A}{2}$$

Now,

$$\Delta = \sum_i C_i^B \left[\frac{M_i^B + M_i^A}{2}\right] - \sum_i C_i^A \left[\frac{M_i^B + M_i^A}{2}\right] + \sum_i M_i^B \left[\frac{C_i^A + C_i^B}{2}\right] - \sum_i M_i^A \left[\frac{C_i^A + C_i^B}{2}\right]$$

So:

$$\Delta = \sum_i \left(C_i^B - C_i^A\right)\left[\frac{M_i^B + M_i^A}{2}\right] + \sum_i \left(M_i^B - M_i^A\right)\left[\frac{C_i^A + C_i^B}{2}\right] \qquad (7.3)$$

$\Delta =$ Difference in age composition * [weighted by average age-specific mortality] + difference in rate schedule * [weighted by average age composition]

$\Delta =$ Contribution of age compositional differences to Δ $+$ Contribution of rate schedule differences to Δ

So, we can decompose the difference into two terms: the contribution of age distributional differences plus the contribution of rate schedule differences and both of the terms completely account for the original differences. Between these, the first one is clearly interpretable and note that the second one "contribution of rate schedule differences" term is precisely the differences between age standardized death rate in population B and A, when the standard population age composition is applied to both populations in the average age composition in population A and B.

We will now apply this method to the Pakistan's urban and rural population to find out the difference between ages and difference in the rate schedule. The first five columns data have taken from the Statistical Bureau of Pakistan's survey named Pakistan Demographic Survey 2007 (PDS).

Table 7.1:

Age-standardization and decomposition of differences between rates								
Age x	**nNx (U)**	**nDx (U)**	**nNx (R)**	**nDx (R)**	**nCx (U)**	**nCx (R)**	**nMx (U)**	**nMx (R)**
0	5761626	99064	13778841	266665	0,109003815	0,141972624	0,017193757	0,019353224
4	6759356	5458	15795276	27571	0,127879802	0,162749304	0,000807473	0,001745522
9	6854584	5825	13401325	14685	0,129681414	0,138082824	0,000849796	0,001095787
14	6630532	5551	10645147	21867	0,12544259	0,109684077	0,000837188	0,002054175
19	5604996	7205	7953588	12987	0,106040543	0,081951142	0,00128546	0,001632848
24	4174036	5700	6659055	18290	0,078968307	0,068612702	0,001365585	0,002746636
29	3112553	4571	5319772	7053	0,058886181	0,054813173	0,001468569	0,001325809
34	3081885	6877	5270532	18003	0,058305975	0,05430582	0,002231427	0,003415784
39	2564348	9089	4212804	21837	0,048514728	0,04340734	0,003544371	0,005183483
44	2458241	6219	3818252	21318	0,046507296	0,039342007	0,002529858	0,005583183
49	1772219	16172	2813897	27234	0,033528492	0,028993465	0,009125283	0,009678393
54	1325845	15100	2218330	35931	0,025083572	0,022856939	0,011388963	0,01619732
59	1003276	20888	1930393	47643	0,018980911	0,019890132	0,020819794	0,024680467
64	763604	18410	1324901	39876	0,014446573	0,013651342	0,024109355	0,030097343
69	499821	24391	964335	57934	0,00945608	0,009936189	0,04879947	0,060076633
74	244044	8951	410044	28405	0,004617052	0,004224958	0,036677812	0,069273054
79	129604	10332	298676	23748	0,002451969	0,003077459	0,079719762	0,079510908
85 +	116534	17713	237634	40972	0,002204699	0,002448502	0,151998558	0,172416405

Continued….

Ci=(U+R)/2	ASCDR U	ASCDR R	CDR U	CDR R	M1	C1	A=M1*C1
0,12548822	0,00216	0,00243	0,00187	0,00275	0,01827349	-0,03296881	-0,00060246
0,145314553	0,00012	0,00025	0,00010	0,00028	0,00127650	-0,03486950	-0,00004451
0,133882119	0,00011	0,00015	0,00011	0,00015	0,00097279	-0,00840141	-0,00000817
0,117563333	0,00010	0,00024	0,00011	0,00023	0,00144568	0,01575851	0,00002278
0,093995843	0,00012	0,00015	0,00014	0,00013	0,00145915	0,02408940	0,00003515
0,073790504	0,00010	0,00020	0,00011	0,00019	0,00205611	0,01035560	0,00002129
0,056849677	0,00008	0,00008	0,00009	0,00007	0,00139719	0,00407301	0,00000569
0,056305898	0,00013	0,00019	0,00013	0,00019	0,00282361	0,00400016	0,00001129
0,045961034	0,00016	0,00024	0,00017	0,00023	0,00436393	0,00510739	0,00002229
0,042924652	0,00011	0,00024	0,00012	0,00022	0,00405652	0,00716529	0,00002907
0,031260979	0,00029	0,00030	0,00031	0,00028	0,00940184	0,00453503	0,00004264
0,023970256	0,00027	0,00039	0,00029	0,00037	0,01379314	0,00222663	0,00003071
0,019435521	0,00040	0,00048	0,00040	0,00049	0,02275013	-0,00090922	-0,00002068
0,014048957	0,00034	0,00042	0,00035	0,00041	0,02710335	0,00079523	0,00002155
0,009696135	0,00047	0,00058	0,00046	0,00060	0,05443805	-0,00048011	-0,00002614
0,004421005	0,00016	0,00031	0,00017	0,00029	0,05297543	0,00039209	0,00002077
0,002764714	0,00022	0,00022	0,00020	0,00024	0,07961533	-0,00062549	-0,00004980
0,002326601	0,00035	0,00040	0,00034	0,00042	0,16220748	-0,00024380	-0,00003955
	0.00570	**0.00728**	**0.00544**	**0.00754**			**-0.00052807**

Continued….

M2	C2	B=M2*C2
-0,00215947	0,12548822	-0,00027099
-0,00093805	0,14531455	-0,00013631
-0,00024599	0,13388212	-0,00003293
-0,00121699	0,11756333	-0,00014307
-0,00034739	0,09399584	-0,00003265
-0,00138105	0,07379050	-0,00010191
0,00014276	0,05684968	0,00000812
-0,00118436	0,05630590	-0,00006669
-0,00163911	0,04596103	-0,00007534
-0,00305333	0,04292465	-0,00013106
-0,00055311	0,03126098	-0,00001729
-0,00480836	0,02397026	-0,00011526
-0,00386067	0,01943552	-0,00007503
-0,00598799	0,01404896	-0,00008412
-0,01127716	0,00969613	-0,00010934
-0,03259524	0,00442100	-0,00014410
0,00020885	0,00276471	0,00000058
-0,02041785	0,00232660	-0,00004750
		-0.00157492

Where M1= (M(U)+M(R))/2, M2= M(U)-M(R), C1=C(U)-C(R), AND C2=(C(U)+C(R))/2 (see equation 7.3 parts)

Source: Decomposition method by Kitawaga (1955) in Demography: Measuring and Modeling Population Process by Samuel Preston

Data source: Pakistan Demographic Survey 2007 by Statistical Bureau of Pakistan.

CDR (Urban)= Σ Ci * Mi = 0.00544 & CDR (Rural)= Σ Ci * Mi = 0.00754

Original difference =CDR(U)–CDR(R)=0.00544 – 0.00754= -0.00210 (by equation 7.2)

