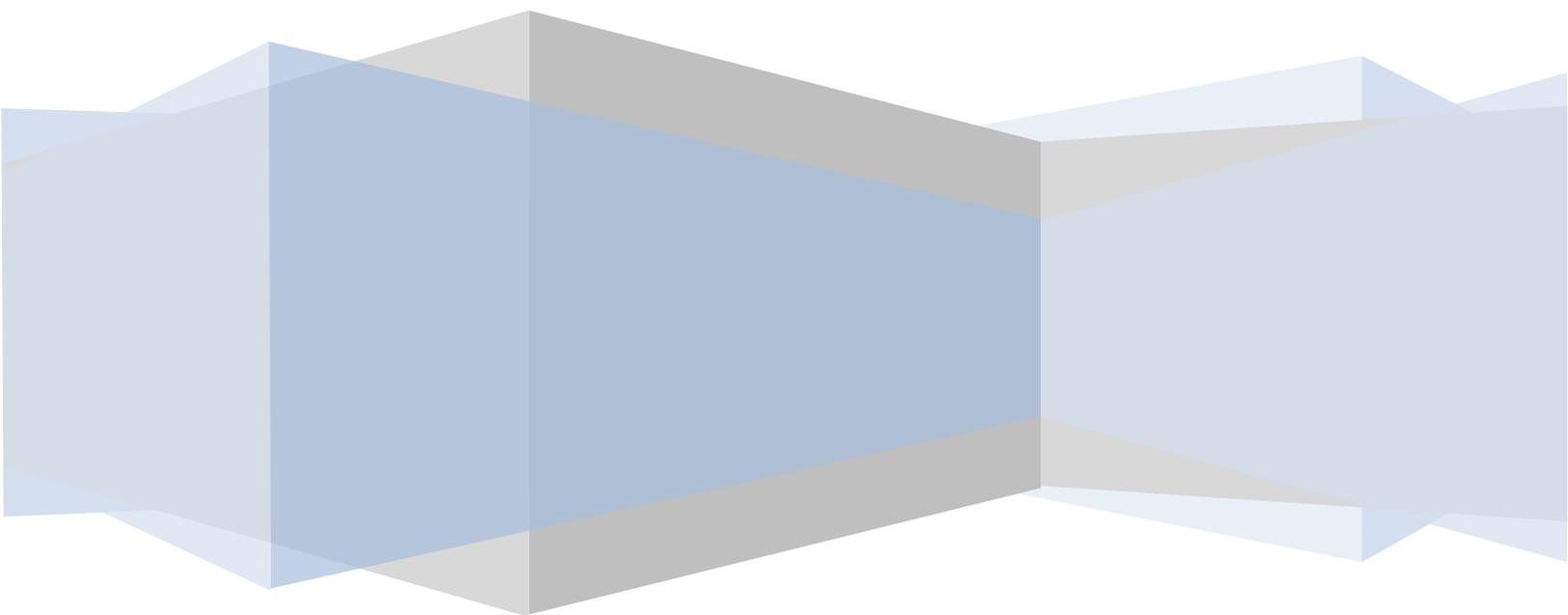


DEMOGRAPHY

STAT H-203

Compiled by: Abdullah Adil Mahmud

Sponsored by, stat.bishwo.com.



Demography

Chapter One: Basic Concepts

Definition: The origin of the word “Demography” (demos =people, graphin = to write) suggests the general nature and scope of the subject matter. That is demography is concerned with the description, analysis and understanding of the population phenomena.

According to **Frank Lorimer**, “*In broad sense, demography includes both demographic analysis and population studies. A broad study of demography studies both qualitative and quantitative aspects of population.*”

According to **Peter R Cox**, “Demography is the study of statistical methods of population involving primarily the measurement, size, growth of the people, the proportions of leaving being born or dying within the same area or region and related functions of-

- Fertility
- Mortality
- Marriage”

According to **Donald J Bogue**, “Demography is the mathematical and statistical study of the size, composition and spatial distributions of human population and of changes over time in these aspects through the operations of five processes-

- Fertility
- Mortality
- Marriage
- Migration and
- Social Mobility

Demography and Economics:

Economics studies economic problems of the people and in these demographic and population studies play a vital role.

When there is more population, economic activities are bound to increase. Population problems are directly linked and concerned with

- Per Capita Income
- Employment
- Education
- Transportation
- Industrialization etc.

Demography influences economics in two important and different ways. On one hand, it as a means of labor force becomes source of production. On the other hand, changes in the population influence labor force and the source of production.

Rehabilitation:

Sponsored by Stat Word: <http://stat.bishwo.com/>

The process of helping a person who has suffered an injury restore lost skill & so regain minimum self-sufficiency.

Some Basic Concepts of Demography:

1. **Live Birth:** Live birth is the complete expulsion or extraction from its mother. Live birth is the beginning of life and influences population. Live birth is very important both in growth of population & social life.
2. **Still Birth:** When one thinks of birth in demography one includes in it both live birth & still birth. Live birth is of course somewhat simpler in counting because in it are counted all babies who take birth. But some problems arise in the case of still birth. In it are included all those children who did not come out of the womb of the mother but the woman remained pregnant for 28 weeks.
3. **Family Size:** Though commonly from family size, we mean total number of persons of a family, yet in demography it does not mean so. In demography, it means number of children a woman or couple has at a point of time. The complete family size indicates the total number of children born by a woman up to the end of her reproductive period.
4. **Child Death:** Usually due to social or other causes there are many deaths in many countries within a period of one year of the birth of the child. In some cases, whereas newly born baby dies, in some both the baby as well as mother dies. This increases death rates as well.

