U.P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N

Demography

Course Design Committee

Dr. Ashutosh Gupta **Chairman**
Director, School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

Prof. Anup Chaturvedi **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. S. Lalitha **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. Himanshu Pandey **Member**
Department of Statistics
D. D. U. Gorakhpur University, Gorakhpur.

Prof. Shruti **Member-Secretary**
Professor, School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Course Preparation Committee

Dr. Anupma Singh **Writer**
Department of Statistics
University of Lucknow, Lucknow

Prof. Shruti **Editor**
School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Prof. Shruti **Course Coordinator**
School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N **DEMOGRAPHY**

First Edition: ***July 2024***
ISBN: 978-93-48987-49-5

Printed and Published by Col. Vinay Kumar, Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2024.
Printed by : Chandrakala Universal Pvt.Ltd. 42/7 JLN Road, Prayagraj, 211002

Blocks & Units Introduction

The present SLM on ***Demography*** consists of Ten units with Four blocks.

The ***Block 1 – Migration*** is the first block of said SLM and is divided into two detailed units.

Unit- 1 Fundamentals of Migration, is the first unit of the Self Learning Material which describes the fundamentals of migration, conceptual definitions, types of migration, sources of migration data, factors affecting migration and direct estimation of migration with the available data.

Unit-2 Methods of Estimation of Migration, is focused on the indirect estimation of migration. Various indirect methods have been discussed with examples including life time survival ratio method, census survival method, life time survival ratio method and vital statistics-based method. This unit also covers the estimation of international migration using the available data.

Block 2 - Stable Population Theory is the second block of this Self learning material.

The Unit 3: Introduction to Stable Population Theory is the first unit of the Self Learning Material which explains the basic concepts and terminologies needed to understand stable Population theory. The unit includes introduction to different type of population structure: Stable, Stationary, Quasi-stable and Non- Stable Populations and their characteristics. It also explains Characteristics and vital rates of Stable Stationary and Quasi- stable Population. Limitations of Stable Population theory.

Unit 4: **Theories and Relationships related to** ***Stable Population Theory*** is the second unit of this Self Learning Material explains the Definitions of intrinsic rate of natural increase. It also explains, basic and vital models developed to explain population growth and projection. It emphasizes on, inter-relationship among intrinsic birth rate and intrinsic death rate, derivation of stable population equation and relationships between birth rate and death rate under condition of stability.

Unit 5: Growth Rates is the third unit of this Self Learning Material explains Computation of intrinsic rate of natural increase under stability condition. It includes derivation of construction of a stable age distribution for a given fertility and mortality schedules, relationship between Net Reproduction Rate, Intrinsic Growth Rate and mean length of generation and the concept of mean interval between generations.

Block 3- ***Fertility and Fertility Models*** is the third block of said SLM and it is divided into three units.

Unit 6: ***Fertility and its Measures****,* is the first unit of the Self Learning Material which explains the concept of fertility in detail. It also covers different period measures of fertility and reproduction such as, Crude birth rate (CBR), General fertility rate (GFR) Age- specific fertility rate (ASFR), Total fertility rate (TFR), Gross reproduction rate (GRR), etc.

Unit 7: ***Cohort Measures and Indirect Estimation of Fertility,*** is the second unit of the Self Learning Material which focuses on different cohort measures of fertility, Use of birth order statistics, child women ratio, own-children method, children ever born (CEB) data and with data on current fertility, Brass P/F ratio for adjusting fertility rates.

Unit 8: ***Fertility Models,*** *is the third and the last unit* of the Self Learning Material which explains indirect estimation of fertility. This unit talks about modelling of fertility and studying some important probability models on time of first birth/conception and number of births/conception n specified time, birth interval models, study of fertility through birth interval analysis.

Block 4 - ***Mortality and Life Table,*** is the fourth block of the said SLM which is divided into two units.

Unit 9- ***Mortality and its Measures,*** is the first unit of the Self Learning Material, which explains the basic concepts and definitions of mortality. It also explains the methods of standardization of death rates in the situations where the mortality condition of varying populations has to be compared.

Unit 10- ***Life Table,*** is the second unit of the Self Learning Material, which focuses on the concept and development of Life Table. It explains the construction of complete and abridged life tables and also the approximations used for different life table function. This unit explains the use of life table in reality.

At the end of every block/unit the summary, self-assessment questions and further readings are given.

U.P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N Demography

Block: 1 Migration

Unit – 1 : Fundamentals of Migration

Unit – 2 : Methods of Estimation of Migration

Course Design Committee

Dr. Ashutosh Gupta **Chairman**
Director, School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

Prof. Anup Chaturvedi **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. S. Lalitha **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. Himanshu Pandey **Member**
Department of Statistics
D. D. U. Gorakhpur University, Gorakhpur.

Prof. Shruti **Member-Secretary**
Professor, School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Course Preparation Committee

Dr. Anupma Singh **Writer**
Department of Statistics
University of Lucknow, Lucknow

Prof. Shruti **Editor**
School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Prof. Shruti **Course Coordinator**
School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N DEMOGRAPHY

First Edition: ***July 2024***
ISBN: 978-93-48987-49-5

Printed and Published by Col. Vinay Kumar, Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2024.
Printed by : Chandrakala Universal Pvt.Ltd. 42/7 JLN Road, Prayagraj, 211002

Block & Units Introduction

The present SLM on ***Demography*** consists of Ten units with Four blocks. This is the first block.

The ***Block 1 – Migration*** is the first block of said SLM and is divided into two detailed units.

Unit- 1 - Fundamentals of Migration, is the first unit of the Self Learning Material which describes the fundamentals of migration, conceptual definitions, types of migration, sources of migration data, factors affecting migration and direct estimation of migration with the available data.

Unit-2 - Methods of Estimation of Migration, is focused on the indirect estimation of migration. Various indirect methods have been discussed with examples including life time survival ratio method, census survival method, life time survival ratio method and vital statistics-based method. This unit also covers the estimation of international migration using the available data.

At the end of every block/unit the summary, self-assessment questions and further readings are given.

UNIT – 1 FUNDAMENTALS OF MIGRATION

Structure

1.1 Introduction

Migration is a fundamental aspect of human history, shaping societies, economies, and cultures across the globe. From ancient nomadic tribes seeking fertile lands to modern-day individuals pursuing better opportunities, the movement of peoples has played a critical role in the development of civilizations. Migration often occurred in response to environmental pressures, such as climate change or natural disasters, which compelled communities to seek more hospitable lands.

In ancient times, migration was pivotal to the formation of early states and empires. For instance, the movements of Indo-European peoples into Europe and South Asia brought about profound changes in language, culture, and societal structures. Similarly, The Arab expansion during the early Islamic period resulted in a vast exchange of knowledge and culture across Indian subcontinent. In early days, the spread of the British Empire across America, Asia, and parts of Africa facilitated the exchange of goods, modern ideas, and technologies, leaving a long-term impact on the regions involved.

The impact of migration in post-independent India has been a significant and transformative force in the nation's demographic and socio-economic landscape. In the immediate aftermath of the 1947 Partition, India experienced one of the largest and most traumatic migrations in history, with an estimated 15 million people displaced across the new borders between India and Pakistan. This massive population exchange had profound implications for regions such as Punjab, Delhi, and West Bengal, where the influx of refugees necessitated urgent resettlement efforts and led to lasting demographic shifts. The rapid and unplanned nature of this migration created significant challenges in terms of resource allocation, social integration, and economic stability. In particular, the states receiving large numbers of refugees had to navigate complex issues of land redistribution, employment, and social services, which, in many cases, laid the groundwork for ongoing regional disparities and tensions.

Beyond the immediate effects of Partition, migration continued to shape India's development trajectory throughout the post-independence period. Internal migration, particularly rural-to-urban migration, became a key driver of urbanization and economic growth. The urban population in India increased from 17.29% in 1951 to 31.16% by 2011, with cities like Mumbai, Delhi, and Kolkata emerging as major economic hubs. This migration brought about significant economic opportunities, contributing to the diversification of labour markets and the expansion of industries. However, it also posed substantial challenges, including overburdened urban infrastructure, housing shortages, and social inequality. The influx of migrants into urban areas often intensified competition for jobs and resources, leading to socio-economic stratification and, in some cases, social unrest. The complex dynamics of migration in post-independent India highlight its dual role as a catalyst for economic transformation and a source of enduring socio-political challenges, necessitating responsive policy interventions to manage its impacts effectively.

Recently, the effects of migration on socio-economic, demographic and cultural factors have attracted increasing attention from administrators, planners, social scientists and researchers seeking a thorough understanding of the process of population movement.

This unit examines the concept of migration, exploring its trends, and its impacts on population change. As we begin this exploration, we will investigate the different forms of migration, including internal and international migrations, explore the reasons behind migration, and review methods for estimating migration.

1.2 Objectives

After completing this unit, you will be able to:

- Understand the fundamental concept and definitions of migration.
- Explore theoretical frameworks and key terminology related to migration.
- Identify and describe different forms of migration, including internal and international migration.
- Analyse the factors affecting the migration.
- Review and assess methods for estimating internal and international migration, including their strengths and limitations.

1.3 Definitions

There are some important definitions:

1.3.1 Migration

The term ‘migration’ is derived from the Latin word ‘migrare’, which means ‘to move or shift’. The term ‘migration’ is defined in the multilingual demographic dictionary of the UN as “spatial mobility between one geographical unit and another, involving a permanent change of residence”. Thus, migration is the movement of people across a specified boundary for the purpose of establishing a new residence. Migration may be permanent or temporary with the intention of returning to the place of origin in future.

1.4 Types of Migration

Migration can be categorized into several distinct types based on various criteria, including the duration of the move, the nature of the migration (voluntary or forced), and the geographical scope of the movement. Understanding these types is essential for comprehending the complexities and implications of migration in contemporary society.

1.4.1 Internal Migration

This refers to the movement of people within a country. Internal migration can occur for various reasons, such as seeking better employment opportunities, education, or living conditions. It can further be divided into:

- ***Rural to Urban Migration:*** Movement from rural areas to cities, often in search of better job prospects and living standards.

- ***Urban to Rural Migration:*** Movement from urban areas back to rural regions, often driven by a desire for a quieter lifestyle or to escape urban challenges such as pollution and overcrowding.
- ***Urban to Urban Migration:*** Relocation from one city to another, typically for job-related reasons or lifestyle changes.
- ***Rural to Rural Migration:*** Movement within rural areas, often related to agriculture or family ties.

Internal migration data is crucial for effective economic planning and resource allocation, guiding infrastructure and public service investments. It helps urban planners manage city growth and address rural depopulation. The data informs public policy, ensuring that education, healthcare, and social services are adequately distributed. Additionally, it provides insights into population dynamics, essential for long-term demographic planning.

1.4.2 International Migration

This involves the movement of people across national borders. It can be categorized into:

- ***Immigration:*** The process by which people move into a country or region from another place with the intent to settle there. Immigrants typically seek to establish permanent residency or citizenship in the new country.

- ***Emigration:*** The act of leaving one's own country or region to settle in another. People who emigrate are often referred to as emigrants, and they are leaving their home country in search of better opportunities, safety, or other reasons.

- ***Refugees:*** Individuals who are forced to flee their home country due to persecution, war, violence, or a well-founded fear of being persecuted for reasons such as race, religion, nationality, membership in a particular social group, or political opinion. Refugees seek safety and asylum in another country, and they are protected under international law.

International migration data is vital for shaping immigration policies and managing border security effectively. It helps governments understand labour market needs, allowing for the strategic recruitment of foreign workers. The data also supports social integration programs by identifying the cultural and demographic impacts of migration. Additionally, it aids in international cooperation and development, addressing global challenges like refugee crises and economic disparities.

1.5 Some Concepts Related to Migration

Following are some basic concepts related to migration

1.5.1 Migrant

Migrant is a person involved in the process of migration. i.e. A household member who previously lived at a different location than their current residence is considered a migrant.

1.5.2 Migration Rate

The migration rate for a specific category of individuals (such as rural or urban, male or female) represents the percentage of migrants within that category. The migration rate can be obtained as

$$m = \frac{M}{P} \times k$$

Where m is the rate of migration for the specified migration interval,

M is the number of migrations or the number of persons migrating during the interval,

P is the population exposed to the likelihood of migration during the interval, and

k is a constant, usually 100 or 1,000

1.5.3 Life-time Migrant and life-time Migration

If a person's place of residence differs from their place of birth at the time of a census or survey, they are referred to as a lifetime migrant, and this phenomenon is known as lifetime migration.

1.5.4 Migration and Out-Migration

The process of entering one administrative subdivision of a country from another subdivision is known as in- migration and any migration from a specified area to outside is known as out- migration. In other words, every move is called out-migration with respect to the place of origin and in-migration with respect to the place of destination.

1.5.5 Place of Origin and Place of Destination

A place from where an individual moves out/leaves his usual place of residence is called the place of origin where as a place in which an individual terminates his move is called the place of destination. Thus, an individual is called as out-migrant with respect to the place of origin whereas the same individual is called in-migrant with respect to the destination.

1.5.6 Gross Migration and Net Migration

The total movements in a specified area during a given interval time is gross migration. The gross migration is defined as

$$\text{Gross Migration} = \text{In-Migration} + \text{Out-Migration}$$

The net effect of in-migration and out-migration on an area's population in a given time period, expressed as increase or decrease is net migration. It is the balance of movements made in opposite directions. i.e.

$$\text{Net Migration} = \text{In-Migration} - \text{Out-Migration}$$

1.5.7 Primary Migration and Secondary Migration

Primary migration refers to the initial movement of people from one place to another. This could be a move from a rural area to an urban area, from one country to another, or any other significant change in location. Whereas, secondary migration refers to subsequent moves made after the initial migration. It involves relocating from one place to another after the first move, often within the same country or region.

1.5.8 Seasonal Migration and Circular Migration

Seasonal migration refers to the temporary movement of people in response to seasonal changes, often seen in agricultural labour where workers migrate to areas with seasonal job opportunities. Circular migration refers to the repeated movement of people between their place of origin and destination, typically for work, on temporary basis. These types of migration is common in regions where individuals move from rural to urban areas (or to different rural areas) for specific periods, often following the agricultural calendar, construction cycles, or other seasonal labour demands.

1.5.9 Return Migration

Return migration describes the process in which migrants return to their original country or place of origin after having in a different region or country. It can occur voluntarily or due to external pressures and may be either a temporary or permanent move.

1.5.10 Voluntary and Forced Migration

Voluntary migration occurs when individuals choose to move, often in search of better economic opportunities, education, or quality of life. Forced migration includes situations where individuals are compelled to leave their homes due to external pressures, such as war, natural disasters, or government policies.

1.5.11 Migration Stream

The total number of movements having common place of origin and destination during a given migration interval is termed as migration stream.

1.6 Sources of Migration Data

The primary source of migration data can be categorised as:

1.6.1 Census and Survey Data

National censuses, household and labour force surveys are traditional methods for collecting migration data. These sources provide insights into the demographic characteristics of migrants, including age, gender, and country of origin. However, they may not capture transient populations or those with irregular status effectively. Apart from census data, NSS data, Economic Survey of India and Periodic Labour Force Survey are primary sources of migration data in India.

1.6.2 United Nation Data

The UN Department of Economic and Social Affairs (UN DESA) is a key source for data on immigrant and emigrant stocks, relying on population censuses, demographic data, and modelled estimates when necessary. This data helps define migrants based on their country of birth and residence, capturing various categories like refugees and unauthorized immigrants.

1.6.3 Administrative Data

Government records, such as population register, immigration and asylum applications, offer additional insights into migration patterns. These records can be useful for tracking legal migration flows but may miss undocumented migrants.

1.6.4 Digital Data

Recent advancements have led to the use of big data and digital trace data, such as mobile phone records and social media activity, to analyse migration trends. These sources can provide real-time insights and capture hard-to-reach populations, including irregular migrants.

1.7 Factor affecting Migration

Migration is influenced by various push and pull factors that drive individuals to leave their home countries or regions and seek opportunities elsewhere. These factors can be broadly categorized into negative aspects that force people to leave (push factors) and positive attributes that attract them to a new location (pull factors).

1.7.1 Push Factors

Push factors are unfavourable conditions in the home area that drive individuals to migrate. Key push factors include

- ***Economic Hardship:*** Limited job opportunities, low wages, and high unemployment rates can compel individuals to seek better economic conditions elsewhere. Many people relocate their usual place of residence primarily because their parent or the primary earner of the household has moved.
- ***Conflict and Violence:*** War, civil unrest, and persecution based on ethnicity, religion, or political beliefs force people to flee their homes for safety.
- ***Natural Disasters:*** Events such as earthquakes, floods, droughts, and other environmental crises can displace populations and lead to migration.
- ***Poor Living Conditions:*** Issues such as inadequate healthcare, poor housing, pollution, and lack of access to education can motivate individuals to leave their current location in search of a better quality of life.
- ***Political Oppression:*** Authoritarian regimes, lack of political freedom, and human rights abuses can push individuals to seek refuge in more democratic and stable countries.
- ***Marriage:*** In many regions, a female has to migrate to her husband's place of residence after marriage.

1.7.2 Pull Factors

Pull factors are the attractive elements of a destination that attract migrants. Important pull factors include:

- ***Job Opportunities:*** The availability of better jobs and higher wages in the destination country is a significant motivator for migration.
- ***Education:*** Students often move across regions to attend prestigious institutions or specialized programs that are not available in their region.
- ***Better Living Standards:*** Access to improved healthcare, education, and overall quality of life can attract individuals to migrate.
- ***Political and Religious Freedom:*** Countries that offer greater political stability and respect for human rights can draw individuals fleeing oppression.
- ***Family and Social Networks:*** Existing family members or communities in the destination area can provide support and resources, making relocation more appealing.
- ***Cultural and Recreational Opportunities:*** Attractive climates, cultural amenities, and lifestyle choices can also serve as pull factors for migrants seeking a different way of life.

Understanding push and pull factors is essential for analysing migration patterns and addressing the challenges and opportunities that arise from human mobility.

1.8 Recent Migration Trend in India

According to the World Migration Report 2024, India is the largest origin country for international migrants, with approximately 18 million Indians living abroad. The primary destinations include the UAE, US, and Saudi Arabia. Notably, male emigrants constitute about 65% of this total, primarily migrating for employment opportunities, while many women remain in India, often migrating for marriage.

According to Migration in India report on PLFS 2020-21, nearly 28.9% of India's population were migrants, with a significant disparity between rural (26.5%) and urban (34.9%)

migration rates. Female migration is notably higher in rural areas, primarily for marriage, while males migrate for work.

Table 1: Migration Rate in India (in per cent) from PLFS 2020-21

Gender	Rural	Urban	Rural + Urban
Male	5.9	22.5	10.7
Female	48.0	47.8	47.9
Male + Female	26.5	34.9	28.9

The data on migration flows between rural and urban areas for internal migration in India is also included in the PLFS 2020-21. The following table illustrates the distribution of internal migration. It shows that rural-to-rural migration is the most common among all migration streams, while urban-to-rural migration has the lowest volume.

Table 2: Percentage Distribution of Internal Migrants over the four Types of Rural-Urban Migration from PLFS 2020-21

Category of migrants	Migration Stream				
	Rural to rural	**Urban to rural**	**Rural to urban**	**Urban to urban**	**All**
Male	18.0	20.8	33.5	27.6	100.0
Female	63.3	7.8	15.6	13.2	100.0
All Person	55.0	10.2	18.9	15.9	100.0

Percentage distribution of Inter-state migrants by location of last usual place of residence in terms of same State, another State or other countries has been given in PLFS 2020-21. The distribution of the migrants in terms of same state, another state or other countries has been shown in following table.

Table 3: Percentage distribution of migrants by location of last usual place of residence in terms of same State, another State or other countries for each category of migrants from PLFS 2020-21.

Category of Migrants	Last usual place of residence in			
	Same State	Another State	Other Countries	All
Rural				
Male	62.5	33.7	3.9	100.0
Female	95.8	4.0	0.2	100.0
All Person	92.1	7.3	0.6	100.0
Urban				
Male	67.9	29.9	2.3	100.0
Female	84.7	14.9	0.4	100.0
All Person	79.0	19.8	1.0	100.0
Rural + Urban				
Male	65.6	31.4	2.9	100.0
Female	92.6	7.2	0.2	100.0
All Person	87.5	11.8	0.7	100.0

1.9 Estimation of Internal Migration

Internal migration refers to the movement of people within a country, which can significantly influence demographic dynamics and economic development. Understanding the patterns and volume of internal migration is crucial for policymakers, as it affects urbanization, labour markets, and social services. Estimation methods for internal migration often rely on census data, surveys, and vital statistics. Despite its importance, comprehensive data on internal migration remains limited, particularly in developing regions, necessitating innovative estimation approaches.

Measuring internal migration involves various methodologies that can be broadly categorized into direct and indirect methods. Each approach has its strengths and weaknesses, and the choice of method often depends on the specific objectives of the study and the available data.

1.9. 1 Direct Methods

Direct method involves direct questions on internal migration which are asked at the time of censuses and surveys. The direct method of measuring internal migration includes:

- Place of birth data;
- Place of last residence data; and
- Place of residence at a fixed prior date data.