ASCDR (Urban) = 5.70 Per 1000 & ASCDR (Rural) = 7.28 Per 1000

Contribution of age compositional differences = A = -0.00052807

Contribution of age specific rate differences = B = -0.00157492

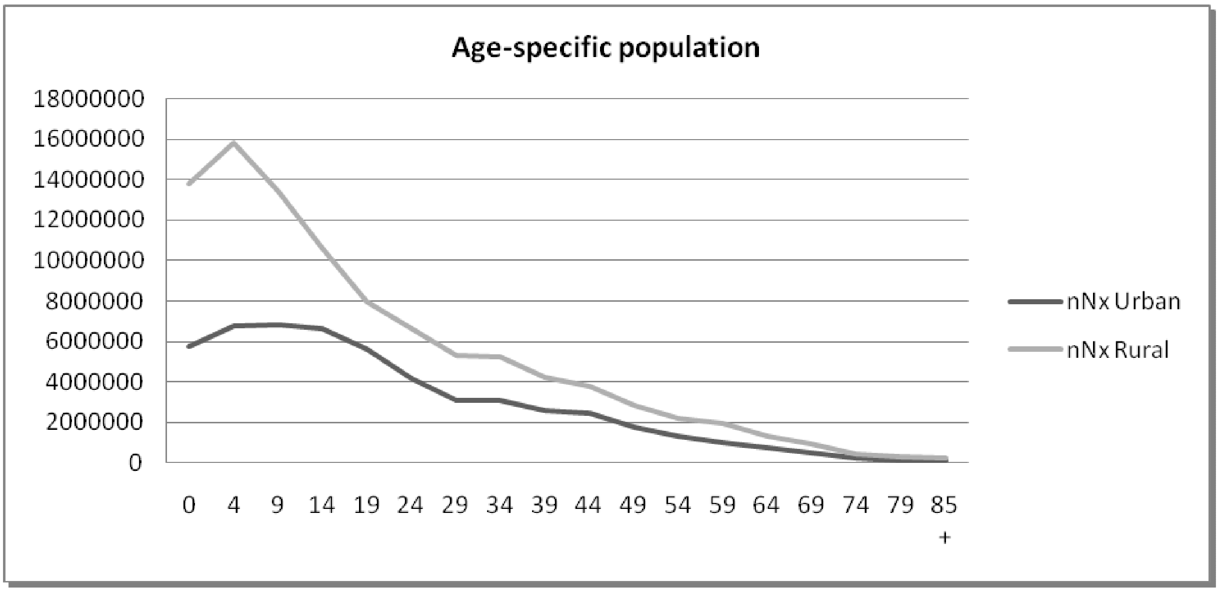
Total Contribution = (-0.00052807) + (-0.00157492) = -0.00210299 (by equation 7.3)

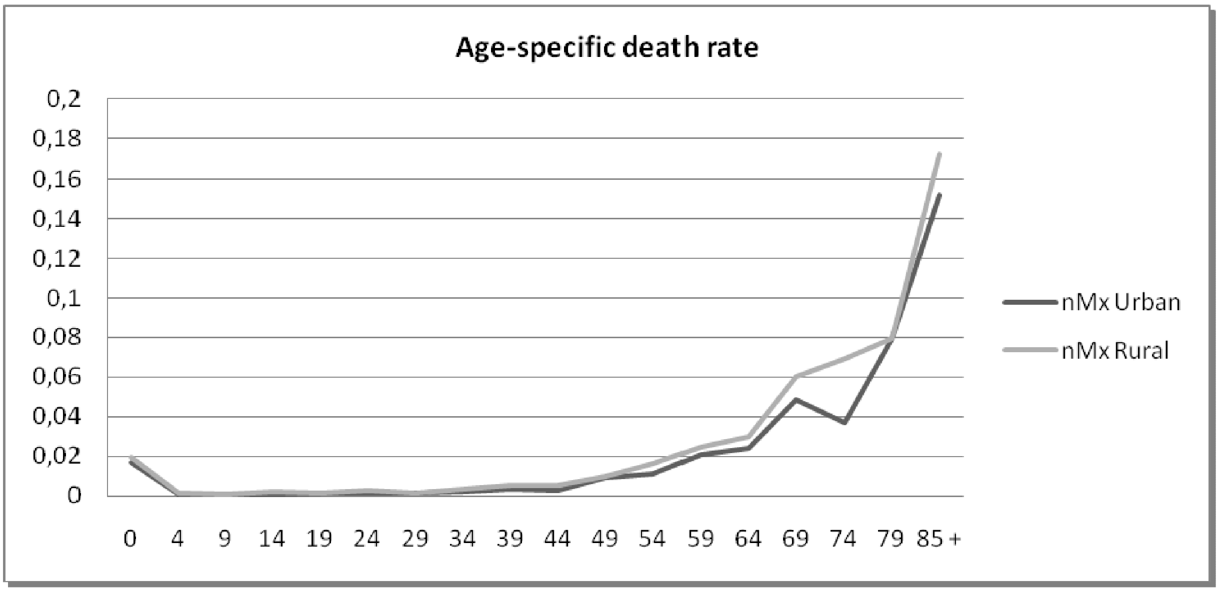
Proportion of difference attributable to differences in age composition =

-0.00052807/-0.00210299 = -0.2511044 = 25%

Proportion of difference attributable to differences in rate schedules =

-0.00157492/-0.00210299 = -0.7488956 = 75%





The table demonstrates the application of the recommended procedure to the decomposition of differences between the crude death rates in Pakistan's urban and rural population (in 2007). The crude death rate of rural population is higher than the urban by 0.00210. Differences in age composition account for 25 percent (-0.00052807/-0.00210299) of the difference between crude death rate while differences in the rate schedule account for the remaining 75 percent. It seems reasonable that both parts are contributing in the same direction to the difference but in many applications, one of the parts will account for more than 100 percent of the original difference and this happens only when the both parts work in the opposite directions and there is no reason to expect that they will normally work in concert. Further, the age specific death rate for urban population 5.7 per 1000 is fewer then the rural 7.28 per 1000 as the urban region are under consideration all the time with having good sanitation facilities, more and effective health care centers then the rural areas and etc. that's why is standard of urban population is quite better then the rural residents.

Both standardization and decomposition procedures can be applied simultaneously to more than one variable, as in the case of standardization, there is nothing to require that age to be one of the variables involved in decomposition. For example, one could decompose a differences between two countries, infant mortality rates into differences due to birth order distributions and differences due to rate schedule differences (differences in their death rates for children of the same birth order). Note that, when age is one of the variables in a standardization or decomposition of demographic rate, it is strongly recommended that age categories be no wider than 5 years interval, (when data permits), because age variation in vital rates is sufficiently great that the age composition with a 10 year age interval can have a substantial effect on the value of an age specific rate pertaining to that interval.

8. Population Projection:-

[43]Population projection (in the field of demography) is an estimate of a future population and is probably the demographic technique that is most frequently requested by demographic planners. In contrast with censuses and intercensal estimates, which usually involve some sort of field data gathering, these population projections usually involve mathematical models based only on pre-existing data. A projection may be done by government departments, non-government organizations. But governments seeks projections of future demographic parameters in order to anticipate needs of all kinds: for schools, roads, heath care centers, medical personnel and national parks etc. while private business seeks these projections in order to estimate the potential size of their future market.

8.1 Fertility Pattern in Pakistan: -

[44]To find-out the population projection for Pakistan, we will start by looking at the fertility pattern of Pakistan, so that it will be easy for us to understand the fertility behavior of Pakistani women. Moreover, the growth rate and the rate of natural increase will also helps us to predict Pakistan's future population.

As mentioned earlier, Pakistan is the world's 6th most populous country with a population of 170 million (2011 estimate by population census organization). To access the fertility pattern in Pakistan, numerous fertility surveys indicate that the total fertility rate in Pakistan remained above 6.0 births per women throughout 1980s (see table 3.1). Table 3.1 shows that after 1990s up to the present time, the demographic research indicates a modest decline in fertility. For example, during 1991-1994 four years period, the estimates indicate a decline of 0.30 births. Then, the Pakistan fertility and family planning survey (PFFPS 1996-97) provides an estimate of total fertility rate to be 5.30, which suggest a slightly more rapid decline during the 1990s than implied by the previous surveys. Collectively, these estimates from different sources indicate the decline in fertility particularly after 1990s.

[43] Internet pages + Wikipedia + Demography: measuring and modeling population by Samuel preston
[44] Reproductivity and age-specific fertility rate in Pakistan after 1981 by Jamal A. Nasir, Munir Akhtar, and M.H.Tahir

Further, according to a survey (PDS-2007 by federal bureau of statistics Pakistan), if the fertility rate/growth rate (see below tables) will remain high then the population of Pakistan will be double in the next 37 years. The below tables (8.1 & 8.2) shows the growth rates and age specific fertility rates of Pakistan in the different years.