Deaths which occur within a period of **4 months** from the birth of a child are known as **Neonatal** period of deaths. The deaths which takes place between a period of **4-12 months** are known as **Post Neonatal** Period of deaths.

5. **Sterility:** A man or woman or couple who has not given birth to child is called sterile. It is used both in connection with individuals or groups consisting either men or women or both. But usually sterility measures are computed only for women.
6. **Parity:** In demography women are classified according to the number of children born alive to them. As for example, first parity women are those who have given birth to one child, second parity women are those who have given birth to two children & so on.
7. **Adulthood:** Adulthood means that it is at this stage that boy or girl becomes capable to produce children.
8. **Reproductive span:** It indicates child bearing period of a woman. Only women can conceive & give birth to children & that too within a limited period & certain age limits.
9. **Fecundity:** Fecundity is that stage in a life of girl by which she becomes capable of becoming mother, no matter whether she actually becomes or not. If she is capable of giving birth to a child she is called **fecund**.

Sponsored by Stat Word: <http://stat.bishwo.com/>

10. **Fertility:** Fertility is the standard of measuring the capacity of the women to produce children. According to **Thomson & Lewis**, “**the born fertility is generally used to indicate the actual reproductive performance of a women or a group of women. The crude birth rate is only one measure of fertility.**”

Thus it is obvious that a **fecund** woman may not be **fertile** but a fertile woman **must be fecund** because if she is not then fertility has no meaning.

11. **Cohort:** In demography the children who are born in one year are called cohort.

12. **Relative Number:** Relative numbers are needed only when comparison is desirable. Both the ratios & rates are relative figures & express a relation between two or more numbers.

13. **Ratio:** Ratio expresses the size of one to another. As for example male female ratio in a country or a region etc.

14. **Rates:** Like other ratios, a rate is also computed & that help in expressing what has happened in terms of a certain unit of time. Two most commonly used rates in Demographic studies are **Birth & Death Rates**.

15. **Birth Rates:** The procedure for calculating birth rate is by finding out total number of births in a given area during a given year and by dividing that by total number of people living in an area at the **mid-point** of that year and multiplying that by 1000.

16. **Death Rate:** Death rate is also computed by the similar way as the birth rate.

17. **Crude Rate:** Both birth rate & death rate are however called CRUDE Rate.

Because, these (birth rate & death rate) use total population as a base number rather than parts of the population.

Usually midyear population is used because it is probably the figure which is both nearest to the average population for the year a readily obtainable.

Chapter Two: Sources of Demographic Data

Methods & Sources of Data Collection:

A **census** of population may be defined as “**Total process of collecting, compiling & publishing demographic, economic & social data pertaining at a specific time or times to all persons in a country or delimited territory**”.

According to V M Dandakar, “**A census of population is the total process of collecting, compiling, evaluating, analyzing & publishing demographic, economic & social data pertaining at a specific time to all persons or in a delimited part of the country.**”

There are two important census **techniques**:

1. **De Facto Method &**
2. **De Jure Method**

Note: A *De Jure* count of the total population is the enumeration of persons who usually reside in a given place. A *De Facto* count is the enumeration of persons physically present at specified place on census night.

1. **De Facto Method:** It is one of the census techniques. Under this system, a date is fixed for taking census of the whole country. Usually such an operation is conducted at night because it is felt that after day long work people will come back to their homes at night. Such a night is called census night. This night is very carefully selected.

At this night all those who are found anywhere are counted. When the people actually present are counted at the same moment in a census technically that is called **De Facto Population**.

Since the census is completed on a particular date, it is also called **Date System** or **One Night Enumeration System**.

Merits of De Facto Method: This method is quiet simple as well as clear.

Disadvantages:

1. Under this system floating population is not counted & the data becomes inaccurate.
2. This system needs a large number of well qualified & trained field workers.
3. Night time is the rest time & enumerator/invigilator is not welcomed at home.

2. **De Jure Method:**

Under this, every person in an area is personally counted & information obtained from him/her. Temporary resident is not included in it. Whoever is enumerated, his/her permanent place of residence is considered. Usually census work is completed within 2 or 3 weeks.

Merits of De Jure Method:

1. Since period is 2/3 weeks so chances of inaccuracy is little bit short.

2. Many question regarding gender, age, social conditions can be asked.

Demerits:

1. Under this system permanent residence, household etc. are to be defined. It is difficult to define uniformly.
2. It becomes difficult to collect information in respect of those who have no permanent residence.

Vital Statistics:

Vital Statistics is defined as that branch of **BIOMETRY** which deals with data & the laws of human mortality, morbidity & demography.

The term vita statistics refers to the numerical data or the techniques used in the analysis of the data pertaining to vital events occurring in the given section of the population.