1.9.1.1 Place of Birth Data

Place of birth gives information about Migrants and Non-migrants. In the census, direct questions are asked about the place of birth, such as, “Where was this person born?”. Depending upon the place of birth, and individual can be classified as a migrant (in or out) and non-migrant. Migrants, defined as persons who were enumerated in a place different from the place of birth. Non-migrants, defined as persons who were enumerated in the place of birth. Note that these terms refer to the life-time migration. For example, in India Census data show place of birth. It shows population classified by state of birth and state of enumeration.

Advantages of this method are:

- Easy to ask and understand.
- Individuals usually remember their place of birth.

However, some disadvantages of this method include:

- There exists some chance of error as answers to the Census questionnaires are generally given by the head of the household or a responsible member of the household. The respondent may not be aware of the exact place of birth of all the members of the household. A person living at a certain place for quite a long time may report it as his place of birth.
- There may be a tendency to report a better-known place as the place of birth instead of the actual place of birth.

- Sometimes some social practice may introduce artificial bias about the information on the place of birth. For example, In India, girls go to their parental homes for their first delivery due to which the child born becomes of life-time migrant though actually he or she is a non-migrant for every other purpose.

- This practice is based upon the assumption of a single movement directly from the place of birth to that of enumeration. Actually, some persons might have moved into the place of enumeration from some place other than the place of birth.

- The birth place data do not convey any idea about the time of movement.

1.9.1.2 Estimation of Life-time Migrants using Place of Birth Data

When the information is collected by place of birth and place of enumeration for each person in the census, the analysis of data involves cross-classifying individuals by their place of enumeration and place of birth. This can be done across various units: state-level, rural and urban areas within a state (rural-urban), locations within the same district (intra-district), other districts within the state (inter-district), outside the state but within the country (inter-state), and internationally (international). The table below illustrates the estimation of lifetime in-migrants and lifetime out-migrants.

Table 4: Tabulation in estimating migrants and non-migrants from place of birth data

Place of Birth	**Place of Enumeration**		
	A_1	A_2	A_3
A_1	a_{11}	a_{12}	a_{13}
A_2	a_{21}	a_{22}	a_{23}
A_3	a_{31}	a_{32}	a_{33}

In the table above, A_1, A_2, and A_3 are the locations where the census data has been collected. The elements a_{ij} indicates the total number of movements from area A_i to area A_j (i, j = 123). The diagonal elements a_{11}, a_{22}, and a_{33} represent total number of non-migrants, and off-diagonal elements represent the total number of migrants. Additionally, we can determine both in-migration and out-migration from the specified locations. For example, the total in-migration to place A_1 can be determined by summing a_{21} and a_{31}, and the total out-migration from place A_1 can be obtained by summing a_{12} and a_{13}.

Taking into account the number of migrants (i.e. in and out both) we can find out the percentage of life-time migrants in a country, i.e. the number of persons who are enumerated at different places other than their place of birth divided by the total population of that country.

1.9.1.3 Estimation of Intercensal Migration Using Place of Birth Data

From the data on the place of birth, it is possible to calculate the extent of migration during an intercensal period. It can be estimated by subtracting the survivals of the migrants counted in the first census from the migrants counted in the second census. Let

- I_t = number of lifetime in-migrants at time *t* in an area,
- I_{t+n} = number of lifetime in-migrants at time *(t + n)* in an area,
- O_t = number of lifetime out-migrants at time *t* from an area,
- O_{t+n} = number of lifetime out-migrants at time *(t + n)* from an area,
- NM = net migration

The indirect estimate of intercensal net migration can be given as:

$$NM = (I_{t+n} - O_{t+n}) - (S_1 I_t - S_0 O_t)$$

Where, S_1 and S_0 are intercensal survival ratios indicating what proportions of I_t and O_t will survive during the intercensal period.

Separating the in-migrants and out-migrants as

$$NM = M_1 - M_2$$

Where $M_1 = I_{t+n} - S_1 I_t$ and $M_2 = O_{t+n} - S_0 O_t$

From this expression, one can find out an estimate of the net balance of intercensal migration along with net-migration among persons born outside the area (M_1) and among the persons born inside the area (M_2).

1.9.1.4 Place of Last Residence Data

The place of birth data significantly underestimates the overall volume of migration. It neither include the details of a person migrated more than once, nor the information about the last residence of in-migrant if it is different from place of birth. Lacking the information about the last residence we lose the information about the migration takes place in between the observation period. It is important to ask questions about the place of last residence of the migrant and the time of in-migration at current place to determine the information about onward and return migration. This ensures a clear distinction between return migrants and non-migrants who have never left their places of birth.

Data categorized by place of last residence can be used similarly to place of birth data for estimating migration. By cross-classifying place of last residence with current or present residence (i.e., place of enumeration), one can estimate the origins of in-migrants to an area and the destinations of out-migrants from an area, thus obtaining an estimate of net migration. The data requirements and estimation methods for this approach are largely the same as those for place of birth data, with the primary difference being that place of last residence is used instead of place of birth.

An important advantage of the place of residence data is that in the place of residence data, direct movement between the places can be estimated while the place of birth data does not give an idea of the intervening movement; it gives idea only of the first residence and the last residence where the person has arrived.

However, there are some limitations of this method as like the place of birth, data on the place of last residence also suffer from absence of a definite time reference. The place of last residence does not indicate a definite period of in- migration. So, persons who have migrated 25 years ago or even before and persons who migrated recently, may be a few days ago will be grouped together and called as “migrants”.

1.9.1.5 Place of Residence at Fixed Prior Date

When using the concept of place of last residence to estimate migration, all migrants are grouped together regardless of their migration timing. For instance, individuals who migrated 20 years ago and those who moved just before the census are both classified as migrants. In this method a question is asked about the place of residence at a fixed prior date. The answers to these questions are important since migration interval in it is given by a comparison of residence at two definite points at time.

A person is considered a migrant if their residence at the fixed prior date (set by the researcher) differs from their place of enumeration. Conversely, those whose residence at the fixed prior date matches their place of enumeration are classified as non-migrants. This approach enables the identification of surviving migrants for a specified duration.

Limitations of this method are:

It does not account for migrants who were alive at the fixed prior date but moved elsewhere and died afterward, nor does it include migrants who returned to their original residence after moving during the migration interval. Additionally, for all those children who are born during the migration interval, i.e. a period from a fixed date to the census date, the question on the place of residence is not applicable, as they did not exist on that date. Similarly, this approach does not take into account those migrants who made moves during the migration interval, but subsequently returned to the place where they were residing at the fixed prior date.

Example: From the following data, estimate the volume of in-migration, out-migration, and net migration for the places A, B, and C as on 1 April 2001. Also calculate total number of migrants during the 1 April 1991 and 1 April 2001.

Table 5: Distribution of Persons Enumerated in Area A on 1 April 2001 according to their Residence on 1 April 1991 and their Place of Birth

Place of Birth	**Place of Residence on 1 April 1991**		
	A	**B**	**C**
A	98,256	4,984	1,098

B	567	2,134	135
C	443	231	877

Table 6: Distribution of Persons Enumerated in Area B on 1 April 2001 according to their Residence on 1 April 1991 and their Place of Birth

Place of Birth	**Place of Residence on 1 April 1991**		
	A	**B**	**C**
A	1,435	876	211
B	7,200	356,678	8,734
C	213	1,256	2,744

Table 7: Distribution of Persons Enumerated in Area C on 1 April 2001 according to their Residence on 1 April 1991 and their Place of Birth

Place of Birth	**Place of Residence on 1 April 1991**		
	A	**B**	**C**
A	453	198	288
B	211	3,277	722
C	2,434	15,960	441,680

To calculate lifetime migrants, we organize the provided data according to the place of birth and place of enumeration as of 1 April 2001. This is done by merging the columns from the provided tables into a single column each, disregarding the information about the place of residence on 1 April 1991, as shown below:

Table 8: Distribution of Persons According to Place of Birth and Place of Enumeration on 1 April 2001

Place of Birth	**Place of Enumeration on 1 April 2001**			
	A	**B**	**C**	**Total**
A	104,338	2,522	939	107,799

B	2,836	372,612	4,210	379,658
C	1,551	4,213	460,074	465,838
Total	108,725	379,347	465,223	953,295

From the tables, it is easy to obtain the details about lifetime migration in 2001. The diagonal values in the table 8 represent the non-migrants of respective areas, e.g. the number of non-migrants in area A is 104,338. The volume of in-migrants in an area can be obtained by summing up respective column values leaving the diagonal element. E.g. the total number of in-migrants for area A is 2,836 + 1,551 = 4,387. Similarly, the volume of out-migrants for an area can be obtained by summing up the respective rows leaving the diagonal element. E.g. the total number of out-migrants in area A is 2,522 + 939 = 3,461. Additionally, the total net number of migrants is determined by subtracting the out-migrants from the in-migrants. The results can be organised in following table:

Table 9: Calculation of in-migrants, out-migrants, and net migrants for areas A, B, and C

Area	In-migrants	Out-migrants	Net Migrants	Non-migrants
A	4,387	3,461	926	104,338
B	6,735	7,046	-311	372,612
C	5,149	5,764	-615	460,074
All Areas	16,271	16,271	0	937,024

Let us examine the migration that occurred between 1 April 1991, and 1 April 2001, for area A of the country. The same procedure can be applied to areas B and C.

In-migrants to area A during 1 April 1991 and 1 April 2001 = *Persons enumerated on 1 April 2001 in area A and who were residing elsewhere (in area B or C) on 1 April 1991.*

Now these in-migrants can be further divided in following categories:

- *Primary migrants to area A = Persons enumerated in A on 2001 but enumerated elsewhere on 1991 which was their place of birth*
  =Persons enumerated in A on 2001, born in B and their

residence at B on 1991 + Persons enumerated in A on 2001, born in C and their residence at C on 1991

= 2,134 + 877 = 3,011

- *Secondary migrants to area A = Persons enumerated in A on 2001, enumerated in B on 1991 and born in C + Persons enumerated in A on 2001, enumerated in C on 1991 and born in B*

  = 231 + 135

  = 366

- *Return migrants to area A = Persons enumerated in A on 2001, enumerated elsewhere on 1991, born in area A = Persons enumerated in A on 2001, enumerated in B on 1991, born in A + Persons enumerated in A on 2001, enumerated in C on 1991, born in A*

  = 4,984 + 1,098

  = 6,082

Out-migrants from area A during 1 April 1991 and 1 April 2001 *= Persons enumerated on 1 April 1991 in area A and enumerated elsewhere (in area B or C) on 1 April 2001.*

These out-migrants can be further categories into:

- *Primary out-migrants from A = Born in A, enumerated in A on 1991, enumerated in B or C on 2001*

  = 1,435 + 453

  = 1,888

- *Secondary out-migrants from A = Born in B, enumerated in A on 1991 and enumerated in C on 2001 + Born in C, enumerated in A on 1991 and enumerated in B on 2001*

  = 213 + 211

  = 424

- *Return out-migrants from A = Born in B, enumerated in A on 1991 and enumerated in B on 2001 + Born in C, enumerated in A on 1991 and enumerated in C on 2001*

  = 7,200 + 2,434

  = 9,634

Following the similar procedure, it is easy to obtain the volumes of in-migrants, out-migrants, and net migrants for areas B and C of a hypothetical country.

Direct methods of estimating internal migration provide valuable insights into the movement of populations by relying on firsthand data collected through surveys and censuses. These methods typically involve asking individuals about their migration history, including their place of birth, previous residence, and duration of stay at their current location. One of the primary advantages of direct methods is their ability to yield detailed demographic information about migrants, allowing for a comprehensive understanding of the characteristics, motivations, and patterns of migration.

However, it is essential to acknowledge the limitations of direct methods, such as potential biases in self-reported data and challenges in reaching certain populations. Despite these challenges, the richness of the data obtained through direct methods makes them an indispensable tool for understanding internal migration and its impacts on society. Moreover, direct methods enable researchers to capture the dynamics of migration more accurately, as they rely on actual responses rather than estimates or proxies. This can lead to more reliable data, which is crucial for policymakers and planners who need to address the implications of migration on infrastructure, housing, and social services.

1.10 Self-Assessment Exercises

1) What is migration, and how does it differ from mobility?
2) What are the main types of migration, and what distinguishes in-migration from out-migration?
3) What are the primary sources of internal and international migration data?
4) What are the main motivations for migration? What are the economic, demographic, and cultural effects of migration on both the countries of origin and destination?
5) How does migration influence development, including aspects like remittances and brain drain?

6) How do environmental factors like climate change, natural disasters influence migration?

7) The following table contains the distribution of Persons Enumerated in Area A on 1 April 1980 according to their Residence on 1 April 1970 and their Place of Birth.

Place of Birth	**Place of Residence on 1 April 1970**		
	A	**B**	**C**
A	65,504	3,345	795
B	363	1,443	85
C	287	166	556

Next table contains the distribution of Persons Enumerated in Area B on 1 April 1980 according to their Residence on 1 April 1970 and their Place of Birth.

Place of Birth	**Place of Residence on 1 April 1970**		
	A	**B**	**C**
A	995	560	132
B	4,887	238,765	5,765
C	142	837	1,826

Next given table contains the distribution of Persons Enumerated in Area C on 1 April 1980 according to their Residence on 1 April 1970 and their Place of Birth.

Place of Birth	**Place of Residence on 1 April 1970**		
	A	**B**	**C**
A	295	134	194
B	154	2,180	473
C	1,623	11,240	310,112

Determine the amount of in-migration, out-migration, and net migration for the regions A, B, and C as of April 1, 1980, based on the statistics below. Compute the total number of migrants between April 1, 1970, and April 1, 1980.

1.11 Summary

Migration, driven by determinants such as economic prospects, political unrest, and environmental transformations, exercises a significant influence on global population patterns and societal structures. Whether voluntary or compelled, internal or transnational, migration profoundly affects both the individuals who migrate and the communities that receive them. As global challenges, including climate change and economic inequality, become more pronounced, migration is projected to escalate, necessitating more comprehensive policy frameworks and enhanced international collaboration. The evolving dynamics of migration will be critical in addressing labour market demands, promoting cultural exchange, and facilitating social integration. Thus, a deep understanding and strategic management of these processes will be crucial in cultivating inclusive and resilient societies.

1.12 References

- International Organization for Migration. (2024). World Migration Report 2024. Geneva: International Organization for Migration.
- Ministry of Statistics and Programme Implementation. (2023). Periodic Labour Force Survey (PLFS), 2022-23. Government of India. Retrieved from https://mospi.gov.in
- Ministry of Statistics and Programme Implementation. (2022). Migration in India, 2020-2021. Government of India. Retrieved from https://mospi.gov.in
- United Nations Department of Economic and Social Affairs. (1992). Measuring international migration: Many questions, few answers. In Measuring international migration: Manual VI (Chapter 4). United Nations

1.13 Further Readings

- Ram, F., & Pathak, K. B. (1998). Techniques of Demographic Analysis. Mumbai: Himalaya Publishing House.

UNIT – 2 METHODS OF ESTIMATION OF MIGRATION

Structure

2.1 Introduction

The estimation of internal migration is often challenging due to the lack of continuous and comprehensive data, especially in developing countries. In many cases, migration data is only collected during population censuses, which occur at infrequent intervals. To address this data gap, researchers employ indirect estimation methods, such as the residual method, cohort analysis, and vital statistics approaches, which infer migration flows based on available demographic data. These methods, though not without limitations, provide valuable insights into migration trends and patterns that would otherwise be obscured.

In this unit, we delve into the methods used to estimate internal migration, particularly when direct data is scarce or unavailable. We introduce several indirect estimation techniques, such as the residual method, and approaches based on vital statistics. These methods are essential for capturing migration flows and patterns, especially in contexts where reliable data is not consistently collected.

Following this, the unit shifts focus to international migration, the movement of people across national borders. This form of migration has become increasingly significant in the context of globalization, with profound implications for both origin and destination countries. We examine various estimation methods for international migration, including the use of census data, administrative records, and survey data. These methods are essential for capturing the complexity of migration flows, which often involve multiple stages and diverse factors influencing decisions to migrate.

Finally, we explore the theoretical frameworks that have been developed to understand both internal and international migration. Classical theories, such as Ravenstein’s law of migration, gravity model, Lee’s push-pull model, etc. By integrating these theories this unit aims to offer a holistic view of migration dynamics and their broader socio-economic impacts.

2.2 Objectives

After completing this unit, you will be able to:

- Use the indirect method of estimation of net internal migration,
- Understand the basics of international migration and its recent trends with the help of data,
- Explore the sources of data on international migration,
- Utilise the estimation techniques of international migration, and
- Review the various migration theories and models.

2.3 Indirect Estimation of Net Internal Migration

Indirect estimation of internal migration can be derived from the population counts by age and sex or even from overall population totals. The methods of computing indirect estimates of internal migration are:

1) The growth rate method, and
2) The residual methods, comprising of
   a) The vital statistics method, and
   b) The survival rate method.

2.3.1 National Growth Rate Method

This is a very easy and popular method. Let P_T^0 and P_T^1 be the total national population at the beginning and end of the intercensal period, respectively. Also, let P_i^0 and P_i^1 be the populations of i^{th} geographical divisions at the beginning and end of the period. The estimated migration rate m_i, for area i is given by

$$m_i = \left[\frac{P_i^1 - P_i^0}{P_i^0} - \frac{P_T^1 - P_T^0}{P_T^0}\right] \times K$$

Where K is a constant, commonly taken to be equal to 1,000. The positive value of m_i denotes the net in-migration rate and the negative value shows the net out-migration rate. You can use the above procedure to determine migration rates by age and sex groups. This method assumes that the rates of natural increase and net immigration are uniform across different regions of the country. The main advantage is that it does not require vital statistics, making it a valuable tool for estimating internal migration in developing countries where the registration of vital events may be unreliable.

2.3.2 Residual Methods

Some residual methods are

2.3.2.1 Vital Statistics Method

Vital statistics method is based on the use of the **balancing equation** viz.

$$M = (P^1 - P^0) - (B - D)$$

Where

M is volume of international migration in a region,

P^0 & P^1 are population of the region on first and second censuses, respectively, and

B & D are total births and deaths, respectively, between the intercensal period in the same region.

It provides an estimate of the net migration volume for the area for which births and deaths reported there throughout the intercensal period, as well as population counts at the start and end of the period, are known. The in- and out-migrants who passed away before to the second census are both included in this estimate of net migration. If the population and the essential data are published according to a characteristic, this approach may be used to estimate net migration for a certain sex, race, or nativity group, as well as for any other characteristics that do not vary over time.

A shift in the borders of the geographic region is one factor that might cause an inaccuracy in these migration estimates of net migration. The necessary data are often accessible whenever territory is transferred from one state to another. Important statistics, as determined by comparing old and new maps of the higher units, such as the state, may be credited to the new region in the event that a portion of these administrative units is moved. Since further transfers take place throughout the intercensal period, there is no need to adjust vital statistics for the upcoming years in the new geographic region. However, a specific section has to be prepared if the transfer happens throughout the year and vital data are available for the entire area for the entire year.

2.3.2.2 Survival Rate Method

The survival ratio method is a widely used approach for estimating net internal migration between two census periods. This method relies on applying survival ratios to the population at the first census to derive an estimate of the expected population at the second census. The difference between the expected population and the actual enumerated population at the second census is then used as an estimate of net migration

The formula used to estimate net migration is:

$$M_x = P_{x+t}^1 - S\, P_x^0$$

Where;

M_x represent net migration,

P_x^0 and P_{x+t}^1 denoted the population aged x and $x + t$ years at first and second census respectively,

t denotes the gap between two censuses, and

S is the survival rate of the population aged x at the first census.

$(S \times P_x^0)$ expressed the expected population based on the survival rate S. So, the net migration is the difference between observed population and expected population.

Mainly, there are two types of survival rates used to estimate net migration volume namely:

- Census Survival Rate
- Life table Survival Rate

2.3.2.3 Census Survival Rate Method

A census survival rate at specific age x is the ratio of the population aged $x + t$ at the second census to the population aged x at the first census. The censuses are taken t years apart.

The census survival rate denoted by S_x^t is given as

$$S_x^t = \frac{P_{x+t}^1}{P_x^0}$$

Now, the estimate of net migration for an area 'a' for specific age group can be calculated as

$$M_x = P_{x+t,a}^1 - S_x^t \, P_{x,a}^0$$

Where P_x^0 and P_{x+t}^1 denoted the total population aged x and $x + t$ years at first and second census respectively, and

$P_{x,a}^0$ and $P_{x+t,a}^1$ are the population of area 'a' aged x and $x + t$ years at first and second census respectively.

Two fundamental assumptions underlie this method: (1) the local area's mortality conditions are the same as those of the country; and (2) the national and local area levels of census age data exhibit the same pattern of errors.

The estimates of the local areas produced by this method reflect net immigration and net internal movement instead of representing net internal migration alone. It is not surprising that some of the survival rates thus determined could be very low or surpass unity due to net immigration, census misreporting of age, and coverage issues.

2.3.2.4 Life Table Survival Rate

Life table survival rate depends on several key components. The columns of life table used to determine the survival rate are:

- ${}_nL_x$ represents the number of individuals surviving in the age interval $(x, x + n)$ and
- ${}_nT_x$ indicates the total number of person-years lived by the survivors aged between $(x, x + n)$ in the future.