[45]Table 8.1

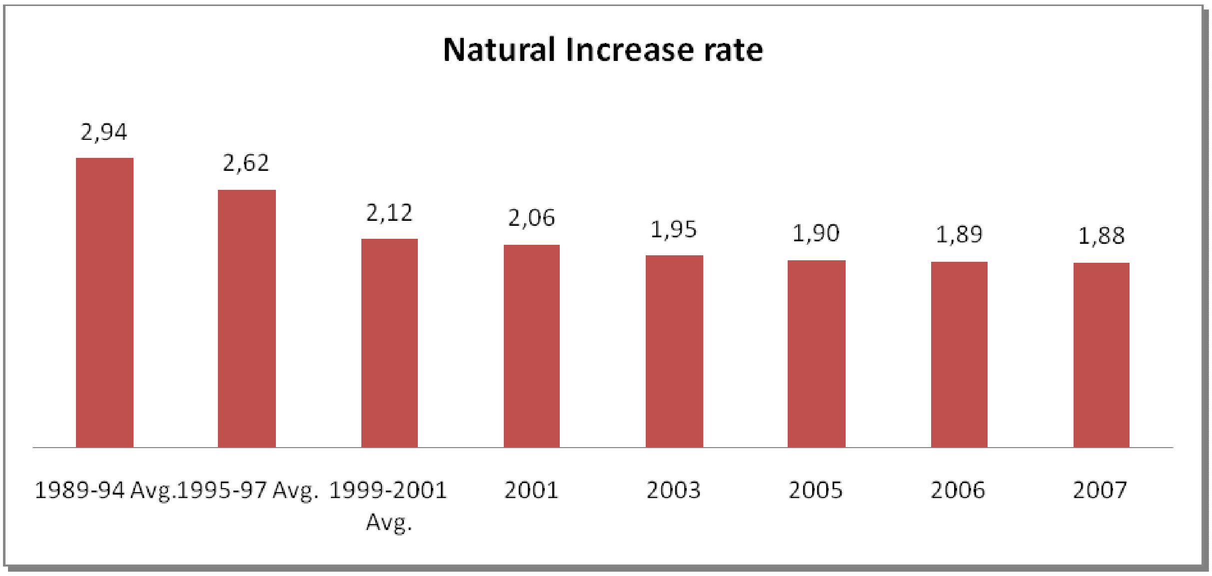
Population Growth Rate of Pakistan (1951-2007)			
Intercensal Growth Rate		***Rate of Natural Increase***	
Census	**Growth rate**	**Survey**	**Growth rate**
1951	1.9	1989-94 (Average)	2.9
1961	1.8	1995-97 (Average)	2.6
1972	2.4	1999-2001 (Average)	2.4
1981	3.6	2001	2.06
1998	3.0	2003	1.95
		2005	1.90
		2006	1.89
		2007	1.88
Sources : Pakistan statistical year book 2010 by Federal Bureau of Statistics, Pakistan and United Nations publications (www.un.org/esa/population/publications/)			

[46]The intercensal growth rate is an estimate of population between official census dates. These intercensal growth estimates can be less informative or more informative than official census figures, depending on methodology, accuracy and the date of data, completeness, and can be released by government or other organizations while the rate of natural increase is the crude birth rate minus crude death rate of a population. If we neglect migration factor, then a positive rate of natural increase means that the population increases and a negative number will represent the decrease in population. Usually developing countries have a positive and quite high rate of natural increase.

---

[45] Publication of Federal Bureau of statistics, Pakistan
[46] Population projection publication (manual 3) by United Nations

According to the Pakistan statistical year book 2010, intercensal growth rate presents the growth rate between the different censuses years and it is an estimated/forecasted growth rate till the next census time while, the rate of natural increase describes the actual increase in population from last year to the next. Moreover, table 8.2 seems to show quite dramatic decrease in fertility rate within last 20 year in each and every age group and on every survey we can see a little bit decrease in natural increase rate but still the population of Pakistan is increasing very fast that really need to implement the family planning policies.



The following table gives information about the age specific fertility rates to get awareness of fertility pattern of Pakistan since 1984.

Table 8.2:

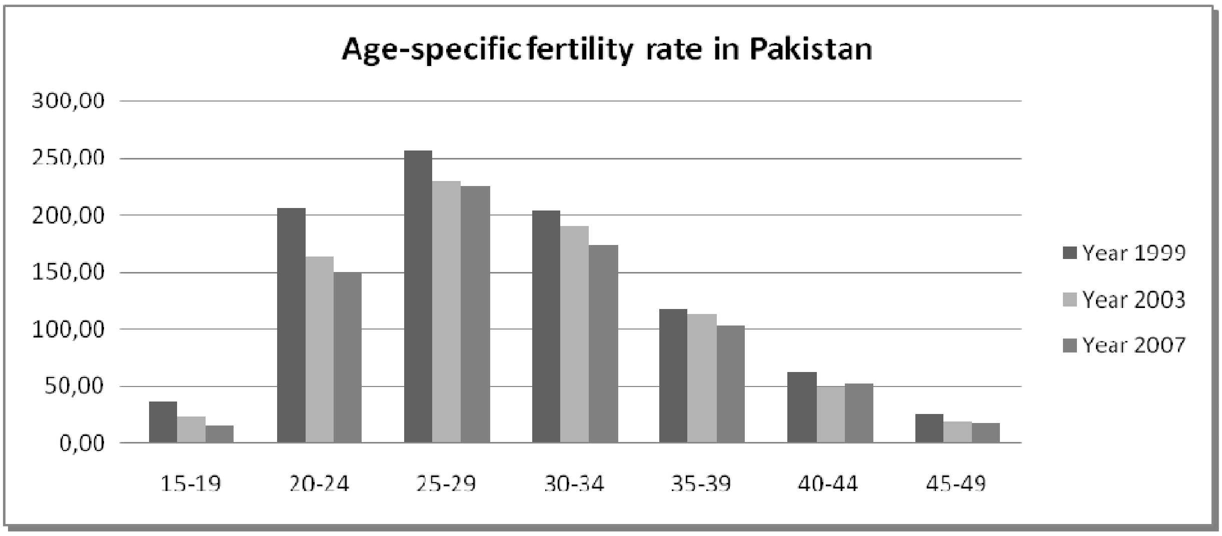
Age specific fertility rates (ASFRs) per 1000 women of Pakistan during 1980s-2000s.							
Years	**Age groups**						
	15-19	**20-24**	**25-29**	**30-34**	**35-39**	**40-44**	**45-49**
1984	65,76	268,33	367,57	314,42	226,07	109,56	37,88
1985	59,15	272,98	350,79	326,98	235,29	108,57	47,88
1986	54,31	265,75	360,26	303,12	226,22	125,98	52,16
1988	66,00	263,60	333,00	278,30	203,30	111,20	41,80
1989	75,70	265,80	323,40	274,30	197,10	102,00	41,60
1990	75,50	274,80	313,20	276,00	175,90	97,00	30,50
1991	69,00	258,20	315,40	259,00	186,50	82,30	27,40
1992	73,30	261,40	312,90	254,50	162,60	74,50	27,80
1995	59,10	243,40	305,10	241,90	148,10	90,10	29,60
1996	54,70	258,20	295,90	255,40	143,00	65,50	23,20
1997	52,30	231,00	273,20	211,20	142,90	68,40	30,70
1999	36,20	205,60	256,90	203,60	118,30	61,70	25,80
2000	32,90	195,10	244,20	203,80	114,50	54,40	22,90
2001	24,20	162,00	242,90	197,20	118,50	57,90	21,90
2003	23,70	163,10	229,60	190,00	112,70	49,00	18,80
2005	20,30	157,60	225,50	179,90	106,60	50,10	18,10
2006	18,10	149,80	225,80	176,60	107,70	53,60	15,60
2007	16,10	150,00	225,30	173,30	102,60	52,60	18,00
Source: Pakistan Demographic Survey 2007 by Federal Bureau of statistics, Pakistan.							

Where:

$$\text{ASFR} = \frac{\#\, \textit{of births in a year to women}}{\#\, \textit{of all women in that age group}} * 1000 \qquad (8.1)$$

This table, from the various multiple and independent sources ASFRs in Pakistan shows the declining trend and the traditional reciprocal of v-shape pattern. It can also be seen that the age interval 25-29 year is the most fertile period while the age group 45-49 is the least fertile period in the reproductive life of a Pakistani women. Further, the parameters TFR and GRR (defined in the next pages) of reproductivity started to decrease modestly after the year 1996 upto 2000.

Here two different graphs also shows the different ASFRs in different time period.