Vital Events:

By Vital events, we mean such events of human life as-

- Fertility
- Mortality (Deaths)
- Marriage
- Divorce
- Separation etc.

According to **ARTHUR NEWSWHOLME**, “Vital Statistics forms perhaps the most important branch of statistics as it deals with mankind in the aggregate. It is science of numbers applied to life history of communities & nations.”

Use of Vital Statistics:

1. Through vital statistics we can study the population trend.
2. Vital Statistics is used in public administration.
3. Vital statistics provide guidelines for the researchers & pharmaceutical proposition.
4. The facts & relating to
 - births
 - deaths
 - marriages etc.

are of extreme importance to vital official agencies to a variety of administrative purposes.

5. The whole of actuarial science, including life insurance is based on the mortality.

Methods of obtaining vital statistics:

1. Registration Method
2. Census Method

Registration Method: The most important source of obtaining vital statistics is the registration method which consists in continuous & permanent recording of vital events pertaining to

- births
- deaths
- marriages
- migration etc.

[+p.49 Kpedekpo]

These data in addition to their statistical utility also have their values as legal documents. Registration of births provide information on place of birth, sex (gender), age, religion of the parents, number of previous issues and their sexes, father's occupation, birth place of parents etc. Similar Information is also obtained with respect to

- Deaths
- Marriages
- Migration etc.

Chapter THREE: Demographic Rates & Ratios

Rates and Ratios of Vital Events:

Generally the vital events i.e. n_1 persons die, n_2 persons are born, n_3 persons get married during a given time 't' in any community or region or country. These are generally transferred to rates or ratios.

Sex ratio:

The ratio of the total number of males (mP) to the total number of females (fP) in the given population is termed as sex ratio.

$$\text{Sex ratio} = \frac{mP}{fP}$$

Sex ratio at birth:

It measures or relates male births to female births in a given population in an area (district, region or country)

$$\text{Sex Ratio at birth} = \frac{\text{Male births}}{\text{Female Births}} \times 100$$

A sex ratio at birth with value more than 100 indicates an excess of male over female births. A value less than 100 indicate an excess of female over male.

Example:

In an African country in 1962, Male birth = 16,517; Female birth = 15,986

$$\text{Sex ratio at birth} = \frac{16517}{15968} \times 100$$

$$= 103$$

Interpretation: The sex ratio at birth is therefore about 103 male births for 100 female births in this population.

Important uses of sex ratio at birth:

Firstly, it is needed in situation where total births have to be separated into male births and female births.

Secondly, the sex ratio at birth can also be used as a first approximation for the evaluation of the quantity at the registration.

General Sex Ratio:

$$\text{General Sex Ratio} = \frac{\text{All males}}{\text{All females}} \times 100$$

Dependency Ratio: The dependency ratio is defined as the ratio of youths under 15 years of age plus persons aged 65 & over to the adults aged (15-64) years.

$$\text{DR} = \frac{\text{Children under 15 years} + \text{Parents aged 65 \& over}}{\text{All females Persons aged (15-64) years}} \times 100$$

CRUDE DEATH RATE (CDR):

This is the simplest of all the indices of mortality & is defined as the number of death (from all causes) per k persons in the population of any given region or community during a given period.

Thus, in particular, the annual death rate (CDR), say, m for any given region or community is given by,

$$m = \frac{\text{Annual deaths}}{\text{Annual mean population}} \times k;$$

Where, $k = 1000$ (usually)

##Most serious drawback of CDR is that it completely ignores the age & sex distribution of population.

Experience shows that mortality is different in different segments of the population. Children in the early ages of their life and older generation are exposed to higher risk of mortality as compared to younger people. Moreover, the mortality rate is different for female irrespective of age group than their male counterparts.

Specific Death Rate (SDR):

In order to arrive at a more useful figure than CDR, we must take into account the fact that mortality pattern is different in different segments of the population. Various segments may be

- Age
- Sex (Gender)
- Occupation
- Social status etc.

As for example, the people engaged in infant child welfare work would be interested to know the mortality condition in the age groups

- Below 1 year
- 1-4 year
- 5-9 year

Another example, those engaged in maternal health programs would like to know the number of deaths occurring among women in the reproductive period (usually 15-49 years). Insurance authorities would be interested in the mortality pattern at different years of population.

Death rate computed for a particular specified section of the population is termed as “Specific Death Rate (SDR)”.

SDR for given geographical region during a given period is defined as-

$$\text{SDR} = \frac{\text{Total number of deaths in the specific section of the population in the given period}}{\text{Total population of the specified section in the same period}} \times k;$$

Where, $k = 1000$ (usually)

Usually SDR is computed specific to

- Age & Sex

Some exercises:

In 1969, in a census in Uganda, it has been found that the population aged (0-14) as enumerated was 4, 404, 229, 1. Those aged 65 years and over were 365,465 & the population (15-64) years were 4,766,566. Find DR & interpret.

Solution: It is given that population aged (0-14) was 4,404,291 & aged (65-over) was 365,465.

Sponsored by Stat Word: <http://stat.bishwo.com/>

$$\text{Now, DR} = \frac{4404291+365465}{4766556} \times 100 = 100.070$$

Interpretation:

100.07 dependents per 100 persons in the (15-64) age group.