To calculate survival rate for specific age interval (for example 10 years), the formula is given as:

$$_{10}S_x = {}_5L_{x+10} / {}_5L_x$$

If the census age distribution has an open-ended interval, such as 70+,

$$_{10}S_{70+} = T_{80} / T_{70}$$

After determining survival rate, we can use it to calculate net migration as

$$M_x = P^1_{x+t} - S_x\, P^0_x$$

Example: Suppose a region has 10,000 individuals aged 30-35 in 2010, and total population of same age group after 10 years in the region in 2020 is 10,300. Further suppose the life table survival rate from age 30-35 to 40-45 is 0.95. What is the net migration volume into the region?

Solution: We have given,

${}_5P^0_{30}$ = 10,000; ${}_5P^1_{40}$ = 10,300; ${}_{10}S_{30}$ = 0.95;

The expected survivors of the region in 2020,

${}_{10}S_{30}\ {}_5P^0_{30}$= 0.95 × 10,000 = 9,500

The estimated net migration volume,

M = 10,300 – 9,500 = 800

i.e. the total volume of net migrants is 800 in the region.

It is crucial that the life table chosen reflects the average mortality conditions during the intercensal period and is suitable for the specific area where estimates are needed. If a life table specific to the area isn't available, but the average mortality rate for the period is roughly known, model life tables can be used to estimate survival rates. Additionally, if the intercensal period differs from the standard 5 or 10 years, complications arise. Survival rates must be calculated for the exact duration of the period, which may require additional effort, especially if only an abridged life table is available. If survival data will appear in non-standard 5-year age groups, it must be adjusted to fit conventional 5-year age groups for comparison with the second census.

2.4 International Migration

Population movements which cross national boundaries merit separate study because they are the fundamental aspect of the modern world, deeply influencing the demographic, economic, and social landscapes of nations. This movement, which involves crossing international borders, is driven by a variety of circumstances, such as the desire to flee unstable political environments, better economic possibilities, and environmental changes. Migration has become an important factor in the dynamics of the world population as a result of the increased prominence of cross-border movement brought about by globalization. This process is a reflection of the complex interactions between personal goals and global trends, which collectively influence the political and socioeconomic systems of both the countries of origin and destination.

International migration has a particularly significant effect on the economy. In their new countries of residence, immigrants frequently help close skills and perspective gaps in the job market, stimulate economic growth, and broaden the pool of available talent. The host economy may gain from this inflow of human capital by encouraging entrepreneurship and innovation. But there are drawbacks as well, like the requirement for social integration and the strain on public services. Furthermore, the presence of migrants can occasionally result in xenophobia and social unrest, calling for the implementation of policies that support inclusion and attend to the worries of both host populations and migrants.

In countries of origin, the effects of migration are similarly complex. While migration can lead to a loss of skilled labour, or brain drain, which may impede economic development, it also generates remittances that are crucial for the livelihoods of many families and for the broader economy. These financial flows can reduce poverty, improve education and health outcomes, and contribute to economic growth. However, the reliance on remittances can also create dependencies and may discourage local investment and development. Balancing these outcomes requires careful policy considerations that address the needs of both migrants and their home countries.

Beyond the economic aspects, international migration has significant social and cultural implications. Migrants often bring diverse cultural practices and perspectives, enriching the social fabric of destination countries. This cultural exchange can lead to greater mutual understanding and the broadening of societal norms. However, it can also challenge existing social structures and lead to conflicts over identity and values. In origin countries, migration can result in changes in family dynamics and community structures, as well as shifts in cultural practices. These social changes, while often positive, can also be sources of tension and require careful management.

On a personal level, the decision to migrate is often driven by the desire for a better life, whether through improved economic prospects, educational opportunities, or enhanced safety. However, the journey is not without its challenges. Migrants may face legal hurdles, social exclusion, and the emotional toll of leaving their home and loved ones behind. Despite these difficulties, migration remains a powerful force in shaping individual destinies and global demographics. Understanding the intricacies of international migration is crucial for developing policies that promote inclusive, equitable, and sustainable societies in an increasingly interconnected world.

2.4.1 Recent Trends of International Migration

The World Migration Report 2024 indicates that the total number of international migrants worldwide has reached 281 million, with a notable increase in the number of displaced individuals due to various crises, highlighting the complexity of global migration dynamics.

Key Statistics

- International Migration at a Glance: The following table shows key statistics of international migration data on 2024.

Estimated number of international migrants	281 million
Estimated proportion of world population who are migrants	3.6%
Estimated proportion of female international migrants	48.0%
Estimated proportion of international migrants who are children	10.1%
Region with the highest proportion of international migrants	Oceania
Country with the highest proportion of international migrants	U.A.E.
Global international remittances (USD)	831 billion
Number of refugees	35.4 million
Number of migrant workers	169 million
Number of internally displaced persons	71.4 million

Figure 1: International migration at a glance (Source: World Migration Report 2024)

- **Emigration Growth:** Recent patterns and trends in international migration from India reveal significant changes in both the scale and nature of migration, driven largely by educational factors and economic opportunities abroad. Between 1990 and 2020, the number of Indian emigrants more than doubled, reaching approximately 17.9 million, making India the largest country of origin for international migrants. Despite this increase, India's share of global migration rose only slightly from 4.4% to 6.3% during this period.

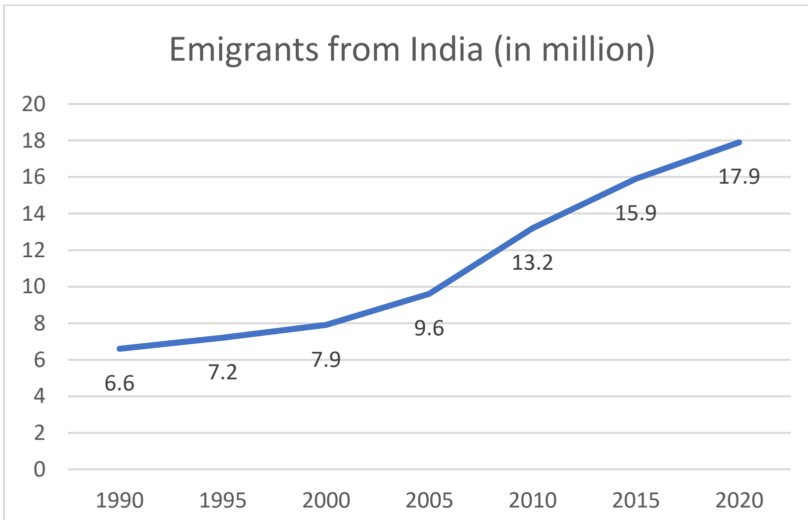


Figure 2: Emigrants from India (Source: UN DESA 2020)

- **Destination Countries:** The primary destinations for Indian emigrants include the United Arab Emirates (UAE), United States (US), and Saudi Arabia. These countries host the largest Indian diaspora, reflecting ongoing trends in labour migration.

- **Demographic Patterns:** The demographic composition of Indian migrants shows that 65% of emigrants are male, primarily migrating for work, while many women remain in India, often due to socio-cultural reasons. In contrast, 86.8% of female migrants typically move for marriage.

- **Remittances:** Indian migrants contribute significantly to the economy through remittances, which have surged globally. In 2022, international remittances sent to low- and middle-income countries reached $831 billion, with a substantial portion attributed to

Indian migrants. This trend underscores the economic importance of migration for both the migrants and their home country.

In 2022, India, Mexico, and China were (in descending order) the top three remittance recipient countries, although India was well above the rest, with total inward remittances exceeding $111 billion, the first country to reach and even exceed $100 billion.

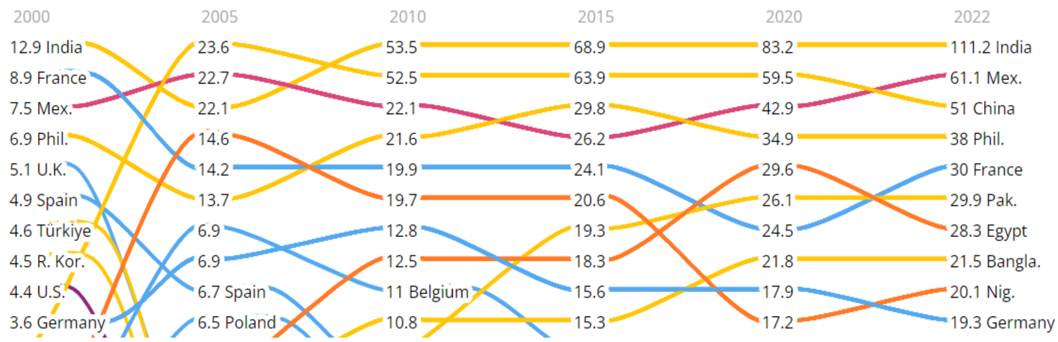


Figure 3: Changing Trends in Remittance Recipient countries (Source: World Migration Report 2024).

High-income countries are almost always the main source of remittances. For decades, the United States has consistently been the top remittance-sending country, with a total outflow of $79 billion in 2022, followed by Saudi Arabia ($39 billion), Switzerland ($31.9 billion), and Germany ($25.6 billion).

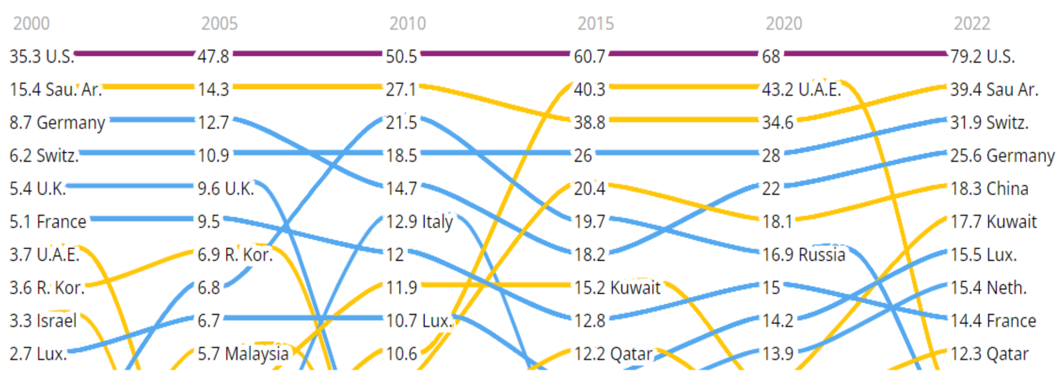


Figure 4: Changing Trends in Remittance Sources (Source: World Migration Report 2024)

- **Net Migration of India:** Net international migration of India refers to the difference between the number of people leaving the country (emigrants) and those entering it (immigrants). India has historically had a high-rate of emigration, with millions of Indians moving abroad for better economic opportunities, education, and improved living standards, particularly to countries like the United States, the United Kingdom, Canada, and the Gulf nations. This has resulted in a large Indian diaspora contributing to remittances, which play a crucial role in the country's economy. On the other hand, immigration to India is relatively low, with most immigrants coming from neighbouring countries such as Bangladesh, Nepal, and Myanmar, often driven by socio-political unrest or economic opportunities. The net effect is that India has a negative net international migration rate, meaning more people are leaving the country than entering it.

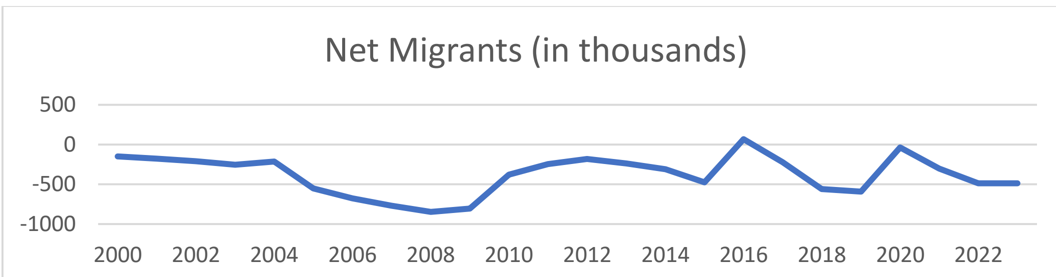


Figure 5: Net Migration of India (Source: World Bank)

2.4.2 Need of Data on Internation Migration

- **Understanding Global Migration Trend**: International migration data is essential for understanding the patterns and trends of people moving across borders. As globalization increases, so does the movement of individuals seeking better opportunities, safety, or improved living conditions. Accurate data helps policymakers, researchers, and international organizations track these movements, analyse demographic changes, and predict future trends by researchers. Without reliable data, it would be challenging to understand the full scope of migration and its impacts on both sending and receiving countries.

- **Informing Policy and Decision Making**: Governments and international organizations rely on robust migration data to craft informed policies that address the challenges and opportunities presented by migration. For instance, data on the skills and professions of migrants can help receiving countries adjust their immigration policies accordingly to fill labour shortages. Similarly, information on emigration patterns can help sending nations develop policies that will reduce brain drain and increase diaspora involvement. Moreover, accurate data is crucial for developing effective refugee and asylum policies, ensuring that those in need receive appropriate protection and support.

- **Economic and Social Impact**: International migration has significant economic and social implications, both for the countries migrants leave and those they move to. Remittances sent by migrants back to their home countries are a critical source of income for many developing nations, contributing to poverty reduction and economic development. On the other hand, the arrival of migrants can affect local labour markets, public services, and social dynamics in receiving countries. Reliable migration data allows for a better understanding of these impacts, enabling governments to balance economic benefits with social integration and cohesion. Accurate data is also vital for addressing challenges such as human trafficking, ensuring that international migration occurs in a safe, orderly, and regulated manner.

- **Managing Population Growth and Urbanization**: International migration data plays a crucial role in managing population growth and urbanization in both sending and receiving countries. In countries with high emigration rates, understanding who is leaving and where they are going helps to manage population decline, address labour shortages, and adjust economic policies accordingly. Conversely, in countries with high immigration rates, data helps cities and regions prepare for population growth by informing infrastructure development, housing needs, and public service provision. As urban areas often become the primary destinations for migrants, data is essential for ensuring that cities can accommodate the influx of new residents while maintaining sustainable growth and quality of life.

- **Human Right and Social Justice**: Accurate data on international migration is also essential for protecting the human rights of migrants and promoting social justice. Migrants, particularly those who move under precarious conditions, can be vulnerable to exploitation, discrimination, and abuse. Data helps to identify these vulnerable populations, enabling targeted interventions to protect their rights and improve their living conditions. Moreover, data on migration flows and experiences can inform the development of legal frameworks that promote fair treatment of migrants, combat xenophobia, and support their integration into new communities

- **Global Cooperation and Development**: International migration is a global phenomenon that requires coordinated efforts across borders. Reliable migration data supports international cooperation by providing a common understanding of migration trends and challenges. This data is crucial for multilateral agreements, such as the Global Compact for Migration, which aims to facilitate safe, orderly, and regular migration worldwide. Furthermore, data on migration contributes to global development goals, particularly in areas such as reducing inequality, promoting decent work, and ensuring access to education and healthcare for all, including migrants. By sharing data and insights, countries can work together to harness the benefits of migration while addressing its complexities in a collaborative and effective manner.

2.4.3 Sources of Data on Internation Migration

The collection of international migration data is multi-layered, involving various methods that capture different aspects of the migration process. These methods can be categorized into three main sources: border collection, registration, and field inquiry. Each of these categories plays a distinct role in gathering vital information about who is migrating, why they are moving, and how migration affects both the countries of origin and destination.

2.4.3.1 Border Control Records

Border collection refers to data gathered directly at points of entry and exit, such as airports, seaports, and land borders. This method typically involves the use of immigration and customs controls, where officials record information about individuals crossing national borders. Key data points include the number of arrivals and departures, nationality, visa type, and purpose of travel. While border collection provides real-time and relatively accurate data on migration flows, it has limitations, such as the difficulty in tracking individuals who overstay their visas or those entering through irregular channels. Nevertheless, it remains a vital source of information for understanding short-term migration patterns and monitoring compliance with immigration regulations

2.4.3.2 Registration

Registration involves the systematic recording of individuals' information by government authorities or international organizations. This category includes a wide range of data sources, such as population registers, asylum applications, work permits, and residence permits. For example, many countries require immigrants to register their presence with local authorities upon arrival, providing detailed data on their demographic characteristics, employment status, and family composition. Registration data is particularly valuable for tracking long-term migration trends and understanding the socioeconomic integration of migrants. Additionally, refugee registration by agencies like the United Nations High Commissioner for Refugees (UNHCR) offers critical insights into forced migration and the needs of displaced populations.

2.4.3.3 Field Inquiry

Field inquiry involves collecting migration data through surveys, censuses, and interviews conducted by researchers, government agencies, or international organizations. This method is often used to gather detailed, qualitative data on migrants' experiences, motivations, and living conditions, which may not be captured through border collection or registration. Field inquiries can include national censuses that ask about migration history, household surveys that explore remittance patterns, or targeted studies on specific migrant groups. Although field inquiry can be time-consuming and resource-intensive, it provides a deeper understanding of migration dynamics, including irregular migration and the challenges faced by migrant communities. This method is particularly important for capturing data on undocumented migrants, who might not appear in official records.

Together, these three categories—border collection, registration, and field inquiry—form a comprehensive framework for collecting and analysing international migration data. Each method has its strengths and limitations, but when combined, they offer a robust picture of global migration trends and their implications.

2.4.4 Development of International Migration Data Sources

Recent developments in the data sources of international migration have been driven by advancements in technology, greater international cooperation, and the need for more accurate and timely information. These developments are enhancing our ability to monitor and understand migration patterns globally. Here are some key recent trends.

- **Big Data and Digital Technologies:** The use of big data and digital technologies is revolutionizing the collection of migration data. Sources such as mobile phone records, social media activity, and online job portals are increasingly being used to track migration flows and migrant behaviour in real-time. These non-traditional data sources provide insights into migration patterns that were previously difficult to capture, such as short-term or circular migration and undocumented movements. Additionally, digital tools like biometric identification and e-visa systems are improving the accuracy of data collected at borders and registration points.

- **Enhanced International Cooperation:** There has been a growing emphasis on international cooperation to improve the collection and sharing of migration data. Initiatives like the Global Compact for Safe, Orderly, and Regular Migration and the establishment of the International Migration Review Forum are promoting the standardization of data collection methods and encouraging countries to share data more transparently. These efforts aim to create a more comprehensive and harmonized global migration database, which can be used to inform policy decisions and enhance global governance of migration.

- **Integrated Data System:** Countries and international organizations are increasingly integrating different data sources to create more comprehensive migration databases. For example, linking border control data with national population registers and labour market information allows for a more detailed analysis of migration's impact on both origin and destination countries. The integration of different data systems also helps to reduce duplication, improve data quality, and provide a more holistic view of migration trends.

- **Improved Data on Refugees and Forced Migration:** There have been significant improvements in the data collection processes related to refugees and forced migration, driven by the need to respond to humanitarian crises more effectively. The United Nations High Commissioner for Refugees (UNHCR) and other organizations are increasingly using digital tools for refugee registration, enabling real-time updates and more accurate tracking of displaced populations. These advancements have improved the ability to monitor refugee flows, assess needs, and allocate resources more efficiently.

- **Citizen Science and Crowdsourcing:** Another emerging trend is the use of citizen science and crowdsourcing to collect migration data. Projects that engage local communities and migrants themselves in data collection efforts are providing new insights into migration dynamics. These approaches can capture data on undocumented migrants or migration routes that are not covered by traditional methods, offering a more inclusive understanding of migration.

These recent developments are significantly enhancing the quality, timeliness, and comprehensiveness of international migration data, providing a more accurate basis for research, policy-making, and international collaboration.

2.4.5 Estimation of International Migration

We can estimate of the volume of net immigration in addition to the data on international migration that comes from direct or indirect sources. The number of net immigrants is the difference between the number of people entering a country (immigrants) and the number of people leaving the country (emigrants) over a specific period of time.

2.4.5.1 Estimation of Net Immigration by Intercensal Component Method

To estimate the total volume of net immigration during an intercensal period, a modified version of the standard intercensal component equation, known as the balancing equation, is used:

$$I - E = (P_1 - P_0) - (B - D)$$

The method simply involves subtracting an estimate of natural increase $(B - D)$ during the period from the net change in population during the period $(P_1 - P_0)$. Net immigration $(I - E)$ is thus derived as residual. However, since census data and records of births (B) and deaths (D) may contain errors, the resulting estimate of net immigration can be inaccurate. The relative error in estimate of net immigration may be considerable when the amount of international migration is small. This method has been applied below to estimate net immigration for a hypothetical population.

S.No.		
1	Population according to census 2001	1,554,370
2	Population according to census 2011	1,610,487
3	Net change $(2) - (1)$	56,117
4	Births	144,879
5	Deaths	99,854
6	Natural increase $(4) - (5)$	45,025
7	Residual estimate $(3) - (6)$	11,092

Theoretically, since immigrants often belong to a more limited universe, it makes sense to estimate net immigration more accurately from the census data on foreign-born people rather than from the whole population. However, this approach makes the assumption that all migrants are foreign-born individuals whose fertility and mortality can be precisely determined. When calculating net immigration, it is best to utilize the total or native population, since this will indicate that the majority of immigrants are native-born people.

2.4.5.2 Intercensal Cohort Component Method

This approach may be used to estimate net migration by age for both the entire population and specific demographic segments. Compiling death statistics by age cohorts to account for the mortality component is a particularly time-consuming step in creating estimates by age. Utilizing survival rates, such as those from life tables or national census, is generally advised. In order to account for age cohorts other than those born during the intercensal era, the estimation formula for this purpose is

$$I_a - E_a = P_a^1 - S(P_{a-t}^0)$$

Where I_a and E_a represent immigrants and emigrants in a cohort defined by age ‘a’ at the end of the period,

P_{a-t}^0 is the population t year younger at the first census, and

S is the survivorship rate for the age a during the intercensal period of t years.