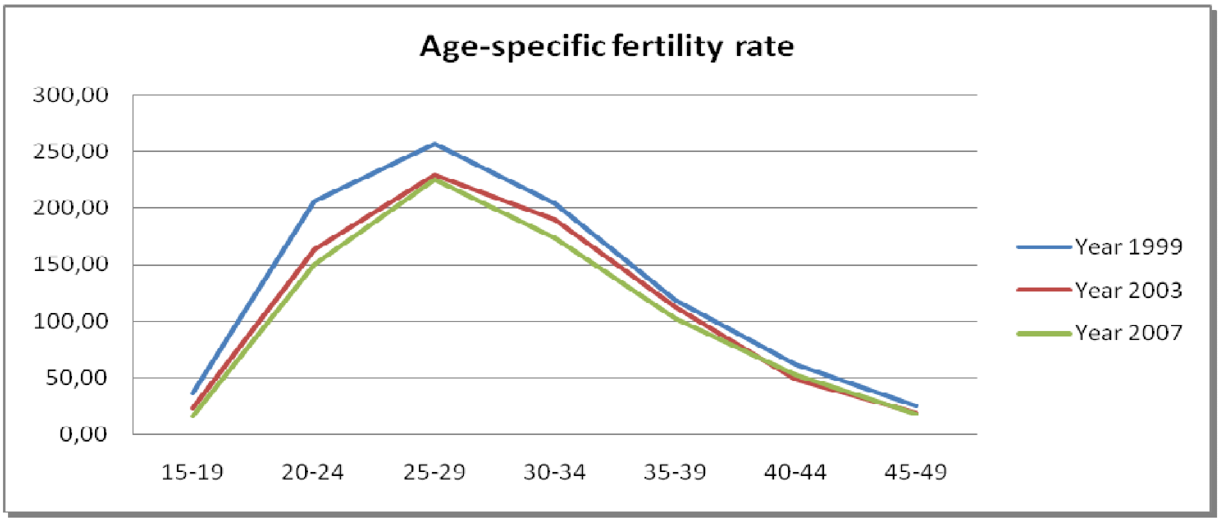


Table 8.3:

Total Fertility Rate and Gross Reproduction Rate of Pakistan during 1980s-2000s.		
Year	**TFR**	**GRR**
1984	6.95	3.33
1985	7.01	3.41
1986	6.94	3.34
1988	6.49	3.10
1989	6.40	3.07
1990	6.21	3.03
1991	5.99	2.92
1992	5.84	2.86
1995	5.59	2.70
1996	5.48	2,.46
1997	5.05	2.43
1999	4.54	2.17
2000	4.34	2.09
2001	4.12	1.97
2003	3.93	2.03
2005	3.79	2.00

The above table represents the total fertility rate (TFR) and gross Reproductivity rate (GRR) in Pakistan. Where:

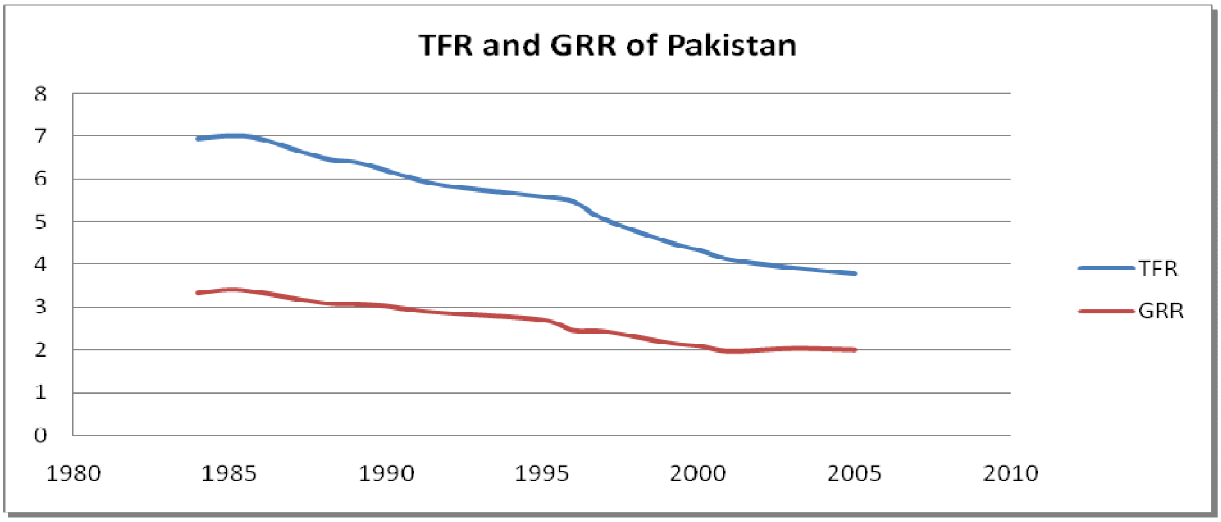
$$GRR = \frac{B^F}{B^T}\int_{\alpha}^{\beta} f_y dy \quad \text{OR} \quad GRR = 5 * \frac{B^F}{B^T}\sum_{\alpha}^{\beta} f_y \qquad (8.2)$$

and also:

$$GRR = \left(\frac{SRB}{SRB+100} * TFR\right) \text{ (By Hinde 1998)} \qquad (8.3)$$

$\int_{\alpha}^{\beta} f_y dy$ is the age specific fertility rate from 15-49 when α, β are 15 and 49 respectively and same as the Fx column given in table 5, and SRB is the sex ratio at birth (males per females). Further, B^F and B^T indicates the female and total number of births respectively. This gross reproduction rate (GRR) is quite different from the total fertility rate (TFR), GRR is the average number of daughters that would be born to a women (or a group of women) if she survived to the age of 45 and conformed to the age specific fertility rate (ASFR) of a given year. GRR is often regarded as the extent to which the generation of daughters replaces the proceeding generation of females. This GRR is the similar to net reproduction rate ([47]NRR) but it ignores the fact that some of the women will die before completing their childbearing age (15-49 years).

These TFR and GRR rates may also help us to know the TFR and GRR in a certain year, so that we can imagine/predict future population as the population's TFR is 3.79 in 2005 means that the population is increasing at the rate of 3.79 birth per year, whereas 2.00 daughters would be born to a women (if she survived till the age of 45).



[47] The net reproduction rate is the average number of daughters that would be born to a female (or a group of females) if she passed through her life time conforming to the age specific fertility and mortality rates of a given year.

8.2 Data and Methodology: -

A secondary data on age specific fertility rate (ASFR) and age specific population by males and females for the year 2007 has been taken from Pakistan demographic survey (PDS 2007) by the federal bureau of statistics, Pakistan and person year lived from age x to x+n from the life table of male and female is used to calculate projection.

Before starting the different methods, we will start with the main projection method; assume that we know the total population size at time 0 and want to estimate the total population at time t, and then the following general equation for size of the population must be attributable to the magnitude of these flows, in particular:

$$N(T) = N(0) + B\{0,T\} - D\{0,T\} + I\{0,T\} - O\{0,T\} \qquad (8.4)$$

Where:

N(T) = No. of persons alive in the population at time T.

N(0) = No. of persons alive in the population at time 0.

B{0,T} = No. of births in the population between time 0 and T.

D{0,T} = No. of deaths in the population between time 0 and T.

I{0,T} = No. of in-migrants between time 0 and T.

O{0,T} = No. of out-migrants between time 0 and T.

The choice of population projection methodology implies a set of necessary projection inputs and achievable projection outputs. One should select a methodology that will provide the desired level of detail in the output and must also select a model whose data requirement can be met. There are several methods to calculate the population forecasting in which a few of them are mentioned and described here:

- Growth rate method
- Compound rate of growth method
- Mathematical methods
- Method of least squares
- Cohort component method
- Demographic/Economic modeling

- Extrapolation method
- Ratio method

By the assumption of equation (8.4), we will see below that the two quantities are related by:

$$N(T) = N(0)e^{\int_0^T r(t)dt} = N(0)e^{\bar{r}[0,T]T} \qquad (8.5)$$

Where $\bar{r}\left[0,T\right]$ is the mean annualized growth rate between time 0 and time T and is a continuous function as it is measured in time units of years. To compute the growth rate during a very short period of time, the population change might be $\Delta N(t) = N(t+\Delta t) - N(t)$. Where $N(t)$ denotes the current and $N(t+\Delta t)$ the future populations. Further, since the person-year lived over the short time interval $[t+\Delta t]$ is now $N(t)\Delta t$, then the crude growth rate is $r(t) = \frac{\Delta N(t)}{N(t)\Delta t}$ and taking the limit of $\frac{\Delta N(t)}{\Delta t}$ is simply the derivative of $N(t)$, when Δt approaches to 0, therefore:

$$r(t) = \lim_{\Delta t \to 0} \frac{\Delta N(t)}{N(t)\Delta t} = \frac{\frac{dN(t)}{dt}}{N(t)} = \frac{d \ln N(t)}{dt} \qquad (8.5a)$$

Where ln refers to as natural log and dt is the time interval. The concept of growth rate enables us to develop a new expression for population change/forecast over a longer time interval. Integrating the above eq. (8.5a) between the exact time 0 & T:

$$\int_0^T r(t)dt = \int_0^T \frac{dN(t)}{dt} dt = \ln N(t)\Big]_0^T$$

$$\Rightarrow \int_0^T r(t)dt = \ln N(T) - \ln N(0) = \ln \frac{N(t)}{N(0)} \qquad (8.5b)$$

Taking exponentials on both sides:

$$e^{\int_0^T r(t)dt} = \frac{N(t)}{N(0)}$$

$$\Rightarrow N(T) = N(0)e^{\int_0^T r(t)dt} \qquad (8.5c)$$

So this is a most important formula in demography. The growth rate is also called the instantaneous growth rate and it expresses the change in population size during a particular discrete time period (say 0 to T) as a simple function of the set of instantaneous growth rates that prevailed during this period. Moreover, the proportionate growth in population over the period N(t)/N(0) is a simple function of the sum of growth rates. The equation (8.5) can be used as projection methodology if we can correctly estimate N(0) and make an accurate assumption about mean growth rate over the time 0 to T, then we can accurately estimate N(T).