Exercise 2:

Age Specific Death Rate (ASDR):

nD_x = Number of deaths in the age group (x,x+n)

That is the number of deaths among the persons with age x or more but less than x+n, in a given region during a given period, t, say. Then the ASDR for the age group x to x+n, usually

$$n m_x = \frac{n^D_x}{n^P_x} \times 1000$$

n^P_x = Total population of the age group x to (x+n)

ASDR for male is

$$m m_x = \frac{m^D_x}{m^P_x} \times 1000$$

Where, m^P_x is the number of males in the population in the age group x to (x+n) & m^D_x is the number of deaths among the population.

ASDR for **female** is $f m_x = \frac{f^D_x}{f^P_x} \times 1000$

Where, f^P_x is the number of females in the population in the age group x to (x+n) & f^D_x is the number of deaths among the population.

Merits:

1. Death rate specific to age & sex overcome the drawback of CDR, since they are computed by taking into consideration the age & sex composition of the population.
2. ASDR specific for age & sex is one of the most important & widely applicable types of death rates.

Demerits of ASDR:

1. The death rate specific to age & sex are not of much utility for overall comparison of mortality conditions prevailing in two different regions, say A & B. It might happen that for certain age groups the mortality pattern for region A is greater than that for B.
2. In addition, the death rates specific to age & sex completely ignores social, occupational & topographical factors.

Problem: [From the following data calculate CDR- to be edited]

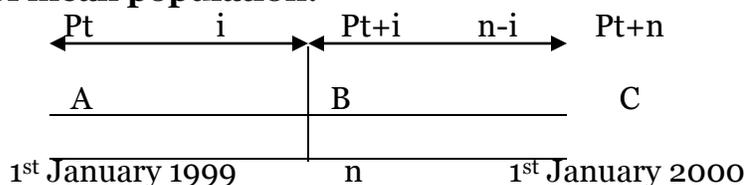
Mid-Year Population:

The midyear population is the population as at 30th June or 1st July of the year in question. Estimates of the midyear population (mean population) are most suitable as standards of comparability with regard to time reference.

It is practically a sufficient approximation for the computation of annual rates (birth, deaths & marriage rates etc.) When the population is subject to seasonal fluctuations, the midyear population differs from the mean population, otherwise for all other cases, the difference is practically negligible.

Therefore, the AM of the population for the two 1st January dates that bound the year in equations is often used as an approximation of the mean population or the midyear population.

The estimation of mean population:



The mean population can be estimated by the methods of arithmetic linear interpolation (estimation of a value lying between two values)

When the interval between the given values is small, we can assume that a linear relationship holds within these intervals. Suppose,

P_{t+n} = the size of the population at time $t+n$

P_t = the size of the population at time t

n = length of time t

h = the amount of growth in number per unit of time.

Then we define,

$$h = \frac{P_{t+n} - P_t}{n} \dots\dots\dots (1)$$

Let us consider i be the length of time between the earlier known date & the estimated date i ($t+i$), then we can show that any value, P_{t+i} within intervals t & ($t+n$) can be obtained using arithmetic interpolation as follows:

$$P_{t+i} = P_t + \frac{i}{n} (P_{t+n} - P_t) = P_t + i \times \frac{P_{t+n} - P_t}{n} = P_t + i \times h \text{ [using (1)]}$$

Where, P_{t+i} is any value between points A & C (in figure). Thus concept can be expressed in the figure (given). Again,

$$\begin{aligned} P_{t+i} &= P_t + \frac{i}{n} (P_{t+n} - P_t) \\ \Rightarrow P_{t+n} - P_t &= \frac{i}{n} (P_{t+n} - P_t) \\ \Rightarrow \frac{i}{n} &= \frac{P_{t+i} - P_t}{P_{t+n} - P_t} \dots\dots\dots(3) \end{aligned}$$

$$P_{t+i} = P_t + \frac{i}{n} (P_{t+n} - P_t)$$

$$P_{t+i} = P_t + \frac{i}{n} (P_{t+n} - P_t) \dots\dots\dots(4)$$

In general the formula (4) is a general formula for locating any value between two dates or two points in time. In figure we have two dates – 1st January 1999 & 1st January 2000. Here, P_{t+i} = Mean population = \bar{P} (say).

If we choose month as a unit of time, the length of time between two points A & C is n, 12 months & length of time between A & B is which is 6 month. Then,

$$\begin{aligned}
 \bar{P} &= P_{t+i} = P_{t+6} = P_t + \frac{6}{12} (P_{t+n} - P_t) \\
 &= P_t + \frac{1}{2} (P_{t+n} - P_t) \\
 &= P_t + \frac{1}{2} P_{t+n} - \frac{1}{2} P_t \\
 &= \frac{1}{2} P_t + \frac{1}{2} P_{t+n} \\
 &= \frac{1}{2} (P_t + P_{t+n}) \\
 &= \frac{1}{2} (P_t + P_{t+12}) = \text{mean population}
 \end{aligned}$$

Problem:

Suppose in a baseline enumeration the total population at the beginning t, the survey 1st January 1999 was 10572. After 12 months 1st January 2000 the total population counted 13124. Find the midyear population on June 1999.