The intercensal cohort component method estimates net immigration by tracking changes in age-sex cohorts between two censuses. First, the population is divided into cohorts by age and sex at the time of the first census. Each cohort is then projected forward to the second census, accounting for natural changes like aging, births, and deaths using survival rates. The expected population for each cohort is calculated, and this is compared to the actual population recorded in the second census. The difference between the expected and actual populations for each cohort is attributed to net immigration, with any excess indicating positive net immigration (more arrivals than departures) and any shortfall suggesting net emigration (more departures than arrivals).

The population's age distribution at census locations is generally reported in quinquennial age-groups. Additionally, inaccurate age reporting of individuals of both sexes affects the age data. In these circumstances, the age data can be adjusted using a suitable smoothing technique. However, excessive data smoothing might have an impact on the estimate of net immigration by age without changing the estimate of total net immigration.

2.5 Migration Model

Any process may be modelled in two steps: (i) a theoretical description of the process; and (ii) the conversion of a theoretical description into a mathematical model. The theories of migration attempt to explain the causes for migration. In the second step, researchers attempt to parameterize the procedure in order to make it more broadly applicable and useful for planning. Following are the theories which have emerged from various studies on migration

2.5.1 Ravenstein's Law of Migration

Ernst Georg Ravenstein, a pioneering geographer, developed a set of principles known as "Ravenstein's Laws of Migration" in the late 19th century. These laws, derived from his analysis of British census data, offer foundational insights into migration patterns and have significantly influenced the study of migration.

- **The law of Short Distances:** Ravenstein observed that most migrants move short distances. This is because nearby destinations are easier and less costly to reach, and allow migrants to maintain connections with their origin communities. As the distance between two locations increases, the likelihood of migration decreases—a concept known as the distance-decay effect.

- **The Law of Migration by Stages:** Migration often occurs in steps rather than in one direct move. For example, a migrant might first move from a rural area to a nearby town, and then to a larger city. This "step migration" reflects a gradual approach to reaching larger, more economically vibrant areas.

- **The law of Long-Distance Migrants Heading to Major Cities:** When migrants do move long distances, they tend to head towards major urban centres. These cities attract migrants with their greater economic opportunities, better infrastructure, and broader range of services. Ravenstein noted that large cities serve as focal points for long-distance migration.

- **The Law of Urban to Rural Migration:** Ravenstein found that urban residents are less likely to migrate compared to those in rural areas. Urban areas typically offer more opportunities and better living conditions, reducing the incentive to move. In contrast, limited opportunities in rural areas often push residents to migrate.

- **The Law of Counter-Flow Migration:** Ravenstein observed that every migration flow generates a counter-flow. For instance, while many people migrate from rural to urban areas, some urban dwellers move back to rural areas for reasons like retirement or lifestyle changes. This demonstrates the dynamic nature of migration.

- **The Law of Rural to Urban Migration:** One of the most significant trends Ravenstein identified is the movement from rural to urban areas. This reflects the pull of cities, driven by industrialization and the promise of better jobs and living conditions, which contrasts with the push factors of rural areas, such as limited land and employment.

- **The Law of Migration by Gender and Age:** Ravenstein noted that migration patterns often vary by gender and age. Women were more likely to migrate short distances, often within their own country, while men were more likely to migrate over long distances, including internationally. Young adults, seeking employment and new opportunities, were the most likely to migrate.

- **The Law of Technology and Migration:** Ravenstein acknowledged that advancements in transportation and communication played a crucial role in facilitating migration. As technologies improved, the cost and difficulty of long-distance travel decreased, making migration more accessible. Improved communication also helped migrants stay connected with their origin, easing the transition.

Ravenstein's Laws of Migration provided some of the earliest systematic insights into migration behaviour and continue to be influential in migration studies. While some of his laws are specific to the 19th century, many remain relevant today, offering valuable perspectives on the forces that shape migration patterns. His work laid the groundwork for future theories that incorporate more complex factors, such as economic conditions, social networks, and policy influences on migration.

2.5.2 Gravity Model

The Zipf's Gravity Model of Migration is a widely used framework to predict and analyse migration flows between locations. This model draws an analogy from Newton's Law of Gravity, which asserts that the force between two objects is proportional to their masses and inversely proportional to the square of the distance between them. Similarly, the Gravity Model suggests that migration flows between two locations are influenced by the sizes of their populations and the distance between them. Essentially, it posits that larger locations with higher populations will attract more migrants, while the distance between locations will act as a deterrent to migration.

The fundamental equation of the Gravity Model is expressed as

$$M_{ij} = \frac{P_i \times P_j}{D_{ij}}$$

Where M_{ij} represents the migration flow from location i to location j;

P_i and P_j are the population of the two locations, and

D_{ij} is the distance between them.

According to this model, the migration flow increases with the population sizes of the locations and decreases as the distance between them grows. This implies that larger cities or regions, which offer more job opportunities and better amenities, are more likely to attract migrants, while greater distances reduce the likelihood of migration due to increased travel costs and logistical challenges.

The Gravity Model has several practical applications. In urban planning, it helps predict migration flows between cities or regions, aiding in the development of infrastructure and transportation systems. For regional development, it provides insights into how changes in population size or economic conditions can impact migration patterns, informing strategies to address regional imbalances. Additionally, businesses can use the model to estimate potential market sizes in different regions based on migration patterns, guiding decisions on where to expand operations.

However, the Gravity Model has limitations. Its simplicity means it may overlook other critical factors influencing migration, such as social networks, personal preferences, or political conditions. While the model incorporates distance as a variable, it may not fully capture how improvements in transportation and communication can mitigate the impact of distance. The model also assumes that migration decisions are based solely on population size and distance, potentially neglecting other variables such as cultural ties or legal constraints. Furthermore, it often does not account for temporal changes in migration patterns, which can vary due to economic cycles, policy shifts, or other factors.

To address these limitations, modern adaptations of the Gravity Model often incorporate additional variables, such as economic indicators, quality of life measures, or migration policies, to provide a more nuanced understanding of migration flows. Expanded models may also integrate spatial interactions and the impact of intervening opportunities, while network models acknowledge the role of existing migrant communities in shaping new migration patterns. Despite its limitations, the Gravity Model remains a foundational tool in migration studies, offering valuable insights into how population size and distance influence migration behaviour.

2.5.3 Lee's Theory of Migration

Everett S. Lee, an American sociologist, expanded the study of migration by developing a comprehensive framework that addresses both the causes and processes of migration. His theory, presented in the 1966 paper "A Theory of Migration," built on earlier models, such as those proposed by Ravenstein, but introduced new concepts to explain the complexities of migration. Lee's theory is particularly notable for its focus on the interplay of push and pull factors, intervening obstacles, and personal factors that influence an individual's decision to migrate.

- **Push and Pull Factors:** At the core of Lee's theory are the concepts of push and pull factors, which describe the forces that drive people to leave their current location (push factors) and attract them to a new one (pull factors). Push Factors are negative conditions at the place of origin that encourage people to move away. Examples include unemployment, poverty, political instability, environmental disasters, and lack of opportunities. Push factors essentially create pressure on individuals to seek better circumstances elsewhere. Pull Factors are positive conditions at the destination that attract individuals to move there. Examples include better job opportunities, higher wages, political stability, better living conditions, and educational opportunities. Pull factors create a sense of opportunity or improvement that draws people to a new location.
- **Intervening Obstacles:** Lee also highlighted intervening obstacles, which are barriers that can hinder migration. These include physical barriers (like distance and geography), economic barriers (such as the cost of moving), legal obstacles (immigration laws and visa requirements), and social factors (family ties or cultural differences). The ability to overcome these obstacles significantly impacts whether migration occurs.

- **Personal Factors:** Personal characteristics also play a crucial role in migration decisions. Factors such as age, gender, education, and personal preferences influence how individuals perceive push and pull factors and how they navigate obstacles. For example, young adults are more likely to migrate in search of opportunities, while gender roles can shape migration patterns in different cultures.

- **Perception and Decision-Making:** Lee emphasized that migration is a selective process, heavily influenced by individual perceptions. People weigh the perceived benefits of migrating against the costs and obstacles. If the advantages of moving seem greater than the risks, migration is more likely to occur. This subjective evaluation explains why migration patterns can vary widely even under similar conditions.

Lee's Theory of Migration provides a nuanced understanding of migration by considering the interplay of various factors. It explains both voluntary and involuntary migration, as well as internal and international migration. The theory's flexibility and focus on individual decision-making have made it a foundational tool in migration studies, offering insights into the diverse and complex nature of migration.

2.5.4 Harris-Todaro Model of Migration

The Harris-Todaro model, developed by economists John R. Harris and Michael P. Todaro in 1970, is a key theoretical framework used to explain rural-to-urban migration in developing countries. The model was created in response to the rapid urbanization observed in many developing nations during the mid-20th century, which often led to urban unemployment and the expansion of informal sectors. Unlike earlier models that assumed people migrate solely for higher wages, the Harris-Todaro model introduces the concept of expected income and incorporates the role of urban unemployment in migration decisions.

The Harris-Todaro model is built around the idea that individuals make migration decisions based not on actual wage differentials between rural and urban areas but on expected wage differentials. The expected wage is calculated as the urban wage multiplied by the probability of finding a job in the urban area. This probability is determined by the ratio of jobs available to the urban labour force, which includes both employed and unemployed workers.

The basic equation of the model is:

$$E(W_u) = W_u \times \left(\frac{L_u}{L_s \times L_u}\right)$$

Where:

$E(W_u)$ is the expected urban wage,

W_u is the actual wage in the urban sector,

L_u is the number of jobs available in the urban sector, and

L_s is the number of unemployed workers in the urban sector.

Individuals in rural areas decide whether to migrate based on a comparison between the expected urban wage and the rural wage. Migration continues until the expected urban wage equals the rural wage, leading to an equilibrium where no additional rural workers are incentivized to move to the city.

2.5.4.1 Proposition of the Model

- **Urban Unemployment:** The Harris-Todaro model explains why rapid urbanization in developing countries often leads to high urban unemployment. As rural migrants move to cities in search of better opportunities, many cannot find immediate employment, resulting in a surplus labour force and rising unemployment rates.

- **Informal Sector Growth:** The model highlights the expansion of informal sectors in urban areas. With more people migrating than the formal economy can absorb, many end up in low-wage, informal jobs, which become a significant part of the urban economy.

- **Policy Implications:** The model suggests that focusing solely on urban job creation may worsen unemployment. Policies that increase urban wages can attract more migrants, exacerbating the problem. Effective strategies should also promote rural development, improving rural incomes and reducing the incentive to migrate.

- **Rational Migration Decision:** Migration is shown as a rational choice based on expected, not actual, income. Even with high urban unemployment, rural workers may still migrate if the potential urban income, factoring in the chance of finding a job, exceeds their rural earnings.

- **Dynamic Equilibrium:** The model predicts that migration continues until expected urban and rural incomes equalize. However, this equilibrium is often unstable due to changing economic conditions and policies, leading to persistent urban growth, unemployment, and underemployment in many developing countries.

The Harris-Todaro model has been extended and critiqued over the years. Some extensions include incorporating factors like housing costs, migration costs, and government interventions. For example, if housing costs in urban areas are high, this could reduce the expected urban income, potentially slowing migration. Additionally, government policies aimed at reducing urban unemployment or improving rural livelihoods can alter the migration dynamics predicted by the model. Critics of the Harris-Todaro model argue that it oversimplifies the migration decision by focusing primarily on economic factors, neglecting social, cultural, and psychological influences on migration. Others point out that the model assumes rational behaviour and perfect information, which may not always reflect real-world conditions, especially in developing countries where access to information can be limited.

Despite its limitations, the Harris-Todaro model remains a foundational theory in the study of migration economics. It provides a clear framework for understanding the relationship between rural-to-urban migration, urban unemployment, and expected wages. The model has been particularly influential in shaping development policies and has inspired a wide range of empirical studies and policy interventions aimed at managing urbanization and its associated challenges in developing countries.

2.5.5 New Economics of Labour Migration

The New Economics of Labor Migration (NELM) emerged in the 1980s as a response to the limitations of earlier migration theories, emphasizing the role of households, social networks, and market imperfections in migration decisions. Unlike traditional models that focused on individual decisions based on wage differentials, NELM views migration as a household strategy aimed at improving overall welfare.

- **Household Decision-Making and Risk Diversification:** NELM focuses on households rather than individuals as the primary decision-making units. Households use migration as a strategy to diversify income sources and manage risks, particularly in the face of local economic uncertainties like crop failures or income instability. Migration helps households overcome market failures, such as the lack of access to credit or insurance, by sending remittances that can finance investments in agriculture, education, or small businesses.

- **Role of Remittances:** Remittances are central to NELM and are viewed as a critical aspect of migration. They are used to improve household welfare, reduce poverty, and finance productive investments. Unlike previous models that saw remittances as secondary, NELM emphasizes their role in economic development and social mobility. Remittances also help households achieve specific goals, such as building homes or funding education, making migration a temporary strategy in many cases.

- **Social Networks and Cumulative Causation:** NELM highlights the importance of social networks in facilitating migration. These networks reduce the costs and risks associated with migration by providing information, financial support, and assistance in the destination area. The concept of cumulative causation, where migration becomes self-perpetuating as networks grow, is also a key element of NELM.

NELM has been criticized for potentially overemphasizing the rationality of households and focusing primarily on economic motivations, possibly neglecting other factors like political or environmental influences. However, its contributions to understanding migration have been significant, particularly in highlighting the role of remittances and household strategies. NELM has influenced policy discussions on rural development, poverty reduction, and the management of remittances, making it a pivotal framework in migration studies and development economics.

2.6 Self-Assessment Exercises

1) What are the key challenges in estimating international migration, and how do they impact the accuracy of migration data?
2) Compare and contrast the different methods used to estimate international migration, including direct and indirect estimation techniques. What are the strengths and limitations of each approach?
3) How do demographic factors such as age, sex, and education influence international migration patterns? Use specific examples to illustrate your answer.
4) What role do migration networks and social capital play in shaping international migration flows?
5) Evaluate the use of administrative data sources (e.g., border control records) in estimating international migration. What are the advantages and limitations of these data sources, and how can they be used in conjunction with other estimation methods?

6) A town had a population of 75,000 at the beginning of 2022. By the beginning of 2023, the population had increased to 76,500. During 2022, the town recorded 1,200 births and 900 deaths. Calculate the net migration for the town in 2022.

7) A country conducted a census in 2015 and reported a population of 5,000,000. Another census was conducted in 2020, revealing a population of 5,200,000. During this period, the country experienced 300,000 births and 200,000 deaths. Assume the country's life table indicates that the survival rate for individuals aged 0-4 years is 0.98, and for those aged 5-9 years is 0.95. Estimate the number of people who would have been in the 5-9 age group in 2015 if there had been no migration, using the survival rate. Also, calculate the net migration for the period 2015-2020 using the residual method.

8) The national statistical office of Country A is trying to estimate its annual net international migration for the past year. They have gathered the following data:

Data Type	**Country**		
	B	**C**	**D**
Immigration Data	45,000	30,000	25,000
Emigration Data	15,000	20,000	10,000

Also,

Total population of Country A at the beginning of the year: 10,000,000

Total population of Country A at the end of the year: 10,025,000

Using the data provided, answer the following questions:

a) Calculate the total number of immigrants and emigrants for Country A.

b) Determine the net international migration for Country A for the past year.

c) Estimate the net migration rate per 1,000 population.

2.7 Summary

This unit examines the estimation of indirect estimation of internal migration volume, explore the concept of international migration, its recent trends in the world and in India, sources of international migration data, and various methods to estimate international migration volume. It highlights the challenges in accurately estimating migration due to diverse data sources, including border records like visa applications, field enquiries like surveys, and other administrative records, each with its limitations. The chapter then explores key migration models, such as gravity model and Lee's push and pull theory, and the New Economics of Labor Migration (NELM). These models collectively provide a nuanced understanding of the drivers and patterns of international migration.

2.8 References

- International Organization for Migration. (2024). World Migration Report 2024. Geneva: International Organization for Migration. Retrieve from https://worldmigrationreport.iom.int/msite/wmr-2024-interactive

- United Nations Department of Economic and Social Affairs. (2004). International migration policies: Chapter 1 (Manual VI). United Nations. Retrieve from https://www.un.org/en/development/desa/population/publications/pdf/manuals/migration/manual6/chap1.pdf

- United Nations Population Division. (2022). World population prospects: 2022 revision. United Nations. Retrieve from https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/wpp2022_summary_of_results.pdf

- World Bank. (2024). World Bank Data. Retrieve from https://data.worldbank.org

2.9 Further Readings

- Ram, F., & Pathak, K. B. (1998). Techniques of Demographic Analysis. Mumbai: Himalaya Publishing House.

U.P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N
Demography

Block: 2 Stable Population Theory

Course Design Committee

Course Preparation Committee

Dr. Anupma Singh **Writer**
Department of Statistics
University of Lucknow, Lucknow

Prof. Shruti **Editor**
School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Prof. Shruti **Course Coordinator**
School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N **DEMOGRAPHY**

First Edition: ***July 2024***
ISBN: 978-93-48987-49-5

Printed and Published by Col. Vinay Kumar, Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2024.

Block & Units Introduction

The present SLM on ***Demography*** consists of Ten units with Four blocks. This is the second unit of the SLM.

Block 2 - Stable Population Theory is the second block of this Self learning material.

The Unit 3: Introduction to Stable Population Theory is the first unit of the Self Learning Material which explains the basic concepts and terminologies needed to understand stable Population theory. The unit includes introduction to different type of population structure: Stable, Stationary, Quasi-stable and Non- Stable Populations and their characteristics. It also explains Characteristics and vital rates of Stable Stationary and Quasi- stable Population. Limitations of Stable Population theory.

Unit 4: **Theories and Relationships related to** ***Stable Population Theory*** is the second unit of this Self Learning Material explains the Definitions of intrinsic rate of natural increase. It also explains, basic and vital models developed to explain population growth and projection. It emphasizes on, inter-relationship among intrinsic birth rate and intrinsic death rate, derivation of stable population equation and relationships between birth rate and death rate under condition of stability.

Unit 5: Growth Rates is the third unit of this Self Learning Material explains Computation of intrinsic rate of natural increase under stability condition. It includes derivation of construction of a stable age distribution for a given fertility and mortality schedules, relationship between Net Reproduction Rate, Intrinsic Growth Rate and mean length of generation and the concept of mean interval between generations.

At the end of every block/unit the summary, self-assessment questions and further readings are given.

UNIT- 3 INTRODUCTION TO STABLE POPULATION THEORY

Structure

3.1 Introduction

In the field of demographic models, there are two hypothetical population models widely known as Stationary population and Stable population, both known as model populations. The notion of stable population says that when a particular population having no migration is exposed to same birth and death rates for considerably long period of time, it will eventually develop into a fixed age composition and attains constant of rate of increase. In other words, a population that is governed by a regime of unchanging fertility and mortality schedule for a long time (with no migration) is called stable population. The age structure of this population remains fixed and the size changes with a constant rate of increase r, say, which depends on such fertility and mortality schedules. Whereas, stationarity of a population is achieved when the growth rate of the population becomes zero and population size becomes fixed.

The concept of Stable population was first introduced by The American mathematician and demographer Alfred J. Lotka (1907) which provides a framework for understanding population dynamics under constant birth and death rates. He based his idea on the exponential growth theory in population given by Thomas Malthus (1798) and introduced mathematical methods to analyze population dynamics. His work laid the groundwork for formal Stable population models which was later expanded by the pioneer demographer and mathematician Nathan Keyfitz (1971). He experimented with stable population models to investigate how changes in fertility rates affect population size and equilibrium. His work introduced new perspectives on the relationship between fertility rates and population dynamics. Further J. W. G. Lee (1935), Coale, A. J. (1957), William Brass (1960), Preston S. H. (1974, 2001) etc. are also some of the notable demographers who significantly contributed to the stable population theory.

In stable population model when the value of growth rate r becomes zero, the size of the total population also becomes fixed over time and such a population is known a stationary population. It has the property that the birth rate and death rate will be equal and will have a fixed age composition also. The age structure of a stationary population is the same as that of the life table defining its mortality. It is almost impossible for any population to conforming strictly to the stationary population model yet some earlier human populations can be thought of to have been developed in a stationary manner at certain periods in their history. For instance, before demographic transition, the human populations undergoing high rates of mortality and fertility lived in a primitive demographic regime, the natural growth was almost nil and lead to smaller increments in world population at the beginning of the nineteenth century and can be said to have been stationary state. Even in the present time, some highly developed countries are achieving fertility rates almost equal to the levels of mortality rates.

Up until approximately 1950, the populations across Asia, Africa, and Latin America were relatively stable in terms of their demographic patterns. This means that despite some improvements in health and a general decline in death rates, these changes were modest enough that the established demographic models, which assume a stable population structure, still applied effectively. In other words, the overall population dynamics such as birth and death rates did not deviate dramatically from stability. Demographers frequently used these stable population models to analyse and understand the demographic characteristics of these regions. These models allowed them to investigate various demographic scenarios and trends, despite the gradual changes occurring. Various ways have been developed so that stable population models could be applied or adjusted to explore different demographic phenomena and outcomes.