Viewing growth rate, i.e. r (t) as a continuously varying function, raises question about the commonly encountered term, exponential growth. Any growth that occurs, either zero or negative growth must obey equation (8.5). An exponential appears in eq. (8.5) because we have defined our measure of growth-the growth rate-proportionate terms. In this sense, the term exponential growth is a redundancy; all growth is exponential by our measure of growth as the proportionate rate of change in the size of the population. When someone use term exponential growth rate, they are often (but not invariably) referring to an N(t) sequence produced by a constant and positive growth rate within some time period and such a sequence is perhaps more precisely characterized by the term geometric growth rate, is in fact constant between time 0 & T at some value r, then eq. (8.5) becomes:

$$N(T) = N(0)e^{r.T} \qquad (8.6)$$

This equation comes from the fact that:

$$\Rightarrow \int_0^T rdt = r\Big]_0^T = r.T - r.0 = r.T$$

Taking natural log of eq. (8.6):

$$\ln \frac{N(t)}{N(0)} = r.T$$

$$\Rightarrow r = \frac{\ln\left(\frac{N(t)}{N(0)}\right)}{T} \qquad (8.6a)$$

So the above equation shows that, if the instantaneous growth rate is constant during the time interval 0 to T, one can solve for its value by observing the population size at the beginning and the end of the interval (time 0 to T). Moreover, if we divide both sides of equation (8.5b) by T, the length of the time interval over which growth is occurring, we get:

$$\frac{\int_0^T r(t)dt}{T} = \frac{\ln\left(\frac{N(t)}{N(0)}\right)}{T} \qquad (8.6b)$$

The left hand side of eq. (8.6b) is simply representing the mean value of the instantaneous growth rate over the time period 0 to T, which we will designate as $\bar{r}[0,T]$. It is the area under the $r(t)$ function between the time periods 0 to T divided by the length of the time intervals, thus:

$$\bar{r}[0,T] = \frac{\ln\left(\frac{N(t)}{N(0)}\right)}{T} \qquad (8.7)$$

Note that the right hand side of equation (8.7) is identical to that of equation (8.6a), if the growth rate is constant between 0 and T; equation (8.7) provides a way of estimating its value. But this is clearly a more general expression since it requires no assumption of constancy. So the expression given in eq. (8.7) can provides the mean annualized growth rate between 0 and T.

In the lack of information about growth rate, the simplest forecast would be to assume that population size will remain constant in future. Since, in most of the populations the annual growth rate is a few percent or less, and then this assumption will often provide a fair approximation about projection for a short period (say a year or less). Then, for the short term, this assumption amounts to saying that the main component of future population size is the size of the population already at the previous date.

8.2.1Growth rate method:-

[48]For computing the annual growth rate, the following simple formula can be applied to the information at any two points of time:

$$r = \left(\frac{1}{n}\right)\left(\frac{P_n - P_0}{P_0}\right)*100 \qquad (8.8)$$

Where:

r = growth rate

Pn = population in the current year

Po = population in the base year

n = number of intermediary years

8.2.2 Compound rate of growth method:-

Another method to find the population projection is the method of compound rate of growth. A simple but slightly improved method can be computed with help of the following:

$$R = \left[\left(\frac{P_n}{P_0}\right)^{1/n} - 1\right]*100 \qquad (8.9)$$

Where, 'R' is the annual growth rate. Apply this to the following:

$$P_n = P_0 * \left(1 + \frac{R}{100}\right)^n \qquad (8.10)$$

Future population can be found as, divide the growth rate by 100 and add 1 then take the nth power (where n is years: e.g.10), finally multiply with the base year population. Population in any requisite year can project by these formulas. To apply this, let the population of Pakistan in 1991 was 110,750,020 (Po) and in 2001 it was 141,450,150 (Pn), then the value can be obtained like these:

[48] Projections of populations, Enrolment and Teachers by Arun C. Mehta

$$R = \left[\left(\frac{141,450,150}{110,750,020}\right)^{1/10} - 1\right]*100 = 2.48\% \tag{8.11}$$

Thus, during the period 1991 to 2001, population increased at the rate of 2.48% per annum. This rate can now the applied to know the population figures in any given year, for example the population in the year 2011 would be:

$$P_{2011} = P_{2001} * \left(1 + \frac{2.48}{100}\right)^{10} \tag{8.12}$$

$$P_{2011} = 141,450,150 * (1 + 0.0248)^{10} = 180.71 million$$

So this method can be used for projecting next population forecasting as well. The next populations might be:

$$P_{2021} = 180.71 * (1 + 0.0248)^{10} = 230.87 million$$

$$P_{2031} = 230.87 * (1 + 0.0248)^{10} = 294.96 million$$

$$P_{2041} = 294.96 * (1 + 0.0248)^{10} = 376.85 million \tag{8.13}$$

and so on.

Since these estimates are only includes the population increase rate and previous population, so they might be larger than the estimates of cohort component (in section 8.2.5) as these estimates are not taking into account all the demographic variables like births, deaths and migration.

8.2.3 Mathematical method:-

This is another simple method of estimating future size of a population. The future population is to take the number of individuals as determinant at more or less recent date in the past and to apply it to an assumed rate of increase as a function of time. The rate of increase may be derived from observation on the past growth of the population itself. The calculation of the net rate of population growth or assumed birth rate, death rates, and the rate of emigration and immigration may be calculated separately and added to obtain the growth rate of each future period.

8.2.4 Method of least squares:-

Another method of population forecasting is the method of least squares. This method is applicable when time series data is available. It is a less complicated method, commonly used to make future projections on the basis of past trends. It is a method in which one needs to fit a straight line to the past observations. It is also known as line of the best fit or method of least squares. The regression line which results from the method of least squares is that straight line which, when drawn through the scatter of points, minimizes the sum of squares of the vertical deviations of the points from the line. The general form of the equation is:

$$Y = a + bX$$

Where, ‘X’ as year, ‘a’ as constant and ‘b’ as slope which give the best fitting line. To compute the values of the constant of the above equation, the following two equations are used:

$$\Sigma Y = na + b\Sigma X$$

$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

Where:

ΣX = sum of all observations of X

ΣY = corresponding sum of all Y observations

ΣXY =sum of all products of X and Y

ΣX^2 =sum of all the squares of X

n = total number of observations

8.2.5 Cohort Component Method: -

The cohort component method for estimating the population projection is distinguished by its ability to preserve knowledge of an age distribution of a population (which can be of a single sex, race or etc.) over time. The model most commonly used that does account for the age distribution is called the cohort component model or method. It is a discrete time model and now nearly the only method used for population forecasting/projection representing a rare consensus for social science. It is a special case of a component method which is defined simply by the use of estimates of births, deaths and net migrating to project/update a population. The cohort component method is expressed in this simple equation:

$$P_t = P_{t-1} + B_{t-1,t} - D_{t-1,t} + M_{t-1,t} \qquad (8.14)$$

Where:

P_t = Population at time t

P_{t-1} = population at time t-1

$B_{t-1,t}$ =No. of births from time t-1 to t

$D_{t-1,t}$ =No. of deaths from time t-1 to t

$M_{t-1,t}$ = Net migrations from time t-1 to t

The method described in equation (8.4) is the same as equation (8.14)., both are running the same terms but the only difference is in migrants. Equation (8.14) is using the number of net migrants from time x-1 to x while equation (8.4) is adding in-migrants and minimizing the out-migrants, which is another way of doing it. Each component of population change are estimated or projected separately and applied to the above equation to produce a forecast. The following table 8.4 gives us the different calculations for the male and female and the first three lines of table 8.4 have been taken from the different sources of data; i.e. the age specific population from the survey PDS-2007, the life table for men and women and the age specific fertility rate in the age interval x to x+5 also from the PDS-2007, to find the population forecasting (given in table 8.5).

Table 8.4:

Mathematical Expressions for Cohort Component Method
1. Females
${}_5N_x^F(t)$ =number of women aged x to x+5 at time t

$_5L_x^F$ =number of person-years lived by women from age x to x+5 (from the life table)

$_5F_x$ =age specific fertility rate in the interval x to x+5

$_5N_x^F(t+5) = {_5N_{x-5}^F}(t) * \frac{_5L_x^F}{_5L_{x-5}^F}$ = women still alive five years later (except first and last age interval)

$_\infty N_{85}^F(t+5) = \left[{_5N_{80}^F}(t) + {_\infty N_{85}^F}(t)\right]\frac{T_{85}^F}{T_{80}^F}$ = women still alive at the last interval of age

T_x^F is the person-year lived above age x for the females.