Solution:

$$\begin{aligned}
 \bar{P} &= \frac{1}{2} P_t + \frac{1}{2} P_{t+n} \\
 &= \frac{1}{2} (13124 + 10572) \\
 &= 11848
 \end{aligned}$$

CHAPTER FOUR: Population CHANGE

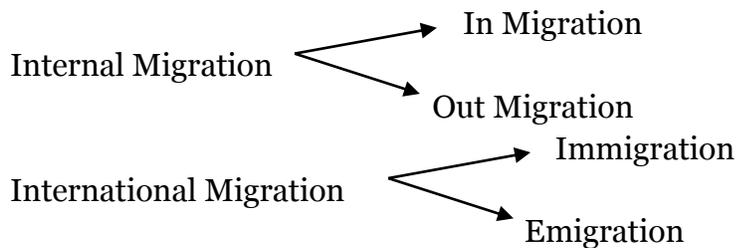
Concepts of Population Change:

Population change is the result of three major process:

1. Natural Change (Birth & Death)
2. Inter-regional Migration
3. International Migration

These changes be they an increase or decrease in the size of the population are usually termed as 'growth'.

It is important to note that natural increase (excess of birth over deaths) is the ultimate source of growth in the world as a whole but within a country migration is likely to be considerable force.



Measurement of Population Growth:

1. Rates of Natural Increase:

The rate of natural increase of a population over a specific time period is usually measured as the difference between the crude birth rate & crude death rate. That is,

$$NI \text{ (Natural Increase)} = CBR - CDR$$

If CBR = 50 per 1000; CDR = 20 per 1000

$$NI = (50-20) \text{ per } 1000 = 30 \text{ per thousand.}$$

2. Rates of Population Growth:

This might be defined as the rate of natural increase adjusted for net migration. It can be measured by a number of methods based on various assumptions about the pattern of growth. The most important thing to note is that, the rate of population growth consists of the elements of births, deaths & migration.

Let us consider,

P_0 = Initial Population

P_t = Population at the end i.e. after t years.

t = no. of years between P_0 & P_t

r = growth rate.

a. Arithmetic Growth Rate:

Ex: 5000, 5100, 5200, 5300, 5400, 5500....

That is $P_t = P_0 + P_0rt$

Sponsored by Stat Word: <http://stat.bishwo.com/>

$$t = 1, 2, 3, \dots$$

$$P_1 = P_0 + P_0 r$$

$$P_2 = P_0 + 2P_0 r$$

$$P_3 = P_0 + 3P_0 r$$

.....

.....

$$P_t = P_0 + P_0 r t$$

$$\Rightarrow P_t = P_0 (1 + r t)$$

$$\Rightarrow \frac{P_t}{P_0} = 1 + r t$$

$$\Rightarrow r t + 1 = \frac{P_t}{P_0}$$

$$\Rightarrow r t = \frac{P_t}{P_0} - 1$$

$$\Rightarrow r = \frac{1}{t} \left(\frac{P_t}{P_0} - 1 \right)$$

$$\Rightarrow r = \frac{P_t - P_0}{P_0 \cdot t}$$

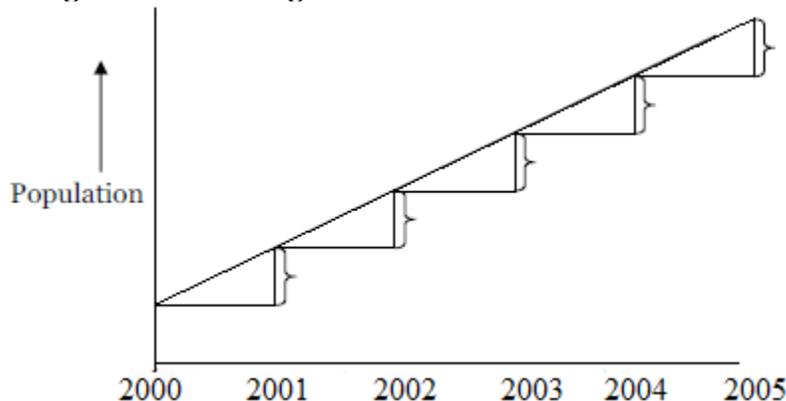
Which is the growth rate under arithmetic assumption.

Arithmetic growth:

#a population growing arithmetically would increase by a constant no. of people in each period. Arithmetic growth produces a linear trend in population growth following a straight line rather than a curve.

- Real examples of Arithmetic population growth are uncommon.
- Despite its limitation for describing actual population changes, it is the basis for widely used measure in demography.

In figure arithmetic growth:



Geometric growth method:

Geometric population growth is analogous to the compound interest.

Compound Interest:

The method is also used in the analysis of interest earned on capital invested at compound interest. The theory of compound interest assumes that the interest earned is automatically reinvested.