Before we embark upon the journey of model populations, we need to have a good understanding of population growth models and how they are useful in making projections for future populations.

3.2 Objectives

This unit will help you understand the following:

- The basic Concepts of Stable, Stationary, Quasi-Stable and Non-Stable Populations
- Populations which exhibit the stability and stationarity in their age-sex structure.
- Characteristics of Stationary, Stable, and Quasi-Stable Populations
- Applications and limitations of stable population model.

3.3 Population Growth Model

The Population growth refers to the change in size of a population at some time and space. The term 'growth' is used irrespective of whether the change is positive or negative. Demographic analysis starts with ascertaining the size of the population which may be of a nation as a whole, or of its subnational level or of particular sub-groups which possess certain characteristics. The population estimates were also required in early days for the purpose of ascertaining the supply for the military forces, for assessing number of persons who may come under the category of slave or to be taxed, etc. To plan certain socio-economic and welfare activities, planners and other population scientists need a fairly accurate knowledge about the country's population, specially related to its size, composition, rate of increase, spatial distribution, and other socio-economic, demographic, cultural characteristics of the population. The population estimates and projections are equally important for the governments as well as private agencies. The researchers also require the estimates of population and its projection as an analytical tool to experiment with demographic process for better understanding of the population dynamics. They (population projections) "Permit experiments out of which we obtain causal knowledge; they explain data; they focus research by identifying theoretical and practical issues; they systematize comparative study across space and time; they reveal formal analogies between problems that on their surface are quite different; they even help assemble data" (Keyfitz, 1971).

A number of social scientists and planners worked out short term and large term population progress by the methodological models for assessing the future population growth of a give population under different conditions. However, various models differ from each other depending on different set of assumptions and circumstances in which they are applied as well as the extended accuracy of the available data.

Projection of a population in future is the manifestation of facts and assumptions that we make. While analysis of population data in the past enables us to comprehend the population dynamics; the knowledge of the current data helps us to understand and foresee the future course of the population change. Thus, with the knowledge of present facts on population and assumptions regarding the future course of change projection provide a link between the past, the present and the future. Hence the population projection virtually becomes result of speculations which are based on the trends in demographic indicators established by the past. In demography, population estimation and projection methods are so intertwined that together they form one separate field of scientific endeavor. Population estimation provides quantification of population facts not obtained by other methods like census and surveys. A population projection may be defined as the numerical outcome of a specific set of assumptions regarding future trends in fertility, mortality and migration. On the other hand, a forecast indicates the specific projection that is believed to be most likely to happen. A population forecast indicated a projection in which the assumptions are to yield a realistic picture of the problem failure develop of a population (United Nations, 1958).

Assume a life table where population remains constant over time and having zero rate of growth based on unchanging fertility and mortality conditions with no migration, this hypothetical population creates a stationary population. Stable population is an advanced stage of population in which instead of constant birth rates, an exponential annual growth in births is assumed. Let p(x) be the survival probability to age x (p(x)is a function of age not time, and the births at time t are $B(0)e^{rt}$ assuming births to follow exponential growth, say,

$$B(t) = B(0)e^{rt}$$

Then to get the expected number of individuals between ages x and x+dx, we need to look back in time from x to x+dx years, when the number of births was $B(0)e^{r(t-x)}$, p(x)would be the fraction of these births that survive to time t and the number of persons aged *x* to *x*+*dx* at time *t* can be written as

$$B(0)e^{r(t-x)}p(x)dx$$

The total population at time t can be given by the integral of this quantity and dividing the above quantity by the total so obtained gives the fraction of population aged x to x+dx at time t, say c(x) as below:

$$c(x)dx = \frac{e^{-rx}p(x)dx}{\int_0^w e^{-rx}p(x)dx} = be^{-rx}p(x)dx$$

Where $b = \frac{1}{\int_0^w e^{-rx}p(x)dx}$

Now since c(x) is the fraction of population aged x to x+dx, its total i.e. $\int_0^w c(x)dx$ must be unity (w being the highest age in the population), hence

$$\int_0^w be^{-rx}p(x)dx = 1$$

And so, we get the expression for b as given in the above equation which gives the birth rate in stable population.

For illustration, let us assume a life table for a hypothetical population having life expectancy 5 years. The hypothetical life table values will be:

Table-1.2: Life Table Values of a Hypothetical Cohort

Exact age(x)	**la**	**La/l0**

0	100000	1.000
1	70000	0.700
2	50000	0.500
3	25000	0.250
4	5000	0.050
5	0	0.000

Now, if we assume 1000 births in this population on January 1, 1900, there will be 700 individuals (1000*.6) at exact age on January 1, 1901; 500 individuals at age exact age 2 on January 1, 1902 and so on.

Table-1.2: Population by age from January, 1, 1900 to January, 1, 1906

Age(x)	**01/01/1900**	**01/01/1901**	**01/01/1902**	**01/01/1903**	**01/01/1904**	**01/01/1905**	**01/01/1906**
0	1000	1000* e^{r}	1000* e^{2r}	1000* e^{3r}	1000* e^{4r}	1000* e^{5r}	1000* e^{6r}
1		700	700*e^{r}	700*e^{2r}	700*e^{3r}	700*e^{4r}	700*e^{5r}
2			500	500*e^{r}	500*e^{2r}	500*e^{3r}	500*e^{4r}
3				250	250*e^{r}	250*e^{2r}	250*e^{3r}
4					50	50*e^{r}	50*e^{2r}
5						0	0

From the table, if we compare between the populations for the years 1905 and 1906, we can observe that at every age, the population is advancing by a proportion of er in the next year making the age distributions proportional for each consecutive year populations. So, the age composition which is the proportional age distribution remains constant over time. The population starting from the year 1905 as shown in the table depicts stable population as it is increasing in each year interval by constant birth rate, growth rate and hence constant death rate and having constant age composition. These characteristics will remain unchanged until the population is destabilized by a change in fertility or mortality schedules.

Although the stable population model can be applied to either sex, it is primarily used for females due to the several reasons: 1. Registration authorities typically gather information on the age of the mothers at childbirth more frequently than they do for fathers; 2. Women usually have children within a more specific age range, such as from 15 to 50 years, whereas men can father children over a broader range, from 15 to 80 years; 3. A woman can give birth only at intervals of 1 or 2 years, which limits her to around 20 children at most due to physiological constraints. In contrast, a man can potentially father hundreds of children.

Characteristics of a Population Under the Condition of Stability, Stationarity and Non-Stability:

- **Age Distribution** : In a stable population, the age distribution reaches a point where the proportions of individuals in each age group remain constant over time. This happens when birth rates, death rates, and migration rates are constant and there are no significant changes in the population structure. On the other hand, the stationary population has a fixed population over time since the birth rate and death rate becomes equal. The population in the early age group starts reducing as compared to the older age group because of declining fertility. The population distribution, therefore will be, narrower in the early age, broader in the middle-ages and will narrow down again in the older ages with high expectation of life. Non-stable populations on the contrary, have population distribution based on the type of fertility and mortality condition (usually higher). The populations in the early age groups are larger and then decrease slowly towards the higher age groups with relatively less life expectancy.
- **Growth Rate :** The growth rate of a stable population is very low. Births and deaths balance out, so the population size changes with a constant rate over time. The equality in birth and death rates makes a population fixed which is called as a stationary population with zero rate of growth. This condition is a result of stability in the population for a long time with declining growth. The condition of stationarity is practically not expected to hold for a very long time because if it does hold, the population will eventually start to shrink. Growth rates in non-stable population is generally either very high, high or declining but not constant or zero.
- **Survival Rates :** In a stable population, mortality rates are consistent across different age groups, leading to a more uniform distribution of individuals across age groups. There are fewer dramatic decreases in the number of individuals as age increases compared to a rapidly growing population. It remains similar in the stationary population leading to an increase in ageing of a population. Whereas, in non-stable populations survival rates keep fluctuating with unstable mortality conditions in different populations.

- **Proportions of Age Groups :** In the long term, each age group will make up a roughly constant proportion of the total population in stable population. For example, if the population pyramid is rectangular, this suggests that the proportions of young, middle-aged, and older individuals are stable. In the stationary population, the proportion of populations in the age group 0-5 years decline due to declined fertility and natural mortality. The countries in IV demographic transition show such traits. The proportion of early age groups is higher in non-stable populations with high fertility and mortality or unbalanced fertility and mortality conditions, usually seen in pre transition or early transition phase of demographic transition.
- **Dependency Ratios :** The ratio of dependents (young and old) to the working-age population tends to stabilize. This is because the number of births and the mortality rates in older age groups remain constant. Similar condition is observed in early stationary stage of population but that due to ageing population the dependency ratio starts increasing.
- **Population Pyramid :** A typical population pyramid shows the age sex structure of a population. This age sex structure varies in stable, non-stable and stationary populations. For a stable population, the age structure often forms a shape resembling a column or a rectangular profile in the population pyramid. This means that the number of individuals in each age cohort is relatively equal, with the pyramid looking more like a rectangle rather than the traditional pyramid shape with a broad base tapering to a point.

Figure: 1.1: Population Pyramid of India for year 1950 and 2010

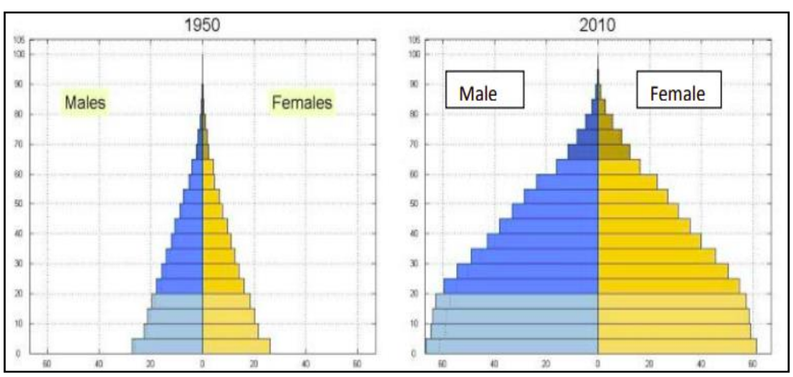


Figure: 1.2: Population Pyramid of India for year 2050 and 2100 (Projected)

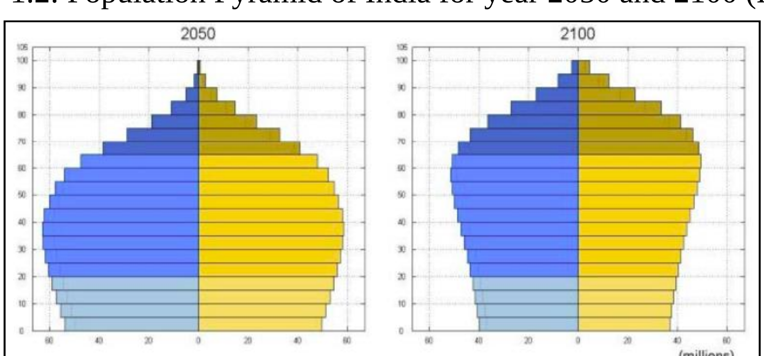


Figure 1.1 gives the picture the age-sex structure India for different years, assuming that by the year 2100 India would have attained stability in population growth. This shows the change in the shape of age pyramid over time showing the change Indian population has gone through over the years and how it might be in the future once it becomes stable.

On comparison of the pyramids for year 1950 and 2010, we can observe the changing life expectancy, higher births and higher deaths in year 1950 as compared to year 2010 for both males and females. As we more further, and we see the 2050 projected population given in figure 1.2, we can see that it consists of large proportions of children and persons in reproductive age group and making it potential for still increasing population because even the fertility and mortality rate have declined considerable over time, the concentrated population in the reproductive ages will still produce the offspring that would lead to growing population, the phenomenon known as population momentum, will be studied in next unit of the chapter. And in the last figure i.e. in year 2100 where the population is assumed to have attained stable state, we see that the shape of the pyramid has totally changed and has almost become cylindrical. This shows less fertility, less mortality and lesser population in the working group and higher concentration of aged population leading to ageing society and thus leading to challenges of higher dependency ration, which is a challenge to stable populations.

There are many populations across the world which are going through the phase of stability and stationarity and can be explained simply by looking at their population pyramids. It includes a lot of European countries like Italy, France, Spain etc. and the nations like Germany, the United States etc. The population pyramid of some such populations is given below:

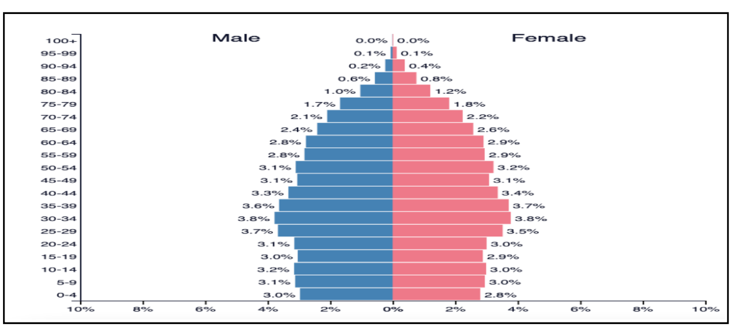


Figure1.3: Population Pyramid of Australia (2024)

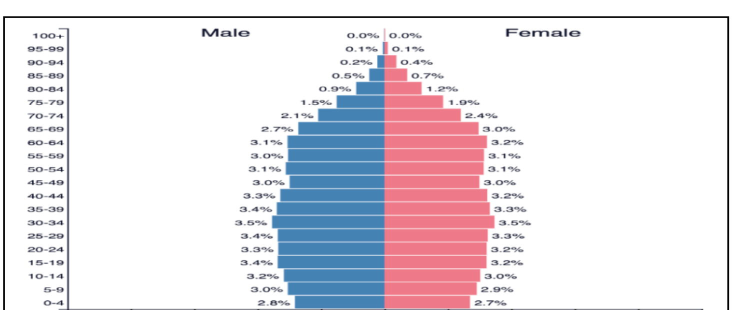


Figure1.4: Population Pyramid of the United States of America (2024)

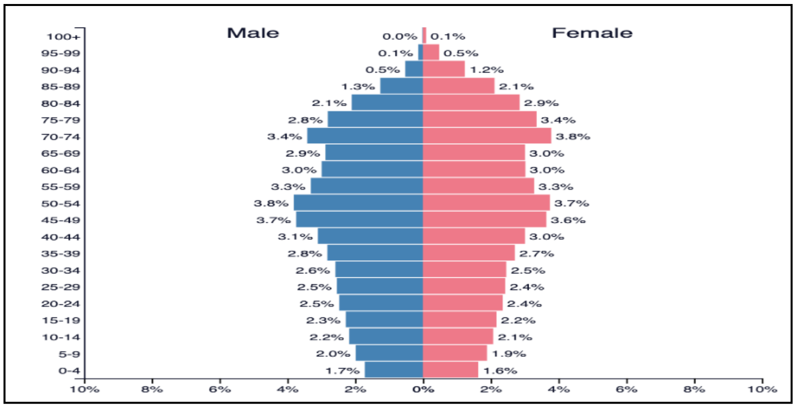


Figure1.5: Population Pyramid of Japan (2024)

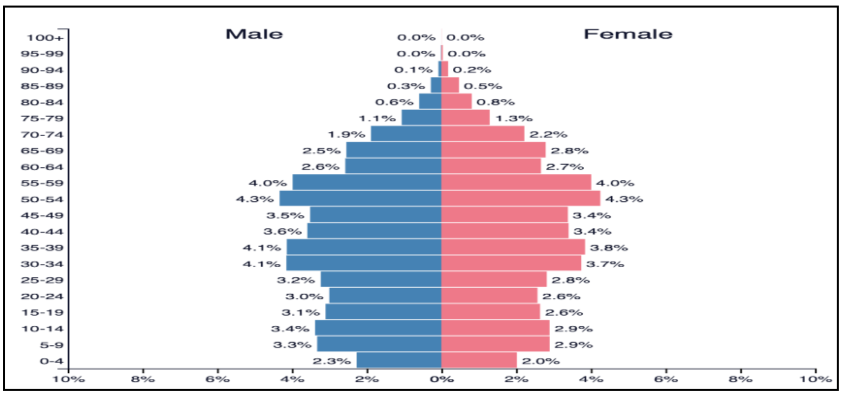


Figure1.6: Population Pyramid of China (2024)

Figure 1.3 and 1.4 are examples of stable population whereas figure 1.5 and 1.6 show the age-sex structure of a stationary population. The population pyramids of above nations depict that they all are developed populations. The fertility and mortality rates in the developed countries are very low. This, obviously contributes lesser number of individuals in the higher age groups. Therefore, we see that the populations which are stable are more like the population pyramid of Australia or the United States of America, shown in the in fig 1.3 and 1.4, i.e., more rectangular in the early and middle ages and slowly declining population with increasing age. Stable population have even distribution of population over earlier age groups and further in the higher age groups with an increasing expectation of life. If we look at such populations in a social aspect, then we can say that these populations are economically more prosperous with higher standard of living. Such populations also have better health and education facilities, which is shown by decreased fertility as well as mortality over the years. The proportion of aging population in both stable and stationary population is higher as compared to the non-stable populations. Since the fertility and mortality are very low the aging population increases in stable population and goes even more higher in stationary population. The stationary populations are more of shrinking populations if we look closely in fig. 1.4 and fig. 1.5. Since the fertility has declined substantially and mortality is also low the early age population is very small. The size of the population, therefore, is fixed. Now due to natural mortality the population is subjected to reduce and hence we get a shrinking population. In reality, if we observe this stage does not go for a longer period because by theory of regulation, we tend to increase the population the moment it starts decreasing. The quasi-stable populations are in a sense result of the situations where a population is stable but it goes off and on to non-stability.

The populations which are **not stable** are considered to non-stable populations. The example of some such populations is given below:

Fig: 1.7 Population Pyramid of Afghanistan (2024)

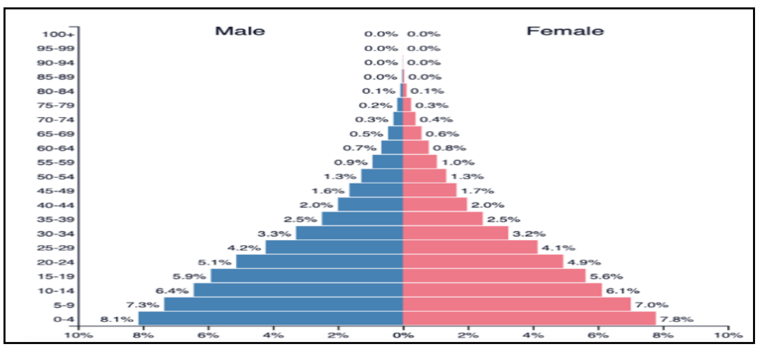
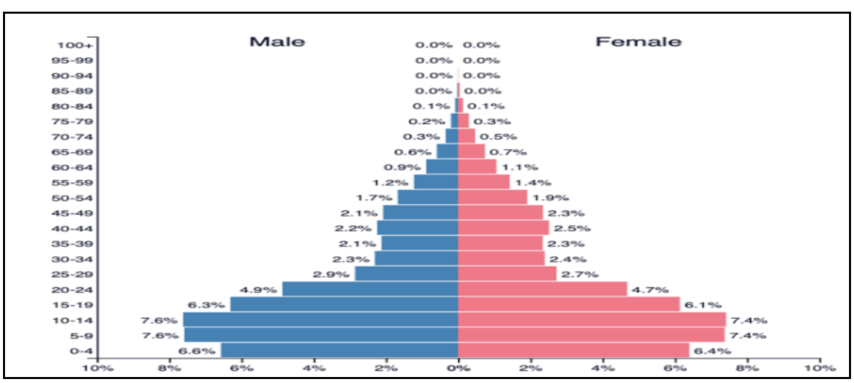


Fig: 1.8 Population Pyramid of South Sudan (2024)



3.4 Definitions

The following sections explains the definitions of Stable, Stationary, Quasi-stable and non-stable populations along with the underlying assumptions.

3.4.1 Stable Population

A population closed to migration having fixed rate of growth and constant age structure over time is known as stable population. In other words, a population closed to migration that is governed by a regime of unchanging fertility and mortality schedule for a long time is called a stable population. The example of a stable population is shown in fig. 1.3 and 1.4.

The fundamental assumptions of a stable population are:

1. The size of the population changes with constant rate of increase (or decrease).
2. The birth rates and death rates remain invariant over time but not equal. And,
3. The population is closed to migration i.e. net migration rates are zero across all ages.

3.4.2 Stationary Population

A population closed to migration with zero rate of growth and a fixed size over time is termed as a stationary population. It is a special case of stable population where growth rate becomes zero and fertility and mortality rates are constant over time but are equal. Since the size of the population is fixed over time, such population tends to shrink in the future under same regime of fertility and mortality. The example of stationary population is shown in fig. 1.5 and 1.6.

A stationary population is attained under the following demographic conditions:

1. The growth rate of population is zero
2. The age structure of the population remains constant
3. Crude birth rates and crude death rates are equal and
4. The population under study is closed to migration.

Thus, almost none of the population variables change over time in the stationary model.

3.4.3 Quasi-Stable (Semi-Stable) Population

For any population to be stable, the condition of constancy of its birth and death rates over time is assumed. Quasi-stable or semi-stable population is a type of population structure that exhibits characteristics of stability but does not fully conform to the conditions of a truly stable population. Unlike a stable population, which has a constant age distribution over time, a quasi-stable population shows some degree of change but maintains certain stable aspects over a period.