$_5B_x[t,t+5] = 5 * {_5F_x} * \frac{_5N_x^F(t) + {_5N_x^F}(t+5)}{2}$ = Births to women age x to x+5 between time t and t+5

$B[t,t+5] = \sum_{x=\alpha}^{\beta} {_5B_x}[t,t+5]$ = Total births between t and t+5, where α, β are the from starting to ending age (e.g. 15-49)

$B^F[t,t+5] = B[t,t+5] * \frac{1}{1+SRB}$ = # of females births between t and t+5 (SRB=sex ratio at birth)

$_5N_0^F(t+5) = B^F[t,t+5] * \frac{_5L_0^F}{5 * l_0}$ = women still alive at the first interval of age, Lx is the person year lived at first age interval

2. Males

$_5N_x^M(t)$ =number of men aged x to x+5 at time t

$_5L_x^M$ =number of person-years lived by men from age x to x+5 (from the life table)

$_5N_x^M(t+5) = {_5N_{x-5}^F}(t) * \frac{_5L_x^M}{_5L_{x-5}^M}$ = Males still alive five years later (except first and last age interval)

$_\infty N_{85}^M(t+5) = \left[{_5N_{80}^M}(t) + {_\infty N_{85}^M}(t)\right]\frac{T_{85}^M}{T_{80}^M}$ = Males still alive at the last interval of age

$B^M[t,t+5] = B[t,t+5] * \frac{SRB}{1+SRB}$ = No. of male births between t and t+5 (SRB=sex ratio at birth)

$_5N_0^M(t+5) = B^M[t,t+5] * \frac{_5L_0^M}{5 * l_0}$ = Males still alive at the first interval of age

Here is the age specific population projection of Pakistan from the year 2007-2032 by the cohort component method.

<u>Table: 8.5</u>

Population Projection of Pakistan by age specific fertility from 2007-2032							
Age x	**Fx Female**	**Nx 2007 Female**	**Lx Female** $l_0 = 100,000$	**Nx 2012 Female**	**Bx 2007-2012**	**Nx 2017 Female**	**Bx 2012-2017**
00-04		9756608	378137	7422841		8635555	
05-09		10844307	466640	12040143		9160158	
10-14		9619874	463848	10779423		11968104	
15-19	0.0161	8211804	460673	9554027	715075	10705639	815452
20-24	0.150	6733861	457613	8157257	5584169	9490565	6617933
25-29	0.2253	5564656	453865	6678708	6896075	8090447	8318727
30-34	0.1733	4474911	450734	5526268	4333011	6632635	5267845
35-39	0.1026	4219507	446280	4430691	2218776	5471659	2539953
40-44	0.0526	3281389	438686	4147707	976926	4355298	1118145
45-49	0.018	2999342	431165	3225132	280101	4076597	328578
50-54		2156822	421571	2932603		3153368	
55-59		1679608	405834	2076309		2823130	
60-64		1296418	376572	1558503		1926600	
65-69		932030	336503	1158473		1392671	
70-74		606846	278227	770620		957847	
75-79		295833	216107	471355		598563	
80-84		177547	190403	260646		415291	
85 +		151288	167298	153797		193836	
		73002651		**81344504**	**21004133**	**90047964**	**25006632**

Continue…

Continued table 8.5

Age x	**Fx** **Female**	**Nx 2022** **Female**	**Bx** **2017-2022**	**Nx 2027 Female**	**Bx** **2022-2027**	**Nx 2032 Female**	**Bx** **2027-2032**
00-04		10203997		11869390		13279814	
05-09		10656708		12592244		14647421	
10-14		9105351		10592946		12516902	
15-19	0.0161	11886184	909321	9043025	842401	10520439	787429
20-24	0.150	10634527	7546909	11807230	8415659	8982957	7796320
25-29	0.2253	9412834	9858723	10547427	11242617	11710525	12536792
30-34	0.1733	8034635	6354595	9347899	7530983	10474665	8588126
35-39	0.1026	6567094	3087940	7955239	3724978	9255526	4414561
40-44	0.0526	5378552	1280001	6455347	1556158	7819871	1877191
45-49	0.018	4280629	376075	5286340	430514	6344674	523396
50-54		3985887		4185379		5168712	
55-59		3035655		3837096		4029141	
60-64		2619573		2816774		3560429	
65-69		1721601		2340838		2517056	
70-74		1151486		1423452		1935449	
75-79		743988		894393		1105637	
80-84		527369		655497		788013	
85 +		284891		379897		484257	
		100230961	**29413565**	**112030415**	**33743309**	**125141489**	**36523816**

Continue…

Continued table 8.5							
Age x	Lx Male l_0 =100,000	Nx 2007 Male	Nx 2012 Male	Nx 2017 Male	Nx 2022 Male	Nx 2027 Male	Nx 2032 Male
00-04	372613	9783859	8338421	10126190	11864820	13450395	14132681
05-09	458581	11710324	12041158	10262234	12462469	14602231	16553625
10-14	456805	10636015	11664972	11994524	10222490	12414205	14545680
15-19	454001	9063876	10570728	11593369	11920899	10159741	12338003
20-24	450089	6824723	8985775	10479643	11493472	11818179	10072198
25-29	445474	5268436	6754746	8893639	10372190	11375624	11697001
30-34	440477	3957414	5209339	6678976	8793877	10255842	11248020
35-39	435226	4132910	3910237	5147237	6599355	8689044	10133581
40-44	426469	3496263	4049753	3831561	5043672	6466572	8514215
45-49	415073	3277150	3402837	3941537	3729175	4908896	6293774
50-54	396804	2429295	3132910	3253064	3768054	3565039	4692836
55-59	366921	1864568	2246347	2896973	3008079	3484285	3296559
60-64	328821	1637251	1670957	2013093	2596160	2695729	3122487
65-69	285031	1106476	1419214	1448431	1745004	2250422	2336731
70-74	229126	857310	889456	1140854	1164341	1402745	1809032
75-79	167973	358255	628497	652063	836364	853582	1028356
80-84	132878	250734	283404	497183	515826	661620	675241
85 +	106548	202880	201865	215952	317355	370777	459431
		76857739	**85400614**	**95066522**	**106453600**	**119424927**	**132949450**

Fx = Age specific fertility rate, Lx = Person year lived from age x to x+n, Nx 2007 = Population at 2007 and so on.
Source: Demographic reports and surveys by Federal Bureau of Statistics, Pakistan

Different columns are used to find projection, first column for age and second one for ASFRs and third one is for number of females in the year 2007 (base year) and forth one represents the person year lived from age x to x+n (taken from the life table).

According to our calculations, the births in each age group (15-49) can be found by the following formulas:

$$_5B_x\left[t, t+5\right] = 5 *_5 F_x * \frac{_5N_x^F(t) +_5 N_x^F(t+5)}{2}$$

$$\Rightarrow_5 B_x\left[2007, 2012\right]_{(15-19)} = 5 * 0.0161 * \frac{8211804 + 9554027}{2} = 715075$$

and similarly the last one:

$$_5B_x\left[2007, 2012\right]_{(45-49)} = 5 * 0.018 * \frac{2999342 + 3225132}{2} = 280101$$

Then the total number of births for the period 2007-12 can be found by adding all the birth in each age group (column 6 table 8.5):

$$B\left[2007, 2012\right] = \sum_{x=\alpha}^{\beta} {}_5B_x\left[2007, 2012\right]$$

$$\Rightarrow B\left[2007, 2012\right] = 21004133$$

$$\Rightarrow B\left[2012 - 2017\right] = 25006632$$

$$\Rightarrow B\left[2017 - 2022\right] = 29413565$$

$$\Rightarrow B\left[2022 - 2027\right] = 33743309$$

$$\& B\left[2027 - 2032\right] = 36523816$$

According to PDS-2007, the total number of births was 3.8 million in a year 2007 and our estimates indicate that in the period of 2007-12 there possibly be around 21 million births, so it looks reasonable to say that there might be around 4 million births each year, which is almost the same as PDS-2007. In the same way, by using above formulation, we can find the next number of birth for the year 2012-17 in column 8 and so on.