The word compound refers to the process of interest being reinvested to earn additional interest. If one considers the investment of I which accumulated to (I+i) at the end of the year where, i is the annual rate of interest. Then the balance of (I+i) can be considered as the principle at the beginning of the second year. The balance at the end of the second year is-

(I+i) + i(I+i) and so on [as $P_t = P_0 + rt$, here, $P_0 = I+i$, $t=1$, $r = i$, $rt = i$]

If I=1 (say),

Then, $(1+i) + i(1+i)$

$$= 1 + i + i + i^2$$

$$= 1 + 2i + i^2$$

$$= (1+i)^2$$

In ending of 2nd year, this will be

$$(1+i)^3 \dots\dots$$

$$(1+i)^4 \dots\dots\dots$$

If the process is repeated for t year, then the amount will be,

$$(1+i)^t, t > 0$$

In the population growth problem, we will set r equal to i.

Geometric Growth Method:

The geometric growth rate in demography is calculated using the compound interest formula. Actual population rarely increases at constant rates.

The main problem with describing population growth as a geometric progression we can write (assuming)

Year	Population
0	P_0
1	P_1
2	P_2
3	P_3
4	P_4
.	.
.	.
.	.

$$P_1 = P_0 + rP_0 = P_0 (1+r) \text{ [taking } t = 1]$$

$$P_2 = P_1(1+r) = P_0(1+r)^2$$

$$P_3 = P_2 + rP_2 = P_2(1+r) = P_0(1+r)^3$$

$$P_4 = P_0(1+r)^4$$

$$P_t = P_0(1+r)^t$$

$$\left(\frac{P_t}{P_0}\right)^{1/t} - 1 = r$$

$$r = \left(\frac{P_t}{P_0}\right)^{1/t} - 1$$

Exponential growth:

Sponsored by Stat Word: <http://stat.bishwo.com/>

Exponential growth method refers to the situation when the growth compounds continuously at every instant of time.

That is it is instantaneous growth.

Exponential Growth (Continuous Model):

In general, the term exponential function $y = f(x)$ can be written as $y = ab^{cx}$; where y & x are variables and a, b are the parameters.

In particular growth situations, it is generally convenient to consider only the case when $b > 1$.

A specific case of an exponential function of considerable use in demographic analysis or economics is where the coefficient b is replaced by constant $e = 2.71828\dots$, the base of natural logarithm & is given by

$$y = ae^{cx}.$$

Such a function is called the natural exponential function & is of considerable use in describing continuous growth process.

A particular form of the equation is the following

$$y = ae^{rn}$$

where, y = level of some variable (as population) after n periods of continuous growth at a rate r , beginning from the level 'a' in the first period.

This type of function could be used to describe the compound interest problem. If the growth rate calculated at the nominal rate of r percent per year were added continuously to the initial population, instead of only once per year. Under this minute by minute compounding arrangement our equation will take the form-

$$P_{t+n} = P_t e^{rn}$$

Where, y is replaced by P_{t+n} & a is by P_t .

This type of growth model (exponential) describes better the nature of the population.

This model can also be used to estimate r .

Initial Population = P_0

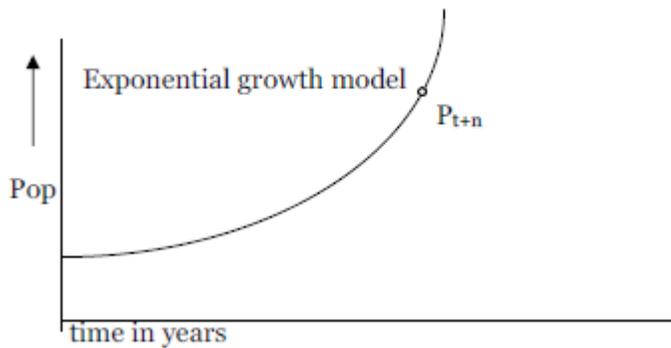
$$P_t = P_0 e^{rt}$$

$$\Rightarrow e^{rt} = \frac{P_t}{P_0}$$

$$\Rightarrow rt = \ln \left(\frac{P_t}{P_0} \right)$$

$$\Rightarrow r = \frac{1}{t} \ln \frac{P_t}{P_0}$$

Which is the growth rate under the assumption of exponential growth.



Problem:

How long will it take to a countries population, say 673 million to be doubled if growing at a constant rate of 2.4 percent per 100 starting from base years 2000, considering arithmetic, geometric & exponential growth?

Interpolation:

By interpolation we mean the process of calculating mathematically the size of a population at a point of time between two dates at which its sizes are known, the average rate of increase between two dates is known & the rate is applied to estimate the population at some point between two dates.

Extrapolation:

By extrapolation we mean estimate of a value outside the range of values given. The two formulae - the geometric growth method or exponential growth method may be used to estimate the size of population at points of time in the future on the assumption that the growth rate is constant & will not change.

Importance of r:

1. We can estimate the value at a later time or calculate how large the population might be at a given time.
2. If we know the rate, then we know how long it would have to reach a given size, i.e. we can estimate the length of time required for the population to reach a given multiple of its present size.

Population Growth Rates, Population size & Per capita Income:

Real income per head (per capita income) is the total income of a country, divided by the number of people. When analyzing the effects of alternative trends of population or the level of economic development, it is necessary to take into their effect both on total

income & ob the number of people among whom this income must be divided. Changes in either will effect per capita income.