3.4.4 Non-Stable Population

The non-stable populations refer to demographic populations that are not in a state of equilibrium or stability. Unlike stable populations, where age distributions remain constant over time and the population growth rate is zero, non-stable populations exhibit changes in their age structures and growth rates.

3.5 Characteristics of Stable, Stationary and Quasi-Stable Population

The following are:

3.5.1 Stable Population Characteristics

Below are given some of the main characteristics of the stable population:

(1) The Fundamental Equation Characterizing Stable Population: given by Alfred J. Lotka is given as-

$$\int_{\alpha}^{\beta} e^{-ra} m(a) p(a) da = 1 \qquad (3.1)$$

Where,

- α and β represent the lower and upper limit of reproduction ages of female of the population under consideration,
- $p(a)$ represents the proportion surviving from birth to age a,
- $m(a)$ represents the annual rate of bearing female children at age a, and
- r represents the natural rate of growth of this population.

As we know, a stable population has an exponential birth series say $B(t) = Be^{rt}$. The value of r that satisfies the above equation for given schedules of $m(a)$ and $p(a)$ is known as intrinsic rate of growth.

(2) Birth Rate in the Stable Population: The number of individuals aged a at time t is determined by the number of births that occurred $t - a$ years earlier, multiplied by the probability of surviving from birth to age a. This can be expressed as

$$N(a, t) = B(t - a) \cdot p(a)$$

Substituting this formula into the previous one results in:

$$N(a,t) = Be^{-r(t-a)} \cdot p(a)$$

$$= B\, e^{rt}\, e^{-ra}\, p(a)$$

$$N(a,t) = B(t)e^{-ra}\, p(a) \tag{3.2}$$

Integrating both sides from 0 to w, w being the highest age in the population, we get

$$\int_0^w N(a,t)da = B(t)\int_0^w e^{-ra}\, p(a)da$$

or

$$\frac{B(t)}{\int_0^w N(a,t)da} = \frac{1}{\int_0^w e^{-ra}\, p(a)da}$$

$$\frac{B(t)}{N(t)} = b(t) = \frac{1}{\int_0^w e^{-ra}\, p(a)da} = b \tag{3.3}$$

This gives the crude birth rate of a stable population, b which remains constant over time.

(3) Age Structure of the Stable Population: Referring back to equation (3.2), if we divide both sides by $N(t)$, the total population size at time t, we can obtain an expression for the proportion of the population in each age group, denoted as $c(a,t)$ which is given as

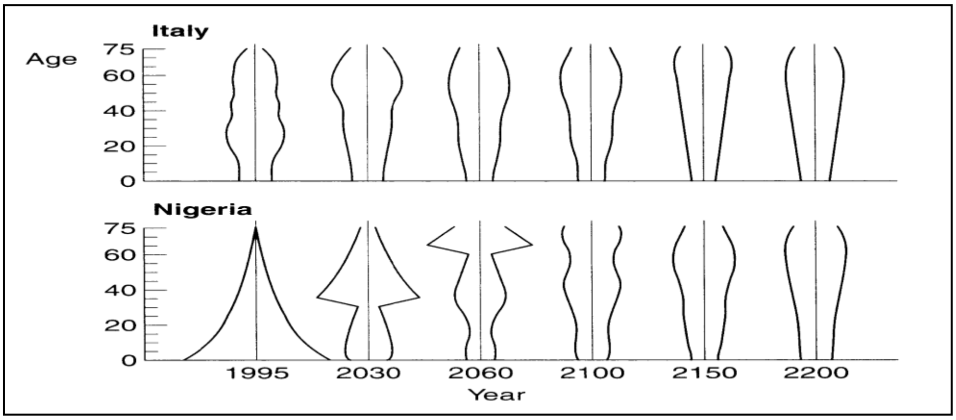
$$c(a,t) = \frac{N(a,t)}{N(t)} = \frac{B(t)e^{-ra}\, p(a)}{N(t)}$$

$$= be^{-ra}p(a)da = c(a) \tag{3.4}$$

Which shows that the proportional age distribution is also fixed over time.

(4) A key characteristic of the stable population model is its ability to "forget" its historical age distribution when exposed to constant rates of change over time. This characteristic is referred to as ergodicity. Strong ergodicity is the long-run behaviour of a population subjected to constant fertility and mortality rates. It means that if two different age distributions are subjected to the same set of constant fertility and mortality rates, they will eventually converge to the same age distribution, regardless of their initial differences. For illustration, Person et. al. (2001) applied the same set of fertility and mortality schedules to Italy and Nigeria, the two very countries inheriting different demographic features and showed that their age structures and hence population pyramid eventually becomes similar irrespective of their sizes as shown in the figure-1.3

Figure 1.3 Relative age distributions of Italy and Nigeria, both projected with the 1995 vital rate of Italy



Data Source: *United Nations, 1995,1996; Picture Source: Samuel H. Preston, Patrick Heuveline, Michel Guillot (2001). Measuring and Modelling Population Processes.*

The time any population will take to stabilize depends on the difference between age distributions of the population on which the constant set of fertility and mortality schedules are imposed i.e. the time when age differences in the population become negligible or say zero over time. The larger the differences, the greater number of years it will take to stabilize. For developing and under developed countries, it can take centuries for their populations to stabilize. This property helps in predicting the time for a population to achieve stability thus making one able to compare populations of different countries in reference to become stable.

3.5.2 Stationary Population Characteristics

(1) Growth Rate: If the value of r becomes zero in a stable population, then the size of the total population also becomes fixed over time and such a population is known a stationary population. It has the property that the birth rate and death rate will equal; of course it will have a fixed age composition also.

(2) Birth and Death Rate: The birth rate in this case is given by

$$b = \frac{1}{\int_0^w p(a)\, d(a)}$$

$$= \frac{1}{\int_o^w \frac{l_{(a)}}{l_0}\, da} = \frac{l_0}{\int_o^w l_{(a)}\, da}$$

$$= \frac{l_0}{\int_0^1 l_{(a)}\, da + \int_1^2 l_{(a)}\, da + \ldots\ldots + \int_{w-1}^{w} l_{(a)}\, da}$$

$$= \frac{\mathrm{l}_0}{\mathrm{L}_0 + \mathrm{L}_1 + \mathrm{l}_2 + \ldots + \mathrm{L}_{10}} = \frac{\mathrm{l}_0}{\mathrm{T}_0} = \frac{1}{\mathrm{e}_0}$$

Thus, birth rate in stationary population is equal to the inverse of the expectation life at birth.

Since the rate of increase is zero, hence the death rate will also be equal to $1/e_0$.

(3) Age Structure in Stationary Population: The age structure in stationary population is given by

$$c(a) = b\, p\, (a) \qquad (\because e^{-ra_1} = 1 \quad for \quad r = 0)$$

$$= \frac{l_0}{T_0} x \frac{l_a}{l_0} = \frac{l_a}{T_0}$$

Hence the proportion between a to a +1 = $\frac{\int_a^{a+1} l(x)\, dx}{T_0} = \frac{L_a}{T_0}$

Thus, we see that L_x column of the life table represents the age composition of the stationary population.

3.5.3 Quasi-Stable Population Characteristics

A population can be unstable in several ways; one such state is called quasi-stability in which birth rates remain constant while death rates decline uniformly. Thus, a population is said to be quasi-stable if it experiences fixed birth rates and constantly declining death rates over time.

Let us consider the effect of a constant drop in mortality at every age say from $\mu(a)$ to

$$\bar{\mu}(a) = \mu(a) - k.$$

Since for a given r, we have

$$\int_\alpha^\beta e^{-ra}\, p(a)\, m(a) da = 1, \qquad (3.5)$$

Let us calculate the new survival function under the changed mortality schedule as

$$\bar{p}(a) = e^{-\int_0^a \bar{\mu}(x)dx} = e^{-\int_0^a \{\bar{\mu}(x-k)\}dx}$$

or $\bar{p}(a) = e^{ka} p(a)$

Let the new rate of increase be $\bar{r}$, hence $\bar{r}$ satisfies the equation

$$\int_{\alpha}^{\beta} e^{-\bar{r}a} p(a)\, m(a)\, da = 1$$

$$\text{i.e.} \int_{\alpha}^{\beta} e^{-\bar{r}a}\; e^{ka}\; p(a)\, m(a)\, da = 1$$

$$or \int_{\alpha}^{\beta} e^{-(\bar{r}-k)a} p(a)\, m(a)\, da = 1 \qquad (3.6)$$

Comparing equation (1.5) and (1.6) we have

$(\bar{r} - k) = r$ which implies that

$$\bar{r} = r + k \qquad (3.7)$$

Now the new age structure is given as

$$\bar{c}(\bar{a}) = \frac{e^{-\bar{r}\bar{a}} . \bar{p}(a)}{\int_0^w e^{-\bar{r}\bar{a}}\; \bar{p}(a)\, da} = \frac{e^{-(r+k)a}\quad e^{ka}\, p(a)}{\int_0^w e^{-(r+k)a}\; e^{ka}\, p(a)\, da}$$

$$= \frac{e^{-ra}\quad p(a)}{\int_0^w e^{-ra}\; p(a)\, da} \quad \text{since } \bar{r} = r + k \text{ and } \bar{p}(a) = e^{ka}\, p(a)$$

$$= c(a). \qquad (3.8)$$

Thus, we see that the population growth rate increased by the same amount k but there is no change in the age structure of the population under uniform decrease in mortality at all ages. However, a uniform decline in mortality at all ages is not realistic. Infant and child mortality, adult and old age mortality all behave differently. But practically it is seen that when mortality declines, it has little differential in mortality by age.

3.6 Applications of Stable Population Theory

Stable population theory has proven to be of great importance in mathematical demography for many reasons:

(1) Stable population theory helps in forecasting future population sizes and structures by assuming that current age-specific birth and death rates will remain constant. This is valuable for planning and policy -making in areas such as urban development and public services. It allows demographers to study different circumstances based on varying assumptions about birth and death rates, providing a range of possible future outcomes. e.g. Keyfitz have applied stable population theory to predict the impact of different population control policies on future population growth. Additionally, the theory can be readily adapted to include migration factors.

(2) By analysing the stable age distribution, researchers can predict how the proportion of different age groups will evolve over time, which is crucial for studying demographic transitions and aging populations.

(3) The theory provides insights into the potential growth or decline of a population based on its intrinsic growth rate which helps in understanding how populations might change in the absence of migration.

(4) Also, stable population theory is crucial for using indirect methods to estimate demographic parameters in countries with incomplete vital registration systems. For instance, mortality rates or intrinsic growth rates can be estimated under the assumption that the population has achieved stability, even when direct data is unavailable.

(5) By projecting the future age distribution of a population, planners can estimate the demand for healthcare services, including those for an aging population. This helps in allocating resources and planning healthcare facilities and services accordingly. The theory aids in evaluating the sustainability of pension systems and social security programs by predicting the future ratio of working-age individuals to retirees. This helps in designing policies to address potential financial imbalances.

3.7 Limitations of Stable Population Theory

Stable population theory has proven to be a very useful demography tool to study and predict demographic phenomenon. While it provides useful insights, there are several issues associated with this theory: The theory assumes no net migration, constant birth and death rates over time, which rarely reflects real-world conditions where these rates often fluctuate due to socio-economic, and environmental factors. Actual populations are often dynamic and don't fit the stable model due to migration. Though these strict assumptions have been relaxed somewhat by various demographers say for example making some changes in the fertility rate only and then assessing the impact of changes in fertility schedules while keeping all other assumptions hold true in the stable population. Similarly relaxing the assumption of fixed mortality rate only or incorporating the migration process while keeping births and deaths and age structure constant. Also, for developing and under developed countries, the usefulness of stable population concept become questionable as the assumptions are hardly applicable on these populations. Recognizing these limitations helps demographers and policymakers use stable population theory more effectively and interpret its findings within the context of real-world complexities.

3.8 Summary

It's understandable why the concept of a stable population is so popular in demographic research. At any given time, mortality and fertility rates can serve as effective indicators of health and reproductive behaviour. A stable population provides a clearer assessment of these aspects. In comparison, although current mortality rates and crude birth rates are useful for understanding health conditions, they present a distorted view of reproductive behaviour because they are influenced by the current age distribution.

3.9 Self-Assessment Exercise

1. Define stable, stationary and quasi-stable population. Explain various characteristics of all three using examples.
2. Derive the expression for birth rate and age composition in stable and stationary populations.
3. Differentiate between stable and stationary population.
4. What are the applications and limitations of stable population theory? Explain.

3.10 References

- Ram, F., & Pathak, K. B. (1998). Techniques of Demographic Analysis. Mumbai: Himalaya Publishing House.
- Samuel H. Preston, Patrick Heuveline, Michel Guillot (2001). Measuring and Modeling Population Processes.

3.11 Further Readings

- Nathan Keyfitz, Hal Caswell (1985). Applied Mathematical Demography.
- The concept of a stable population: Application to the study of populations of countries with incomplete data (1968). United Nations New York.

UNIT - 4: THEORIES AND RELATIONSHIPS RELATED TO STABLE POPULATION THEORY

Structure

4.1 Introduction

As discussed in the earlier chapter, a stable population is a population with a constant growth rate under a fixed regime of fertility and mortality with zero migration assumed. This stage of stability is achieved after long series of transformation called as Demographic transition the concept of which was given by Notestein in the year 1929. Whereas, the concept and derivation of stability equation was given by Lotka (1907). Other than stability, a population goes through different other phases; un-stability, quasi-stability and stationarity. Under the assumptions of stability, the natural growth rate, birth rate and death rate of a population maintain a relationship which can be expressed in a mathematical form.

4.2 Objectives

- To give a general understanding of population growth rate under population growth models.
- To give definitions of intrinsic rates of natural increase, intrinsic birth rate and intrinsic death rate.
- To study the relationships among intrinsic rates of natural increase, intrinsic birth rate and intrinsic death rate.
- derivation of Lotka's formula of fundamental relationship in stable population.
- To Recognise how different fertility and mortality schedules determine the age structure and growth rate of stable populations.

4.3 Population Growth Rates and Projection Models

Before moving towards the notion of intrinsic growth rate of a stable population, it is important to have a general understanding of growth rate or say rate of increase/decrease of a population and other mathematical models to understand growth of a population in a longer period of time. In demography, the simplest measure of population growth is the rate of natural increase. The natural increase is the difference between the number of births and deaths during a given period. It is simply defined as-

Natural Increase = Number of Births- Number of Deaths

Intrinsic Growth Rate = Birth Rate - Death Rate

It should be noted that the third component of population change, namely, migration is not included in the definition of this rate and so it is only a measure of change.

Other important definitions and population projection models are explained in the subsections given below.

4.3.1 Rate of Natural Increase

The rate of increase in population can be measured by change in population size over time. If P_0 be the size of a base population at time 0 and P_t be its size at time t, then the growth in population during the period t is $P_t - P_0$ and the rate of increase is

$$r = \frac{P_t - P_0}{P_0} * K$$

Usually, K is taken as 1000. This gives change per 1000 over time t.

4.3.2 Arithmetic Growth Model

If t is measured in years, the annual rate of increase is

$$r_A = \frac{1}{t}\left(\frac{P_t - P_0}{P_0}\right) * K$$

This measure is called the annual arithmetic growth given as

$$P_t = P_0\,(1 + rt) \qquad (4.1)$$

But it is more appropriate to use 'rate of increase' instead of growth as this rate, in strict sense, does not reflect population growth as it assumes same annual increase in size.

4.3.3 Geometric Growth

In a population every addition has the potential to change the size of the population and hence a geometric pattern of increase is more appropriate. Assuming that the population is compounded annually with growth rate r, the growth equation during time t can be written as

$$P_t = P_0 (1 + r_G)^t \qquad (4.2)$$

r_G : Geometric growth rate

The above expression gives the Geometric law of population growth.

4.3.4 Exponential Growth

In the geometric growth of population, annual compounding is assumed. If we assume that compounding is not occurring annually but n times in a year, then we can write

$$P_t = P_0 (1 + r/n)^{nt} \qquad (4.3)$$

If the year is divided into n intervals and by making n infinitely large, then (3) becomes

$$P_t = P_0 e^{rt} \qquad (4.4)$$

As $\lim_{n\to\infty}(1 + r/n)^n = e^r$

This is known as exponential form of population growth and is continuous version of geometric growth in equation (4.2).

This growth rate can also be derived considering the instantaneous rate of change as

$$r = \frac{1}{P_t}\frac{d}{dt}P_t = \frac{d}{dt}\log P_t$$

On integration, we get

$$\log P_t = rt + \log C$$

At $t = 0, C = P_0$

Therefore,

$$P_t = P_0 e^{rt} \tag{4.5}$$

Below are given the examples to compute the growth rate for these growth models.

Example 1: If population of India as of April, 1 for 1971 and as of March 1, 1981 is 548.0 and 685.0 million respectively. Estimate the population for 1978 and project the population for 1991 using three mathematical methods discussed above.

Solution: It may be noted that the census of India of 1971 and 1981 are not exactly 10 years apart. So, for estimation and projection of the population some adjustment has to be made. The intercensal period during 1971-81 is

$$9years + \frac{11months}{12months} == 9.9167years$$

(a) Using Arithmetic Growth Model:

$$P_{81} = P_{71}\ (1 + 9.9167*r)$$

or $$685 = 548\ (1 + 9.9167*r)$$

Solving this equation, we get

$$r = \frac{1}{9.9167}\left\lfloor\left(\frac{685}{548}\right) - 1\right\rfloor = 0.02521$$

So, estimate of population for 1978 (P_{78}) would be

$$P_{78} = P_{71}\ (1 + 6.9167 * 0.02521)$$

$$= 548 * (1+6.9167 * 0.02521)$$

$$= 643.55 \text{ million.}$$

For projection of population to 1991, we should use 1971 as base to maintain linear growth rate as constant between 1971-91. Nevertheless, we can assume that population would linearly grow after

1981 with observed growth rate ‘r’

$$P_{91} = P_{81}(1 + 10r) = 685\ (1+10 * 0.02521) = 858.69 \text{ million.}$$

(b) Using Geometric Growth Model:

$$P_{81} = P_{71}(1 + r)^{9.9167}$$

On substitution, we get

$$685 = 845\ (1+r)^{.99167}$$

Or
$$r = \left(\frac{685}{548}\right)^{\frac{1}{9.9167}} - 1 = 0.02276$$

The estimate of population for 1978 (P_{78}) would be

$$P_{78} = P_{71}(1 + 0.2276)^{6.9167}$$

$$\hat{P}_{78} = 548\ (1.02276)^{6.9167} = 640.29\ million$$

And as calculated in previous method, we can estimate

$$\hat{P}_{91} = P_{81}(1 + 0.02276)^{10} = 685(1.02276)^{10} = 857.86\ million.$$

(c) Using Exponential Growth Model:

$$P_{81} = P_{71}\, e^{9.9167*r}$$

$$685 = 548\ e^{9.9167*r}$$

$$r = \frac{1}{10} * \log_e\left(\frac{685}{548}\right) = 0.0225$$

so, P_{78} $= 548\ e^{6.9167*\ 0.0225}$ = 640.29 million

and P_{91} $= 685\ e^{10*0.0225}$ = 857.86 million.

These basic models of population growth discussed so far depict indefinite population growth with no upper limit to P_t which is not true in general as it has been observed that the velocity of change slows down after a certain period and hence these models are not suitable to describe real time population growth in nature. Below are given some more realistic population growth models to portray the behaviour of real population growth.

4.3.5 Gompertz Model

The Gompertz Model was given by Benjamin Gompertz in 1825 and was initially used to study human mortality, but it has also been useful to study and predict population growth processes. The Gompertz Law of Population Growth is a mathematical model used to describe population growth that follows a sigmoidal (S-shaped) curve, where population grows exponentially first, slows down over time, and eventually reaches a plateau or asymptote.

The growth model can be written as

$$P_t = K\ B^{st}$$

Taking logarithms both sides, we get

$$\log P_t = \log K + st \log B \qquad (4.6)$$

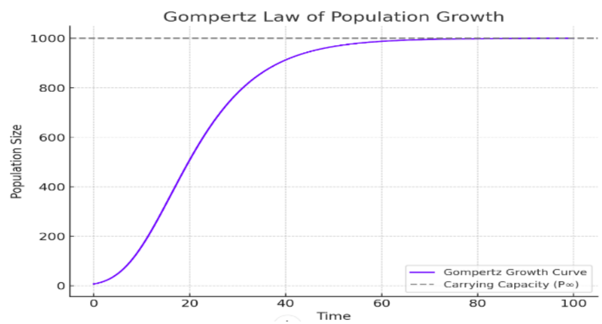
Equation (4.6) becomes a logarithmic model and its fitting be done as by simply writing the model in a linear form as

$$\text{Log}\, P_t = a + bt$$

Where, $a = \log K$ and $b = s \log B$

This is now a linear equation in terms of t (time), which can be fitted using linear regression.

The graph below illustrates the Gompertz Law of Population Growth, where the population size grows rapidly at first, then gradually slows down as it approaches the carrying capacity as indicated by the dashed line. The curve follows the typical sigmoidal shape, with the population growth rate decreasing exponentially over time.