From the total number of births, the number of female and male births can be separated by the following calculations:

$$B^F\left[2007,2012\right]=B\left[2007,2012\right]*\left(\frac{1}{1+SRB}\right)$$

$$\Rightarrow B^F\left[2007,2012\right]=21004133*\left(\frac{1}{1+1.14}\right)=9815016$$

$$B^M\left[2007,2012\right]=B\left[2007,2012\right]*\left(\frac{SRB}{1+SRB}\right)$$

$$\Rightarrow B^M\left[2007,2012\right]=21004133*\left(\frac{1.14}{1+1.14}\right)=11189118$$

Where the sex ratio at birth for the period 2007-2012 is 1.14 that is calculated as; the number of male births in the year 2007-12 divided by the number of female births in the year 2007-12 and so on. So, here we can see that the number of females in the period of 2007-12 might be 9,815,016 and male might be 11,189,118 which states 1.9 million females and 2.2 million males each year. It also looks reasonable as in the data given by statistical bureau of Pakistan; PDS-2007 says that there are total 1.8 million females and 2.0 million males in a year. Finally, our estimates made by the cohort component method looks quite logical as described earlier that the cohort component method gives almost accurate forecasting that is why it is now-a-days the only method on which demographers are relying much.

Further, the procedures of finding number of females and males at the first and last interval (i.e. 00-04 & 85+) are quite different from the rest of calculations as there is a high probability of dying in the infancy and at the open ended interval. So here is some overview to find them:

$${}_5N_0^F(2012)=B^F\left[2007,2012\right]*\frac{{}_5L_0^F}{5*l_o}$$

$$\Rightarrow {}_5N_0^F(2012)=9815016*\frac{378137}{5*100,000}=7422841$$

In the same way:

$${}_5N_0^M(2012)=B^M\left[2007,2012\right]*\frac{{}_5L_0^F}{5*l_o}$$

$$\Rightarrow {}_5N_0^M(2012) = 11189118 * \frac{372613}{5*100,000} = 8338422$$

and:

$${}_5N_\infty^F(2012) = \left[{}_5N_{80}^F(2007) + {}_5N_{80}^F(2007)\right] * \frac{T_{85}^F}{T_{80}^F}$$

$$\Rightarrow {}_5N_\infty^F(2012) = [177547 + 151288] * \frac{167298}{357701} = 153797$$

Similarly:

$${}_5N_\infty^M(2012) = \left[{}_5N_{80}^M(2007) + {}_5N_{80}^M(2007)\right] * \frac{T_{85}^M}{T_{80}^M}$$

$$\Rightarrow {}_5N_\infty^M(2012) = [250734 + 202880] * \frac{106548}{239426} = 201865$$

It says that in the year 2012 there might be 7,422,841 female and 8,338,422 male babies at the first age interval, whereas on the age of 85+ there may be 153,797 and 201,865 female and males left respectively. In general we know that the dying ratio of male is higher than the females which means that the males die early whereas females die late but cohort component estimates indicate that there are more males in the year 2012 (i.e. 201,865) then the females (i.e.153,797) at the age of 85+. This is due to the data (base year 2007) given by the statistical bureau of Pakistan which states that there are 202,880 males in 2007 while 151,288 females at the age of 85+ (see column 3 of table 8.5 in females and in males columns). But in spite of all these figures, our estimates shows that the number of males might be higher in the later years, like in the year 2032 the number males might be 132,949,450 and females 125,141,489. More information can see in the table 8.5 and rest of the calculation can also be found in the following table 8.6 (by using table 8.4).

Table 8.6:

Calculations of table 8.5	
$B^F[2007-2012]=21004133*\frac{1}{1+1.14}=9815016$ SRB = Male Births(2007,2012) / Female Births(2007,2012)=1.14	$B^M[2007-2012]=21004133*\frac{1.14}{1+1.14}=11189118$ SRB = Male Births(2007,2012) / Female Births(2007,2012)=1.14
$B^F[2012-2017]=25006632*\frac{1}{1+1.19}=11418554$ SRB = Male Births(2012,2017) / Female Births(2012,2017)=1.19	$B^M[2012-2017]=25006632*\frac{1.19}{1+1.19}=13588079$ SRB = Male Births(2012,2017) / Female Births(2012,2017)=1.19
$B^F[2017-2022]=29413565*\frac{1}{1+1.18}=13492461$ SRB = Male Births(2017,2022) / Female Births(2017,2022)=1.18	$B^M[2017-2022]=29413565*\frac{1.18}{1+1.18}=15921104$ SRB = Male Births(2017,2022) / Female Births(2017,2022)=1.18
$B^F[2022-2027]=33743309*\frac{1}{1+1.15}=15694562$ SRB = Male Births(2022,2027) / Female Births(2022,2027)=1.15	$B^F[2027-2032]=36523816*\frac{1}{1+1.08}=17559527$ SRB = Male Births(2022,2027) / Female Births(2022,2027)=1.15
$B^F[2027-2032]=36523816*\frac{1}{1+1.08}=17559527$ SRB = Male Births(2027,2032) / Female Births(2027,2032)=1.08	$B^F[2027-2032]=36523816*\frac{1}{1+1.08}=17559527$ SRB = Male Births(2027,2032) / Female Births(2027,2032)=1.08
$B[2007-2012]=21004133$	$B[2012-2017]=25006632$ and so on…
$T_{80}^F=357701$, $T_{85}^F=167298$	$T_{80}^M=239426$, $T_{85}^M=106548$, $l_0=100,000$
$N_x^F(2012)at_{(05-09)}=9756608*\frac{466640}{378137}=12040143$ and so on till: $N_x^F(2012)at_{(80-84)}=295833*\frac{190403}{216107}=260646$	$N_x^M(2012)at_{(00-04)}=9783859*\frac{458581}{372613}=12041158$ and so on till: $N_x^M(2012)at_{(80-84)}=358255*\frac{132878}{167973}=283404$

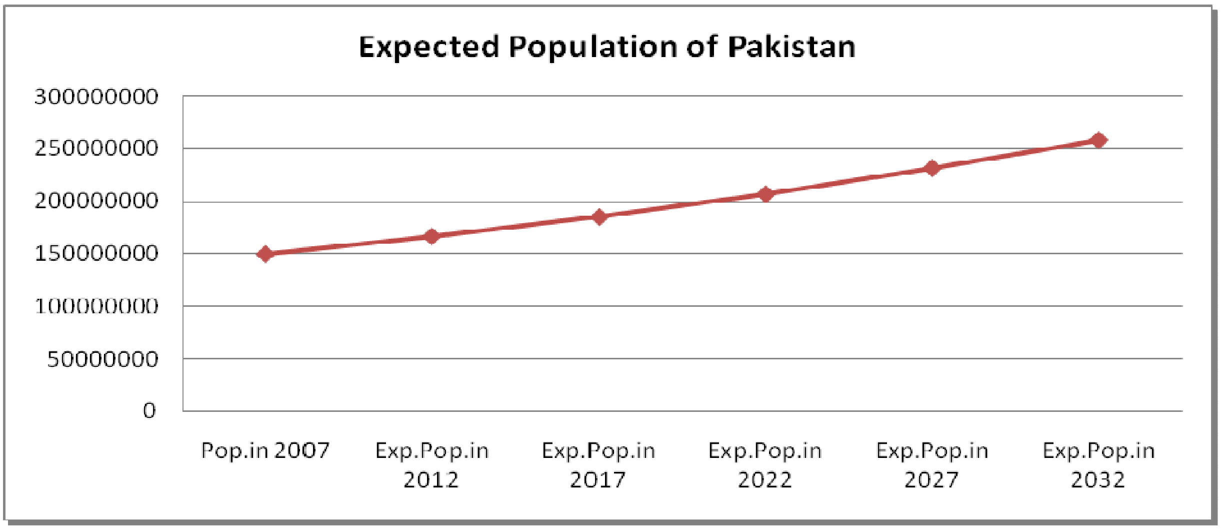
As mentioned earlier, the data has been taken from PDS 2007 and male and females life tables to make prediction. Then, the mathematical expression (table 8.4 & table 8.5) has been used to obtain the future population in a five year age interval. Prediction has been obtained for the year 2007-2032 at five years interval. After each births column, sex ratio at birth has been calculated and used for the next 5-years births. Our estimate shows that the number of births in the year 2007-2012 can be 21,004,133 from which female births might be 9,815,016 and male births might be 11,189,118, where the sex ratio at birth is 1.14. Similarly, the number of births in the year 2027-2032 might be 36,523,816 births whereas female birth may stand at 17,559,527 and male births at 18,964,289 and the sex ratio at birth might be more or less 1.08. So, same as the general perception, I have also found more male births then the female births throughout the projection period (see table 8.6 column 1 & column 2). Finally, the results are shown in the table 8.7 that illustrate the future population of Pakistan. It is almost the same as equation (8.13) results but due to the lack of information about migrants, (I have not found any reliable data on age specific migration for Pakistan), so the migration factor is missing in the calculation and table 8.7 is only relying on base year population 2007 i.e number of males and females, births, and deaths. Similarly, to find the future population for the males and females separately for each age group, table 8.4 is used and results are shown in table 8.5.