END of CHAPTER FOUR

CHAPTER FIVE: ERRORS IN AGE DATA

Measurement of age, causes of errors & their detection etc:

Introduction:

Age is an important demographic variable. It is the primary basis of demographic classification in vital statistics, census, & survey work.

Much of the demographic information in use is expressed in terms of the age at which something happens. As for example, when an individual leaves school, enters employment or dies.

Age data provided an indispensable basis for studies of a nation's manpower potential & its requirements for school, housing, foods, goods & services etc.

With the appropriate vital statistics

-the fundamental measures of population growth,

-Age-specific fertility rates

-Age-specific mortality rates

-reproduction rate

-life table functions can be estimated.

Therefore, the age structure of a population is considered the basic to the study of population problems.

Population Composition by age & sex structure:

A normal age structure derived from a complete & relatively accurate census enumeration follows a typical pattern.

It begins with a large number (or proportion) in the first age group (0-4) of each sex.

This gradually diminishes in subsequent age groups until the final group at around 100 years when the number becomes negligible or nil.

Population Pyramid:

A population Pyramid shows the total picture of a population by age & sex. It is a useful way to illustrate the age structure of population.

Forms of Pyramid:

Pyramids are of different shapes. The shapes are determined by 3 factors.

1. Fertility
2. Mortality
3. Migration

Therefore each rectangle representing an age or a group of ages has length that depends on

- The total population at birth of the corresponding cohort or group of cohorts
- The amount of reduction by mortality &
- The amount of migration

The Construction of a Pyramid:

In a pyramid the absolute number in each group can be represented by a rectangle. Ages are located on the vertical axis & the total of each of the different ages (or age groups) are shown on the horizontal axis.

The rectangles are placed on top of each other, the youngest age group at the bottom & the oldest on top.

The rectangles for males are shown on the left side of the diagram & those for females on the right side. The open-ended age group can be represented only approximately.

As for example, If the open-ended age group is 75+ (75 years and above) we can treat this as 75-80 age group & draw the rectangles in this age group.

Pyramid from following date:

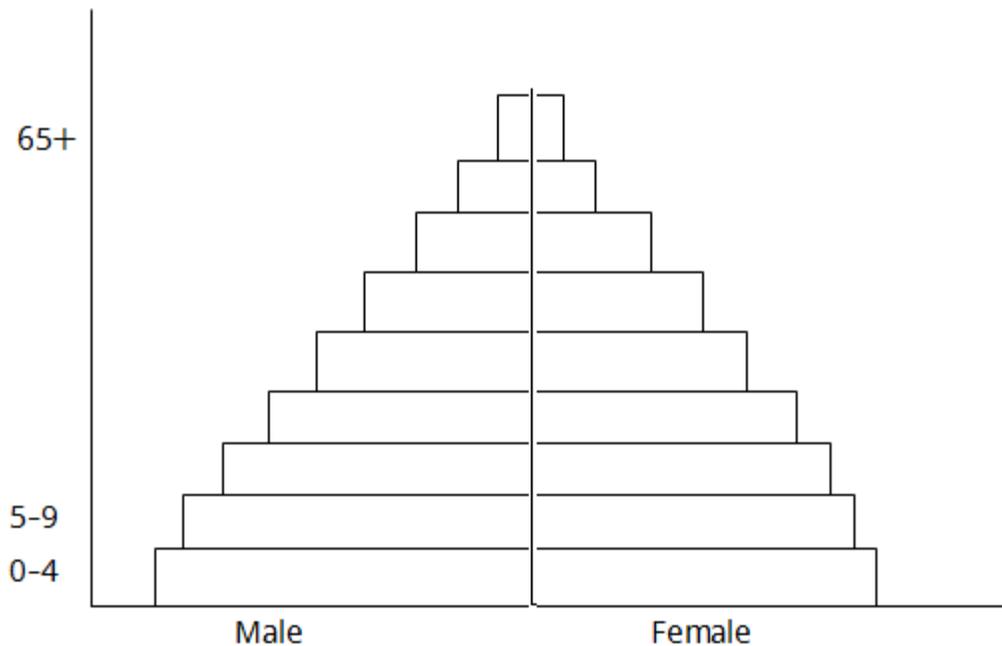
Age group	Male	Female
0-4	904,392	925,080
5-9	731,284	729,787
10-14	568,079	517,416
.		
.		
.		
65+		
	4769860	4686603

The population for males in 0-4 age group is

$$\frac{904392}{4769890+4686603} \times 100 = 9.6$$

Similar calculation can be carried out for the other age groups (both for males & females).

The Pyramid:



Methods of detecting errors in age data:

There are two broad approaches to the problem of identifying errors in age data. The first approach, useful before the data are published involves the case by case checking techniques employing data from interviews & administrative records.

The second approach involves the use of the techniques of demographic analysis.

Age Heaping

#Age Shifting

Age Heaping/Age Preference/Digital Performance:

The tendency of entering on the digits (0, 5) is called age heaping or digital preference or age preference.

Many people have the tendency to repeat age ending with zero & 5. If someone's exact age 24 he/she tells 25.

The tendency is called age heaping.

Because of age heaping, age data suffers from errors.