4.3.6 Logistic Model

The logistic curve of population growth has usually been found to give good fit for those populations which grow under conditions of limited food and space but without any biological restriction on reproduction. But these conditions are not applicable to human species and hence the curve is not a suitable mathematical model of human population growth. Yet it has been found that the curve can be used to describe certain populations that are subject to sudden spurt of growth from a fairly steady level and then decline slowly to reach a low growth rate again. Also, the curve can be fitted only for data of a limited interval of time curve is not accepted as the best it has adequate uses for demographic analysis. It is also called Pearl-Reed curve, after the author who independently derived it in 1920. In fact, the curve was, in fact discovered earlier in 1838 by Verhulst who suggested that the population growth could be "*S*" type and he named it logistic. The logistic curve possesses the characteristics of proceeding from lower limit to a determinate upper limit. It may be remembered that the curve may fit observed data with greater accuracy, yet fails to produce the population for future with same accuracy if the population trends undergo drastic change. The logistic curve is defined as

$$\frac{d}{dt}P_t = rP_t\left(1 - \frac{P_t}{K}\right) \tag{4.7}$$

Where K is maximum population size and gives the upper asymptote. This equation can also be written as

$$r = \frac{K}{P_t(K - P_t)} \times \frac{d}{dt} P_t$$

or

$$r = \left(\frac{1}{P_t} + \frac{1}{K - P_t}\right) \frac{d}{dt} P_t$$

Which on integration gives

$$log P_t - \log(K - P_t) = rt + c$$

Where c is constant of integration.

On rearranging the above expression, we can write

$$\frac{P_t}{K - P_t} = e^{rt+c}$$

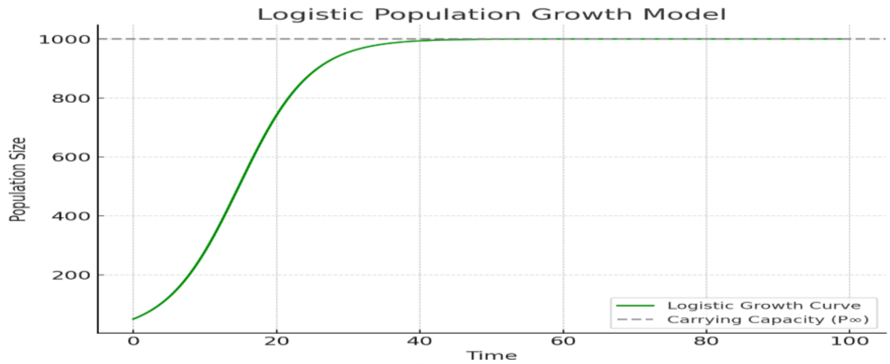
or

$$P_t = \frac{K}{1 + e^{-(rt+c)}}$$

Which can also be written as

$$P_t = \frac{K}{1 + e^{a+bt}}$$

The Logistic curve is useful for projection over a longer period than the simple geometric law particularly if the past series have reached the point of inflexion. The difficulty however, is to select the countries and the period of observations (Shryock and Siegel, 1980). Like all mathematical curves, this curve is also mechanistic. Therefore, it is generally advised not to use this curve for long run projections. The logistic curve is not useful to project population with negative growth rate.



Fitting of the Logistic Curve: In Logistic curve there are three parameters K, a and b which have to be estimated. For estimating three parameters, we need population figures at least from three censuses.

The Logistic curve can also be written as

$$\frac{1}{P_t} = \frac{1 + e^{a+bt}}{K} = \frac{1}{K} + \frac{e^{a+bt}}{K}$$

or

$$\frac{1}{P_t} = \frac{1}{K} + \frac{e^a}{K}(e^b)^t$$

i.e. $$Y_t = A + BC^t \qquad (4.8)$$

where

$$Y_t = \frac{1}{P_t}, \quad A = \frac{1}{K}, \quad B = \frac{e^a}{K}, \quad C = e^b$$

Equation (4.8) is known as modified exponential curve It means that any time series data would follow logistic law if their reciprocal follows a modified exponential law.

To fit this mode, one can use the three selected forms method freely Pearl and Reed. These points are taken at equivalent on the time scale and then are selected in such way that the whole age of observations is more or less evenly lowered.

Let these points be at 0, n, and 2n (in the interval between two points) at which populations be P_0, P_n and P_{2n} respectively. Since logistic curve if agrees, must pass through there points, and as such

$$\frac{1}{P_0} = \frac{1+e^a}{K} \tag{4.9}$$

$$\frac{1}{P_n} = \frac{1+e^{a+bn}}{K} \tag{4.10}$$

$$\frac{1}{P_{2n}} = \frac{1+e^{a+2bn}}{K} \tag{4.11}$$

Let us write

$$d_1 = \frac{1}{P_0} - \frac{1}{P_n} = \frac{e^a}{K}(1-e^{bn}) \tag{4.12}$$

And

$$d_2 = \frac{1}{P_n} - \frac{1}{P_{2n}} = \frac{e^{a+bn}}{K}(1-e^{bn}) \tag{4.13}$$

On division of (4.12) to (4.13), we get

$$\frac{d_1}{d_2} = e^{-bn}$$

Which give us

$$b = \frac{1}{n}[\log d_2 - \log d_1] \tag{4.14}$$

On further simplification, we can write

$$1 - \frac{d_2}{d_1} = \frac{K d_1}{e^a}$$

$$\frac{{d_1}^2}{d_1 - d_2} = \frac{e^a}{K}$$

Which can be written as

$$\frac{1}{K} = \frac{1}{P_0} - \frac{{d_1}^2}{d_1 - d_2} \tag{4.15}$$

From equation (4.9), we get

$$\frac{K}{P_0} = 1 + e^a$$

Which gives

$$a = \log\left[\frac{K}{P_0} - 1\right]$$

Thus, for a given population data, we can estimate the parameters as above in the equation and can predict the population.

4.3.7 Doubling Time

Usually, people show their interest in knowing about the time that will be taken for a population to double itself for a given growth rate r. The time that a population takes to double in size at a constant growth rate is known as its Doubling time. Thus, we need to find t such that

$$P_t = 2P_0$$

To derive this, assuming exponential growth, the doubling time can be calculated by substituting

$P_t = 2P_0$ in equation (2.5), we get

$$2P_0 = P_0 e^{rt}$$

and the doubling time needed for a population with growth rate r is

$$t = \frac{log_e 2}{r}$$

Which gives us

$$t = \frac{0.693}{r}$$

If r is expressed as percentage, we can write t=70/100r approximately which implies that with a growth rate of r=.01, the population will double in 70 years and with r=.02, it will double in 35 years. Although in case of decreasing population, this t can provide the timing in halving the population instead of doubling.

So far, we tried to understand the rate of growth of populations in some of the mathematical population growth models which brings us next to the Malthusian population models which provides the basis for the stable model populations. The previous chapter gives some understanding of stable population model, the conditions for a population to become stable and its characteristics in a detailed manner. This chapter provides a detailed exercise to determine the mathematical expression for the basis of stable population model i.e. the characteristic equation given by Lotka

$$\int_{\alpha}^{\beta} e^{-ra}\, p(a)m(a)da = 1$$

the parameters involved in this equation, and also to find the root of the equation that gives the intrinsic growth rate.

4.4 INTRINSIC RATES AND THEIR INTERRELATIONSHIP

These are as follows;

4.4.1 Intrinsic Birth Rate

The birth rate ‘b’ of a stable population is known as intrinsic birth rate. In other words, the birth rate of a population closed to migration which experiences constant vital rates and increases or decreases with constant rate of growth is termed as intrinsic birth rate.

In notation, it is expressed as

$$b = \frac{1}{\int_0^w e^{-ra}\, p(a)da}$$

4.4.2 Intrinsic Death Rate

The death rate ‘d’ of a stable population is known as Intrinsic death rate. In other words, the death rate of a population closed to migration undergoing constant birth rates and rate of growth is termed as intrinsic death rate.

In notation, it can be expressed as

$$d = \int_0^w c(a)\, \mu(a) da$$

Where $\mu(a)$ is the force of mortality as stated in the previous chapters.

4.4.3 Intrinsic Growth Rate

The growth rate of a stable population is termed as intrinsic growth rate denoted as ‘r’. It is the value of ‘r’ that satisfies the characteristic equation of stable population and is the natural growth rate that is intrinsic to the fertility and survival schedules that produced it. If the intrinsic birth rate is higher than the intrinsic death rate, the population will increase; if it is lower, the population will decrease. When intrinsic birth and death rates are equal, the population stationary in size.

4.4.4 Relation among Intrinsic Rates

In a stable population, the intrinsic growth rate r, the intrinsic birth rate b and the intrinsic death rate d are related by the equation below:

$$r=b-d$$

Where r satisfies the fundamental equation

$$\int_\alpha^\beta e^{-ra}\, p(a)m(a)da = 1$$

and b is expressed as

$$b = \frac{1}{\int_0^w e^{-ra}\, p(a)da}$$

And d can be derided using these equations.

We can also express the death rate as

$$d = \int_0^w c(a)\, \mu(a) da$$

Where $\mu(a)$ is the force of mortality as stated in the previous chapters.

The above expression can be written as

$$d = \int_0^w be^{-ra} p(a)\, \mu(a) da$$

$$= b \int_0^w e^{-ra} p(a)\, \mu(a) da$$

Or we can write on substituting the value of b as

$$d = \frac{\int_0^w e^{-ra} p(a)\, \mu(a) da}{\int_0^w e^{-ra}\, p(a) da}$$

Since d=b-r, we can write

$$\frac{\int_0^w e^{-ra} p(a)\, \mu(a) da}{\int_0^w e^{-ra}\, p(a) da} = \frac{1}{\int_0^w e^{-ra}\, p(a) da} - r$$

And on solving this equation, we get

$$r = \frac{1 - \int_0^w e^{-ra} p(a)\, \mu(a) da}{\int_0^w e^{-ra}\, p(a) da}$$

4.5 Derivation of Lotka's Fundamental Equation

Lotka based his work on three key assumptions:

1) Age-specific fertility rates remain constant over time,

2) Age-specific mortality rates remain constant over time, and

3) Net migration rates are zero across all ages.

It's important to note that Lotka focused on a single-sex population i.e. female population only. He analysed the birth sequence in this population starting from time 0, when these conditions were first applied:

4.5.1 Sequence of Births

Suppose that a constant schedule of fertility and mortality is imposed at time t = 0, when there exists some arbitrary initial age distribution. Subsequent female births will be the sum of two components, namely

1) births occurring to female present in the initial population, and

2) births to females itself borne during the regime of constant fertility and mortality schedules.

Thus if $B(t)$ represents the number of births at time t then we have

$$B(t) = F(t) + \int_0^t B(t-a)p(a)m(a)da \qquad (4.16)$$

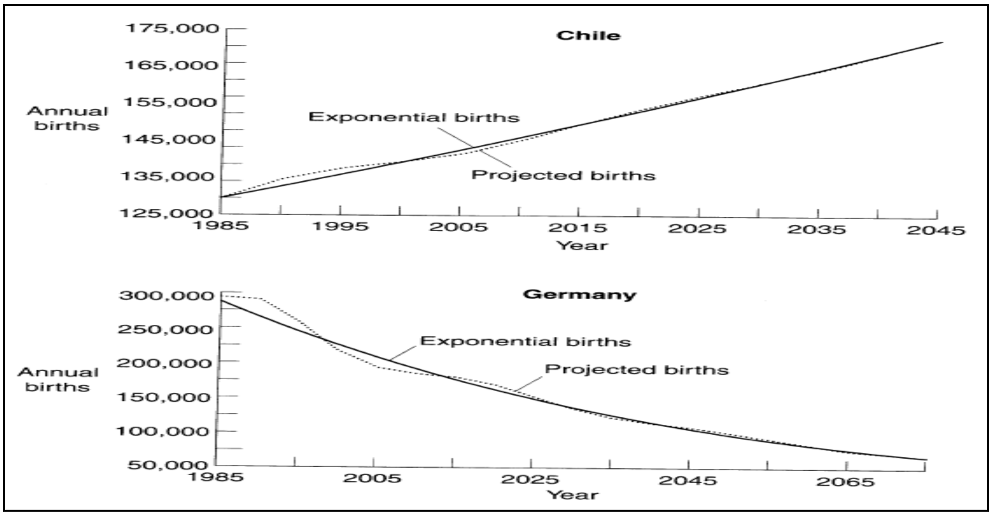
Where $F(t)$ represents the births to females in the initial population, $p(a)$ represents the probability that a female child just born will surviving from birth to age a and $m(a)$ represents the annual rate of bearing female children for women at age a.

Let α and β be the lower and upper limit of the reproductive ages for women which is usually assumed to be 15 and 49 years respectively, then after $t \geq \beta,\ F(t)$ becomes zero and equation (4.16) reduces to

$$B(t) = \int_\alpha^\beta B(t-a)p(a)m(a)da \qquad t \geq \beta \qquad (4.17)$$

Our problem is to solve equation (4.17). In general, the solution of this equation will contain an infinite number of arbitrary constants which are at our disposal. We can choose them so that in addition to satisfying equation (4.17), equation (4.16) is also satisfied. It is seen that equation is homogenous that is to say that the solution remains a solution when multiplied by a constant c (say). Thus if $B(t)$ is the solution then so is $cB(t)$. Moreover, the solutions are additive i.e. if $B_1(t)$ and $B_2(t)$ are solutions of equation (4.17) then $B_1(t) + B_2(t)$ is also a solution of equation (4.17). Thus, a number of solutions can be found that satisfy equation (4.17) and therefore from these special solutions a more general solutions can be built up. Lotka demonstrated that an exponential birth series satisfies this equation, as illustrated in the figure below:

Figure: Birth Trajectories in two Populations subject to Constant Vital Rates



Data Source: Keyfitz and Flieger, 1990; Figure Source: Preston S, Heuveline P, Guillot M.,2001.

To anticipate somewhat if e^{rt} is a solution of equation (2.17) for $r = r_1, r_2, r_3 \ldots r_n \ldots$ then the general solution will be

$$B(t) = Q_1 e^{r_1 t} + Q_2 e^{r_2 t} + Q_3 e^{r_3 t} + \cdots \tag{4.18}$$

where Q_1, Q_2, Q_3 etc. are arbitrary constants. Now Q_i's may be chosen in such a way that equation (4.16) is also satisfied. Thus if $B(t) = e^{rt}$, then putting this in equation (4.17), we get

$$e^{rt} = \int_0^{\beta} e^{r(t-a)} p(a)m(a)da \qquad t \geq \beta$$

or $$1 = \int_0^{\beta} e^{-ra} p(a)m(a)da \qquad \text{for } t \geq \beta \tag{4.19}$$

Equation (2.19) is usually known as fundamental equation or characteristic equation of stable population in r and $p(a)m(a)$ is knows as net maternity function which can be denoted as φ.

There are various methods to determine the value of r, will be given in next unit. Since m(a) and p(a) both are positive valued functions, the right-hand side of equation (2.19) would be a decreasing function of r, say y(r). If r takes values from -Infinity to + infinity, the function y(r) will take values between +infinity to zero and there will be a unique value of r for which y(r) is exactly equal to 1. The value of r for which the right-hand side of the above equation (2.19) equals exactly 1 is termed as **the intrinsic growth rate of the stable population**. The value of r indicates the required annual growth rate that naturally arises from the specific birth and survival schedules m(a) and p(a) that determine stable population dynamics.

4.5.2 Determination of Intrinsic Growth Rate

If α represents the minimum age for childbearing and β denotes the maximum age (with m(a)=0 for all other ages), then equation (4.19) can be expressed as

$$\psi(r) = \int_{\alpha}^{\beta} e^{-ra} p(a)m(a)da = 1 \qquad (4.20)$$

where $p(a)$ is the fraction of female population that survives to age a, $m(a)da$ is the probability, that a female who is of age a will bear a female child in next da period of her life and α and β respectively are the lower and upper limits of the reproductive period and the function $p(a)m(a)$ is known as the net maternity function.

There are a number of procedures for solving equation (4.20). However firstly we will consider some properties of equation (4.20). At first, we will show that equation (4.20) has only one real root. This follows from the fact that the integrand consists of non-negative factors. If the integral is equal to 1 for r, then it cannot also be equal to 1 for ρ > r, since $e^{-\rho a} < e^{-ra}$ for all a. Similarly, a real number less than r cannot be a root of equation (20). Since $\varphi(a) = p(a)m(a) \geq 0$ for all a, and for any value of a, e^{-ra} is a monotonic decreasing function of r, therefore the integral $\psi(r)$ must be decreasing function of r. To show that that $\psi(r)$ is monotonic decreasing function, we note that

$$\psi'(r) = -\int_{\alpha}^{\beta} a e^{-ra} \varphi(a)\, da$$

is always negative as all the factors in the integrand are non- negative. A monotonic decreasing function can cross the line $\psi(r) = 1$ only once at real root say r_1. Also $\psi''(r)$ being positive makes $\psi(r)$ a concave function.

It can be observed that $-\psi'(r)/\psi(r)$ gives the mean age at childbearing in stationary population when evaluated at r =0 and gives mean age at childbearing in stable population for when evaluated at a point r where $\psi(r) = 1$.

Since $\psi(r)$ is a polynomial with only one real root, any additional roots must be complex and will always appear in conjugate pairs.

If we consider a complex root r = x + iy of $\psi(r)$ then

$$e^{ra} = e^{(x+iy)a} = e^{xa}[\cos(ya) + i\sin(ya)]$$

Therefore, $$\psi(r) = \int_{\alpha}^{\beta} e^{-(x+iy)a}\, p(a)\, m(a) da = 1$$

since from De Moivre's theorem

$$e^{-iay} = cos(ay) - i\sin(ay)$$

On substitution in the above equation, we get

$$\psi(r) = \int_{\alpha}^{\beta} e^{xa}[\cos(ya) - i\sin(ya)]\, p(a)\, m(a) da = 1 \tag{4.21}$$

Hence equating the real and imaginary roots of the equation (6), we have,

$$\int_{0}^{\beta} e^{-ax} cos(ay)\, p(a)\, m(a)\, da = 1 \tag{4.22}$$

and

$$\int_{0}^{\beta} e^{-ax} sin(ay)\, p(a)\, m(a)\, da = 0 \tag{4.23}$$

since cos (ay) = cos(-ay) and since equation (4.7) is also satisfied by *-y*, if by *y*, hence we see that complex roots occur in conjugate pairs. i.e. if *x + iy* is a root of equation (4.20) then *x-iy* is also a root of equation (4.20).

Now we will show that all the complex roots have their real parts smaller than r_1 (the real root). $cos(ay)$ has a maximum value 1 and must fall below 1 in any continuous interval of non-zero width while $p(a)m(a)$ is non zero for an extended age range. Therefore, the presence of $cos(ay)$ in the integral must serve to diminish it. Therefore, e^{-ax} should be greater than $e^{-r_1 a}$ which implies that $r_1 > x$.

The argument that the presence of a fixed set of age specific fertility and mortality rates will result in a stable age distribution rests on the fact proved above that real part of any complex root must be less than the real root.

For this purpose, we see that if the real part x of a complex root is less than r_1 then the factor e^{xt} in equation (2.18) will be less than $e^{r_1 t}$ and hence the term $Q_s e^{r_s t}$ involving the factor $e^{r_s t}$ would become small in absolute value compared with the leading term $Q_1 e^{r_1 t}$ as t become large. Hence beyond a suitable t the series in the equation (2.3) may be approximated by its first term and the births are simple exponential $Q_1 e^{r_1 t}$. Henceforth, we shall use r in place of r_1 and Q for Q_1.

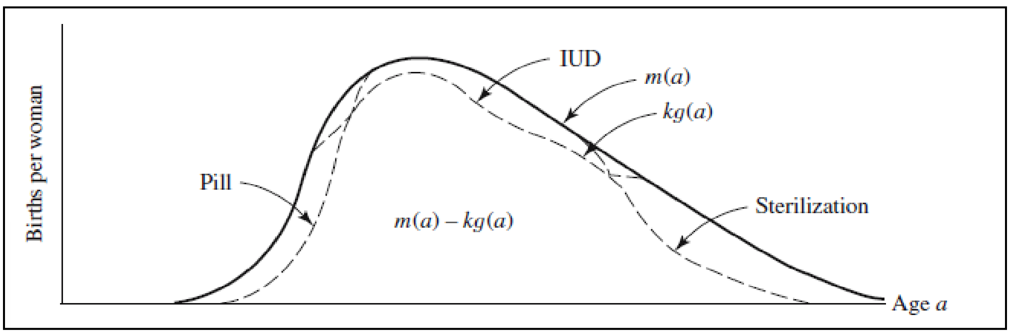
4.5.3 Effects of Changes in Fertility and Mortality on the Intrinsic Growth Rate

The main feature of the stable population model is that when fertility and mortality rates stay consistent over a long period, the age structure of the population will become stable. This implies that whether a population is getting older or younger is not influenced by the current levels of fertility and mortality, but only changes in these rates will affect the age structure. Stable population model is useful for examining how shifts in fertility and mortality impact demographic characteristics over the long term. It helps to compare two stable populations that differ in specific fertility or mortality features. It is also useful to predict the impact of changes in the vital rates on a previously stable population. After implementing a change, the population will eventually reach a new stable state, and its age structure will be compared to the original. The process of moving to this new equilibrium is not considered, as the stable population theorem ensures that a new equilibrium will be reached.