Table 8.7:

Population Projection for Pakistan			
Year	**Male Population**	**Female Population**	**Total Population**
2007	76857739	73002651	149860390
2012	85400614	81344504	166745118
2017	95066522	90047964	185114486
2022	106453600	100230961	206684561
2027	119424927	112030415	231455342
2032	132949450	125141489	258090939
Source: cohort component method			

These results are based on the cohort component method calculation. According to the equation (8.14), the expected population of Pakistan in 2032 might be 296.02 million which is relatively higher then the estimates of cohort component method (equation 8.13) is 258.09 million in 2031 but we know that the cohort component estimates are the more reliable than any other method. But on the other hand, in our calculations since I have mentioned that the migration data is not included, so it might be possible that there is a little-bit change in the projection. However, there might not be too much difference from 258 million because according to some sources (Google and CIA factbook), Pakistan's net migration rate was-0.48 per 1000 people in 2009 where as it was -0.51 per 1000 people in 2008.

Finally, Since the TFR of Pakistan is high and according to the FBS statistics reports, if the population of Pakistan will keep this increase then within just 37 years it will be more than double. So, it is quite possible that Pakistan should keep its national family planning programms in place and not only encourage total fertility to fall below 1.85 but would aim to achieve a considerable period of declining population. [49]As the United Nations report says that the population of Pakistan will be increased around 206 million (144%) in the next 50 years, which looks relatively high. It might be noted that UN analysis was also carried out after the projection were constructed which shows that Pakistan has betatedly and very recently experienced quite steep fertiltiy decline but there is signigicant difference from the past. However, the new data and analysis indicates that Pakistan's attainment of replacement fertiltiy may be far behind that of other south asian countries (like india and bangladesh). There will probably be some gap in the product of its lowere levels of girls schooling, or rural-urban policies or a higher infant mortality rates and related phenomenas.



Moreover to end this, the data given in the PDS-2007 indicates us that there are 97.05 million people residing in the rural areas whereas this number is 52.81 for the urban one. So Pakistan is still considered as a rural and agriculture country. The following graphs will be

[49] World population to 2300 report by United Nations.

the enough information to overview the inhabitants of Pakistan by the urban and rural areas that makes grounds of 258 million population in the year 2032. The record says that the 65% of the population is living in the rural areas whereas 35% of the population is residing in the urban areas of Pakistan. Like many other developing nations, Pakistan's government has also much emphasis on the urban areas instead rural population where there are more people residing, having less attention. Then, off course urban have much better facilities as well then the rural one, either of education system, quality health care centers or sanitation facilities or etc. that really effect the future generation/population.

8.3 Comments: -

By using the data from the different sources either by Pakistan demographic surveys, federal burear of statistics, Pakistan or life table for male and females, first I have tried to give some overview of Pakistan's fertility pattern to get know about fertility behaviour of Pakistani women. I have shown in the different data tables, (like ASFRs table or TFR and GRR table) that Pakistan's total fertility rate is still high specially in the rural areas even though many family program policies have been made but the need is to implement them. As I have mentioned before that Pakistan should implement its family planning programms in place. So that there should have some significant change in the fertility rates.

Second, I have made an effort to predict the Pakistan's future population. First I have described some methods to calculate/project/predict the future population and then applied the Pakistani data on these methods. The first method (equation 8.13) shows that there may be 294.96 million population in 2032 whereas the cohort compoenent methods states that it might be 258.09 million in 2031, that looks reasonable to say that the first growth rate methods's estimates are not considered to be the best estimates by the demographer as it does not take into an account all the demographic variables like births, deaths, and migration while the cohort component method includes all of these demographic variables (i.e. births, deaths, and migration). According to the data (use in table 8.5), the number of females in 2007 were 73,002,651 while the men was 76,857,739 and then by the cohort component calculations, this number may becomes 90,047,964 and 95,066,522 for the females and males respectively in the year 2017, whereas the calculations done via equation (8.13) doesn't categories the population to understand that how many males and females will be in future, and also cohort component method gives population in various age groups which is another aspect of cohort method.

So to conclude this, I can say that since there is a large amount of population living in the rural areas where there are not enough facilities of health, food, sanitation and etc. and the population may becomes double as stated by the FBS in next 37 year, if the TFR will remain high, so the government is really needed to take in place the family planning policies in order to reduce the TFR not only in the urban but especially in the rural areas of Pakistan.

Further in the end, I must mentioned that I have used mostly the data from the Pakistan demographic surveys (like 2007, as it was given continues) and have made the projections but if one start making these projection by the data given the Pakistan statistical year book 2010, there might have big change in projections, as the population given in [50]Pakistan statistical year book 2010 is 158.17 million in the year 2007 while it is 149.86 million in the [51]Pakistan demographic survey 2007, that has almost 8.3 million difference in the two publications of Federal Bureau of Statistics, Pakistan.

[50] http://www.statpak.gov.pk/fbs/sites/default/files/other/yearbook2010/Population/16-1.pdf
[51] http://www.statpak.gov.pk/fbs/sites/default/files/population_satistics/publications/pds2007/tables/t01.pdf

References: -

- Abou Zahr, C. *et. al., Who Counts?4, The way forward*, Series, 2007.
- Anderson, C. J., *Multiple Logistic Regression For Dichotomous Responses* University of Illinois at Urbana-Champaign.
- Banda, J. P., *Main sources of socio-demographic statistic* United Nations 2003.
- Bongaarts, J., *Fertility transitions in developing countries, progress or stagnation* New York: Population council, 2008.
- Coale, A. J., Demeny, P., with Vaughan, B., *Regional Model Life Tables and Stable Populations,* New York: Academic Press, 1983.
- *Countdown to 2015: Maternal, Newborn & Child Survival,* WHO & UNICEF Decade Report (2000-2010).
- *Country Health Profile,* WHO Countries Pakistan, http://www.who.int/countries/pak/en/
- *Country profile: Maternal, Newborn & Child Survival,* WHO & UNICEF Report.
- Daniel, W. W., *Bio Statistics: A Foundation for the Analysis in the Health Sciences* 8th ed. New York: J. Wiley & Sons, 2006.
- *Demographic and Research Statistics,* Federal Bureau of Statistics, Pakistan, http://www.statpak.gov.pk/fbs/population_publications
- Dobson, A. J., Barnett, A. G., *An Introduction to Generalizes Linear Models* 3rd ed. Boca Raton, Fla.: Chapman & Hall/CRC, 2008.
- Fergany, N., *On the Humans Survivorship function and Life Table Construction*, Springer, 1971.
- Hudson V. M., Den Boer, A. M., *Missing Women and Bare Branches, Gender Balance and Conflict*, ECSP report, Issue 11, 2005.
- Keyfitz, N., Flieger, W., *Population, facts and method of demography,* San Francisco, W. H. Freeman, 1971.
- Larget, B., *Multiple Linear Regression* University of Wisconsin-Madison, 2008.
- Masset, E., White, H., *Infant and Child Mortality in Andhra Pradesh* Munich Personal RePEc Archive, 2008.

- Mehta A., *Projections of population, enrolment and teachers*, NIEPA, New Delhi.
- *Methods for Population Projections by Sex and Age,* Manual III, United Nations Population Division, 1956.
- Nasir, J. A., *et al*., *Levels of recent and targeted and determinant of fertility in Pakistan* Pakistan Academy of Sciences, 2009.
- Nasir, J. A., *et al*., *Reproductivity and Age-Specific Fertility Rates in Pakistan,* Pakistan Journal of Statistics, 2009.
- National Database and Registration Authority (NADRA) Pakistan http://www.nadra.gov.pk/
- *Pakistan Statistical Year Book* 2009 &2010, Federal Bureau of Statistics, Pakistan, http://www.statpak.gov.pk/fbs/content/pakistan-statistical-year-book-2010
- Poston, D. L., Bouvier, Leon F., *Population and Society: An Introduction to Demography* Cambridge: Cambridge University Press, 2010.
- Preston, S. H., *et al., Demography: Measuring and Modeling Population Processes* Malden, Mass. Blackwell Publishers, 2001.
- Swanson, D. A., McKibben, J. N., *New Directions in the Development of Population Estimates in the United States*, Population Research and Policy Review, 2010.
- *The Millennium Development Goals,* United Nations report, 2010.
- *The State of the World Children, special edition,* UNICEF Report, 2009.
- *The World Factbook,* CIA.
- United Nations, Pakistan, http://unportal.un.org.pk/sites/unpakistan/pages/default.aspx
- *World Mortality Report 2005,* New York: United Nations, 2006.
- *World Population by 2300,* New York: United Nations, 2004.