#Age Shifting:

Age shifting is the tendency in persons in a particular age range to report them either younger or older than the true age.

As for example, older people tend to exaggerate their ages, young men may overstate their ages, unmarried women may understate their ages.

Age data suffers from error because of age shifting.

Causes of errors in age Data:

- i) Ignorance of correct age
- ii) Carelessness in reporting & recording
- iii) A general tendency to state age in figures ending in certain preferred digits (Age Heaping)
- iv) A tendency to exaggerate length of life at advanced ages (Age Shifting)
- v) Misstatements arising from motives of an economic, social, political or purely personal character.

There are THREE measures of detecting digital preference (or errors in the age data):

- i) Age ratios, sex ratios & age/sex accuracy index
- ii) WHIPPLE's Index
- iii) MYER'S Index

Age ratios, sex ratios & age/sex accuracy index:

Age ratios and sex ratios can be used either separately or jointly in evaluating the quality of the census returns of surveys by age groups.

Age ratio is defined as the ratio in the given age group to one half the populations in two adjacent groups.

Mathematically, let ${}_5P_x$ be the age group from age x to $x+5$, ${}_5P_{x-5}$ & ${}_5P_{x+5}$ by the preceding & following age group respectively, then,

$$\text{Age Ratio} = \frac{{}_5P_x}{\frac{1}{2}\{{}_5P_{x-5} + {}_5P_{x+5}\}} \times 100$$

The computed age ratios are then compared with the expected value, which is usually 100. The discrepancy at each age group is a measure of net age misreporting.

An overall measure of the accuracy of an age distribution called age accuracy index is derived by taking the average deviation (regardless of sign) from 100.0 of the age ratios & summing over all age groups.

The lower this index, the more adequate the census data is on age. The age ratios are usually calculated for males & females separately & can be calculated for each age group (except the youngest and the oldest) provided the intervals are equal.

An age ratio under 100 implies either that members of the group were selectively under enumerated or that error in age reporting resulted in misclassifying persons who belonged to that age group.

A ratio more than 100 suggests the opposite (over enumerated) of one or other.

- i) If Age Ratio (AR) < 100, then it implies that in particular age group the popⁿ is under counted.
- ii) If AR = 100, there is no selective error.
- iii) If AR > 100, then it implies that particular age group is over calculated.

Application of Age ratio:

AR is a technique which can be applied to evaluate the quality of census & survey data on age.

Limitation of AR:

AR cannot be calculated for the first age group.

[An example]

Age-specific sex ratio (number of males per 100 females in age group) can also be used in evaluating census data.

$$\text{Sex Ratio} = \frac{\text{Male}}{\text{female}} \times$$

Joint Score = $3 \times$ (mean of sex ratio score) + (mean of male & female age ratio scores)
 Joint scores are proposed by United Nations which is Age/Sex accuracy index.

Whipple’s Index (WI):

Quality of Data	Whipple’s Index
Highly Accurate	Less than 105
Fairly accurate	105-109.9
Approximate	110-124.9
Rough	125-174.9
Very Rough	175+

Whipple’s index is applicable where age is reported in single years. The index is obtained by summing the ages returns between 23 & 62 years inclusive and finding what percentage is borne by the sum of the returns of years ending with 5 and 0 to one fifth ($\frac{1}{5}$) of the total sum.

By the above definition WI can only measure the extent of age heaping at the digit 5 & 0. It is not an efficient method since heaping occurs on digits other than 5 & 0.

The result varies between a minimum of 100 representing no concentration at all and maximum 500, if no returns are recorded with any digits other than 0 & 5.

Between these extremes the mentioned scales has been used.

$$WI = \frac{P_{25}+P_{30}+P_{35}+P_{40}+P_{45}+P_{50}+P_{55}+P_{60}}{\frac{1}{5}(P_{23}+P_{24}+P_{25}+\dots+P_{62})} \times 100$$

$$\left[\begin{array}{l} \square\square\square min = 100 \\ \max = 500 \end{array} \right]$$

MYER’S INDEX:

Myers’s index is an ingenious device which reflects preferences (or dislikes) for each of the ten digits from 0 to 9.

Like Whipple’s index, it is applicable where age is given in single years. The method derives a blended population which is essentially a weighted sum of the number of persons reporting ages ending in each of the ten terminal digits 0,1,2,3,4,5,6,7,8,9.

Assumption:

The underlying assumption of this method is that there are no systematic irregularities in the reporting of the age, the blended sum of each terminal digit should be approximately 10 (ten) percent of the total blended population.

- i) If the sum at any given digit exceeds 10 percent of the total blended population, it indicates over selection of ages ending in that digit (digit preference)
- ii) On the other hand, a negative deviation (or sum is less than 10 percent of the total blended population) indicates under selection of ages ending in that digit (digit avoidance). An overall measure of extent to which there is digit preference &/or avoidance in a census age distribution is the index of preference which is obtained as the absolute sum of derivatives for each of the ten terminal digits.

MYER's Index can vary from 0 to 180 (representing respectively, the situations where ages are reported accurately and where all ages were reported with the same terminal digits).

Compiled by: Abdullah Adil Mahmud, SBI, DU
Get more documents on <http://stat.bishwo.com/>