4.5.4 Effect of Small Arbitrary Change in Birth Function

The birth rate of India keeps declining over time, it falls more rapidly at some ages than at others due to various socio-economic norms. As the women are getting more educated and empowered, the early age birth rates have declined considerably. Also, the increasing access to birth control methods have dramatic impacts on birth rates. Keyfitz has shown the impact of some birth control methods on the maternity function by assuming a function kg(a) as a reduction in birth function m(a) such that kg(a) is a small arbitrary change in shape taken out of m(a) and is given in the figure below:

Figure: Area *k(g)a* removed from maternity function by several methods of birth control



Source: Nathan Keyfitz, Hal Caswell (1977).

The new value $r+\Delta r$ can be obtained from the characteristic equation

$$\int_{\alpha}^{\beta} e^{-ra} p(a)m(a)da = 1$$

in the form as

$$\int_{\alpha}^{\beta} e^{-(r+\Delta r)a} p(a)[m(a) - kg(a)]da = 1 \qquad (4.24)$$

Expanding the exponential term $e^{-\Delta ra} \approx 1 - \Delta ra$ (assuming k to be small enough that the second-order terms Δr can be ignored) in (4.24), ignoring the terms with $k\Delta r$ and solving for Δr, we can write

$$\Delta r \approx -\frac{\int_{\alpha}^{\beta} e^{-ra} p(a)m(a)da}{\int_{\alpha}^{\beta} ae^{-ra} p(a)m(a)da} k$$

or

$$\Delta r \approx -\frac{\int_\alpha^\beta e^{-ra}\, p(a)m(a)da}{\bar{A}} k \tag{4.25}$$

Where $\bar{A} = \int_\alpha^\beta ae^{-ra}\, p(a)m(a)da$ represents the mean age of childbearing in stable population.

If $g(a) = m(a)$ for all a, (4.25) becomes

$$\Delta r = -\frac{k}{\bar{A}} \quad or \quad k = -\bar{A}\Delta r$$

The question is asking how much the age-specific rates need to change to reduce the overall rate of increase by Δr. The answer is that reducing these rates by $100\bar{A}\Delta$r percent at each age will do it. The approach is useful for comparative study, where birth rates are compared between two populations: one with rates m(a) and another with adjusted rates m(a)−kg(a), and then their stable conditions are analysed.

4.5.5 Change in Birth rate for Drop to Bare Replacement

Bare replacement level means is the threshold value of birth rate at which the value of NRR becomes 1. Keyfitz has given expression for reduction in fertility to bare replacement in the following manner.

The net reproduction rate which gives the expected number of girl children by which a girl child born now will be replaced must be unity for bare replacement level fertility. Thus, if for $m(a) - kg(a)$, the NRR becomes unity, we must have

$$\int_\alpha^\beta p(a)[m(a) - kg(a)]da = 1$$

As $R_0 = \int_\alpha^\beta p(a)m(a)da$

The solution of this equation for k will give the amount required for the reduction to replacement by the change in the shape of g(a) in the age pattern. And the solution is given as

$$k = \frac{R_0 - 1}{\int_\alpha^\beta p(a)g(a)\, da}$$

And for g(a)=m(a), this becomes $k = \frac{R_0 - 1}{R_0}$

4.5.6 Effect of Uniformly Lower Death Rates

Effect of decline in mortality on stable age distribution is more complex in general to understand as a neutral mortality change does not affect largely. A neutral change in mortality occurs when there is a n equal reduction in death rates across all age groups.

Suppose there is a uniform change in mortality at all ages say k, then we can write,

$$\mu_a' = \mu_a - k \text{ for a>0}$$

This gives age specific death rates after change in mortality by amount k at all ages. Due to this change in mortality, the effect on the survival function will become

$$p'(a) = e^{-\int_0^a (\mu_x - k)dx} = p(a)e^{ka}$$

As $p(a) = e^{-\int_0^a \mu_x dx}$,

i.e the change in survival function at age a due to a constant mortality change will be only by a factor of e^{ka} to the original survival function.

Its characteristic equation can be written as

$$\int_\alpha^\beta e^{-r'a}\, p'(a)m(a)da = 1$$

Or

$$\int_\alpha^\beta e^{-(r'-k)a}\, p(a)m(a)da = 1$$

It is identical with the characteristic equation for the population without change except that its intrinsic growth rate becomes $r = (r' - k)$ or $r' = r + k$. It is equivalent to say that the solution with declined mortality is the original solution plus the decline in mortality.

The new age structure due to this change will be

$$c'(a) = \frac{e^{-r'a}p'(a)}{\int_{\alpha}^{\beta} e^{-r'x}\, p'(x)dx}$$

$$= \frac{e^{-(r+k)a}p(a)e^{ka}}{\int_{\alpha}^{\beta} e^{-(r+k)x}p(x)e^{kx}\, dx}$$

$$= \frac{e^{-ra}p(a)}{\int_{\alpha}^{\beta} e^{-rx}p(x)\, dx} = c(a)$$

This shows that a constant decline in mortality does not affect the age distribution. In more general sense it can be said that a constant change in mortality at all ages has no effect on the age distribution of a population.

Suppose an age distribution, however irregular, and a decline of 0.001 in mortality rates at all ages. Then exactly one person in a thousand who would have died on the former regime now survives, age by age. This increases the number at every age by exactly 0.001, that is, multiplies it by 1.001, and multiplying every age by 1.001 can have no effect on the age distribution (Coale 1956).

4.6 Summary

The present chapter provides the understanding of population growth rate and models. It also focuses on intrinsic rates and their relationship to understand their course under the regime of stability and stationarity of a population. Estimation of growth rates and other parameters in under these population models, through the basic mathematical models have also been explained. The fundamental of stable population theory that is the derivation of Lotka's equation of stable population have been determined. The problem of estimation of 'r' for stable population equation has also been explained.

4.7 Self-Assessment Exercises

1. If population of Kerala according to census 1991 and 2001 was 29,011,000 and 31,839,000 respectively. Then, estimate the population for year 1996 using Arithmetic, Geometric and Exponential growth models.

2. For the data given for the population of Kerala in Q.1. Then, project the population for 2005 using Arithmetic, Geometric and Exponential growth models.
3. Derive the equation of stable population given by Lotka. Also explain the effect of small arbitrary changes in the birth function in the stable population equation.

4.8 References

- Gompertz, B. (1825). *On the Nature of the Function Expressive of the Law of Mortality*, Philosophical Transactions, 27: 513-85.

- Coale, A.J. (1956). The effects of changes in mortality and fertility on age composition. Milbank Memorial Fund Quarterly 34:79–114.

- Ram, F., & Pathak, K. B. (1998). Techniques of Demographic Analysis. Mumbai: Himalaya Publishing House.

- Nathan Keyfitz, Hal Caswell (1977). Applied Mathematical Demography.Samuel H. Preston, Patrick Heuveline, Michel Guillot (2001). Measuring and Modeling Population Processes.

- Spiegelman, M., (1980) *Introduction to Demography*, (Revised Edition), Harvard University Press, Cambridge.

2.9 Further Readings

- Samuel H. Preston, Patrick Heuveline, Michel Guillot (2001). Measuring and Modeling Population Processes.

- Shryock, H.S. et al., (1976) *The Methods and Materials of Demography,* Condensed Edition by Stockwell, E. G., Academic press, New York.

- United Nations, (1956) *Manual III Methods for Population Projections by Age and Sex*, New York.

UNIT - 5 GROWTH RATES

Structure

5.1 Introduction

In stable Population theory, the intrinsic growth rate is the rate at which a population closed to migration grows with constant fertility and mortality schedules for a long time. In the previous chapters, we read about the concepts of stable population theory, and its variations depending upon different conditions such as stationary population, semi-stable, quasi stable and non-stable populations. We also came across some population growth models and estimation of growth rates and other parameters in these models that help in the estimation of future and unknown populations. We have also studied definitions of parameters involved in stable model population and the genesis of the basis for stable population i.e. the characteristic equation

$$\int_{\alpha}^{\beta} e^{-ra} p(a) m(a) da = 1$$

And then studied how to determine the intrinsic growth rate in the stable population using this fundamental equation.

The present chapter helps to estimate the natural growth rate of stable population through different methods as per the conditions and availability of data. It will also provide a clear understanding of generational differences in population through the notion of mean length of generation and then we shall find some relationships among different averages or mean ages in stable population. We shall also try to understand some other important concepts such as reproductive value, population momentum etc. and their applicability in real life conditions.

5.2 Objectives

- Use a range of methods to calculate the intrinsic growth rate of a stable population, such as generation length, method of moments and iterative methods
- To study the concept of mean interval between generations.
- To find relationship between Net Reproduction Rate (NRR), intrinsic rate of natural increase and mean length of generation
- To study some advance terminologies occurring in stable population models such as reproductive value, replacement level fertility, population momentum etc.

5.3 Method to Calculate Intrinsic Growth Rate of Stable Population from given Fertility and Mortality Schedules

As stated in the previous chapter, the natural rate of change in a stable population is termed as intrinsic growth rate.

Let us again right the Lotka's fundamental equation

$$\int_{\alpha}^{\beta} e^{-ra} p(a)m(a)da = 1$$

where $p(a)$ is the fraction of female population that survives to age a, $m(a)da$ is the probability, that a female who is of age a will bear a female child in next da period of her life and α and β respectively are the lower and upper limits of the reproductive period and the function $p(a)m(a)$ is known as the net maternity function.

There are several methods to compute this rate depending upon the availability of data and assumptions, some are given below:

5.3.1 Using Method of Moments

Now, by expanding e^{-ra} within the integrand in the equation above and neglecting higher than 2nd orders of r , the above equation can be rewritten as

$$\int_{\alpha}^{\beta} \left(1 - ra + \frac{r^2 a^2}{2}\right) p(a)m(a)da = 1$$

or

$$\int_{\alpha}^{\beta} \varphi(a)da - r\int_{\alpha}^{\beta} a\varphi(a)da + \frac{r^2}{2}\int_{\alpha}^{\beta} a^2\varphi(a)da = 1$$

where $\varphi(a) = p(a)m(a)$ is the net maternity function.

The above equation can be written as

$$R_0\left(1 - \frac{rR_1}{R_0} + \frac{1}{2}\frac{r^2 R_2}{R_0}\right) = 1 \qquad (5.1)$$

Where $R_k = \int_{\alpha}^{\beta} a^k \varphi(a)da$ is the r^{th} moment of the net maternity function φ(a).

Taking logarithm on both the sides,

$$\log_e R_0 + \log_e R_0 \left(1 - \frac{rR_1}{R_0} + \frac{1}{2}\frac{r^2 R_2}{R_0}\right) = 0 \quad (5.2)$$

Again, neglecting the terms containing higher powers of r than two, we have

$$\log_e R_0 - \frac{rR_1}{R_0} + \frac{1}{2}\left[\frac{R_2}{R_0} - \left(\frac{R_1}{R_0}\right)^2\right] r^2 = 0$$

or $$\frac{1}{2}\, x_1 r^2 + x_2 r - \log_e R_0 = 0 \quad (5.3)$$

where $x_2 = \frac{R_1}{R_0}$ and $x_1 = \left(\frac{R_1}{R_0}\right)^2 - \frac{R_2}{R_0}$

As mentioned above, R_0 is the net reproduction rate and R_1 and R_2 are the first and second moments of the net maternity function. Solving we have

$$r = \frac{1}{x_1}\left(-x_2 + \sqrt{x_2^2 + 2x_1 \log_e R_0}\right) \quad (5.4)$$

This gives the intrinsic rate of growth of the population.

Further, a first order approximation of r gives (from 3.3)

$$r = \frac{1}{x_2}\log_e R_0$$

or $$R_0 = e^{x_2 r}\ , \quad (5.5)$$

Using the value of $x_2 = \frac{R_1}{R_0}$, one can easily estimate value of r.

5.3.2 Using Length of Generation

Since the net reproduction rate NRR gives the rate of growth in one generation of length T and r is the compound annual rate of growth in stable population, this gives us

$$R_0 = e^{rT} \quad (5.6)$$

On comparison of these two expressions (3.5) and (3.6) of R_0, we get

$$T = x_2 = \frac{1}{r}\log_e R_0 \qquad (5.7)$$

Substitution of the value of $\log_e R_0 = \frac{1}{2}\, x_1 r^2 + x_2 r$ from (5.3) in above equation gives

$$T = x_2 + \frac{1}{2} x_1 r$$

On simplification, we get

$$r = \frac{2(T - x_2)}{x_1}$$

Every year, stable population increases $(1 + r)$ times than the previous year and so it increases in proportion to the net reproduction rate by the end of a generation and if r is negative, R_0 will be less than one.

5.3.3 Indirect Method Using Census Data

The value of r can be estimated indirectly in case of lack of registration data on births and deaths using intercensal growth rate of the population if population is assumed to be stable and two census populations are known. We can also estimate r using only one census data on age composition of a stable population by choosing any two ages x and y.

From stable population theory, we can write for a fraction of population aged x as

$$c(x) = be^{-rx}p(x)$$

And for the fraction of population aged y,

$$c(y) = be^{-ry}p(y)$$

Taking logarithms of these two equations reduce them to a pair of linear equations of first degree and solving these two linear equations for r gives an estimate of r as

$$\hat{r} = \frac{1}{y-x} ln \frac{c(x)/p(x)}{c(y)/p(y)}$$

For this approximation to hold good, c(x) and c(y) need to be the number of persons at age x and y respectively to the nearest birthdays, A better approximation can be obtained by replacing p(x) and p(y) the life table functions

$$L_x = \frac{1}{2}(l_x + l_{x+1}) \quad \text{and} \quad L_y = \frac{1}{2}(l_y + l_{y+1})$$

Although it is the simplest way to provide an estimate of r using current age distribution, it is a vague one as there can be many estimates using different pairs of age distributions. Therefore, one can use group of ages and then average these estimates to get a single better estimate.

5.3.4 Using Functional Iteration Method

To determine the roots of the characteristic equation for a stable population using data from a life table and age-specific female birth rates, it is common practice to approximate

$$\psi(r_n) = \int_\alpha^\beta e^{-r_n a} p(a)m(a)da = \sum_{a=15,5}^{45} e^{-r_n(a+2.5)} {}_5L_a * {}_5m_a$$

Where, ${}_5L_a$ is the number of person years lived between ages a and a+5 (from female period life table with l_0=1)

${}_5m_a$= Rate of bearing female children between ages a and a+5

Therefore, using the iteration formula as given below

$$r_{n+1} = r_n + \frac{\psi(r_n) - 1}{T} \tag{5.8}$$

Considering length of generation T to be 27.5 years generally, and starting with r_0=ln NRR/T, and continuing till the difference between two consecutive values of r_{n+1} becomes very small will give the desired value of intrinsic r of the stable population.

For calculation purpose, the continuous forms of the equations are of no use, these must be converted to discrete approximations. Thus, we can write for

$$\int_x^{x+5} e^{-ra} p(a)m(a)da = e^{-(x+2.5)}{}_5L_xF_x$$

Where ${}_5L_x = \int_x^{x+5} p(a)da$ and F_x is the observed age specific fertility rate for age x to x+4.

Steps to Calculate r Using Iteration Method:

1. Identify the model life table for the population using Regional Model Life Table of Coale and Demeny 1966 as per the vital situations of the population under consideration.
2. Write the following values in a tabular form using the selected life table to make calculations:
   a. ${}_5L_a$ =Number of person-years lived between ages a and a+5 (from female period life table with l_0=1
   b. ${}_5m_a$= Rate of bearing female children between ages a and a+5
3. Use r_0=ln NRR/T to find the initial guess for r
4. Compute $y(r_n) = \sum_{a=15,5}^{45} e^{-r_n(a+2.5)}{}_5L_a * {}_5m_a$ and using this value, compute $r_{n+1} = r_n + \frac{y(r_n)-1}{T}$
5. Repeat step 3 and 4 till the value of y(r) becomes almost equal to 1. The value of r so estimated will be the intrinsic growth rate of the population.

Example: 1

Table 3.1 Calculation of Intrinsic rate of growth of the female population of India from the ASFRs available for 1971-1972.

Age group a to a+5	Middle age a+2.5	${}_5L_a$	${}_5m_a$	${}_5L_a.{}_5m_a$	$\exp(-r_n(a+2.5))\,{}_5L_a.{}_5m_a$				
					r_0	r_1	r_2	r_3	r_4
15-19	17.5	3.68127	0.04261	0.15686	0.10754	0.10875	0.10875	0.10872	0.10875
20-24	22.5	3.60203	0.12775	0.46016	0.28321	0.28733	0.28733	0.28723	0.28733
25-29	27.5	3.50220	0.13469	0.47171	0.26063	0.26528	0.26527	0.26516	0.26527
30-34	32.5	3.34718	0.10582	0.35420	0.17569	0.17940	0.17940	0.17931	0.17940
35-39	37.5	3.14970	0.07045	0.22190	0.09881	0.10122	0.10122	0.10116	0.10122
40-44	42.5	2.91244	0.03681	0.10721	0.04286	0.04405	0.04404	0.04402	0.04404
45-49	47.5	2.65090	0.01427	0.03783	0.01358	0.01400	0.01440	0.01399	0.01400
				1.80986	0.98232	1.00004	1.00042	0.99959	1.00001

***Source*:** * Office of the Registrar General, India, *Fertility Differentials in India*, 1972, Ministry of Home Affairs, New Delhi.

** Registrar General of India, *Life Tables, Series, 1- India*, Part 1 of India, Government of India, New Delhi (1977), Appendix A

NRR= 1.89 daughters per women

r_0 = log(1.89)/27.5=0.02157 y(r_0)= 0.98232

r_1=0.02157+(0.98232-1)/27.5=0.02093 y(r_1)= 1.00004

r_2=0.02093+(1.00004-1)/27.5=0.02093 y(r_2)= 1.00042

r_3=0.02093+(1.00042-1)/27.5=0.02095 y(r_3)= 0.99959

r_4=0.02095+(1.00042-1)/27.5=0.02093 y(r_4)= 1.00001

And thus using iteration method we get, r=0.02093 as the intrinsic growth rate for female population of India 1971-72.

5.4 Computation of Stable Age Distribution from the Given Fertility and Mortality Schedules

To obtain the age distribution of a stable population corresponding to the prevalent levels of fertility and mortality schedules, we need to first compute the intrinsic growth rate of the population using an estimate of NRR and the mean age of child bearing. To do so, we will first calculate the NRR and the mean age of child bearing with the help of ASFRs and the life table survival probabilities. The nLx column of the life table provides the stationary female population per 100,000 female births (100,000 equivalent to radix l_0). Next, we calculate $e^{-r(x+2.5)}$ for each five year age group where x=15, 20,…,45 and $e^{-r(x+2.5)} * nLx$ gives the female stable population. To get the male equivalent stable population, we can simply multiply $_nL_x$ columns by SRB (generally taken as 1.05) with $1.05e^{-r(x+2.5)}$ and then the total stable population by summing the male and female populations for all corresponding ages. The Illustration being given will help understand the procedure to calculate the intrinsic growth rate and other parameters.

Table 3.2 Calculation of intrinsic growth rate of female population of Uttar Pradesh from ASFR data available from Census 2001

Age Group (Years)	**Age-Specific Fertility Rate (ASFR) Women**	**Women in stationary population $_nL_x/l_0$**	**Births to women in stationary population 2*3**	**Mid-point of age group R_0**	**col 4*5 R_1**	**col 5*6 R_2**
1	2	3	4	5	6	7
15-19	0.0538	4.94	0.265772	17.5	4.65101	81.39268
20-24	0.2412	4.92	1.186704	22.5	26.70084	600.7689
25-29	0.2689	4.9	1.31761	27.5	36.23428	996.4426
30-34	0.1563	4.88	0.762744	32.5	24.78918	805.6484
35-39	0.0635	4.85	0.307975	37.5	11.54906	433.0898
40-44	0.0112	4.81	0.053872	42.5	2.28956	97.3063
45-49	0.001	4.76	0.00476	47.5	0.2261	10.73975
Total	0.7959		3.899437		106.44	3025.388

Note: NRR is computed by multiplying the total of column 4 by S'= $1/(1+S)$ where S is SRB taken as 1.05 male per female.

From the table shown above, we can obtain

NRR=R_0= (1/(1+S)) *∑col (4) = 1.9022

GRR = (1/(1+S)) *5∑col (1) = 1.9412

Since the Age-Specific Fertility Rate (ASFR) encompasses both males and females, the total is divided by the sex ratio at birth (S' = 1 / (1 + S) = 0.4848) to calculate the Net Reproduction Rate (NRR). The stationary population is then derived from the female life tables.

The mean age of childbearing (T) = total of column 6/ Total of column 4

= 106.44/3.899437

= 27.29626

Intrinsic rate of natural increase (r) = log NRR/T

=log (1.9022)/ 27.29626 =0.02355607

This indicates that the population will keep increasing if these levels of fertility and mortality continue in a long time period.

Now it becomes easy to calculate the stable population as the intrinsic growth rate and stationary female population are available. This will be calculated by multiplying the stationary population for each age group by exp(−rx) to determine the stable population within that age group, where x represents the average age of that age group. However, only female population has been considered. To include the male population as well, one need to repeat the process separately using stationary populations for males. But the radix in this life table should not be usually 1000, it instead should be increased to say 1080, making the ration of the radices of male and female life tables equal to the sex ratio at birth. The calculations of the age distributions of stable population for Uttar Pradesh is shown in the table below:

Table 3.3: Calculations of the age distributions of stable population for Uttar Pradesh, 2001

age	lx values		Stationary population nLx		Average age	Factor	Stable Population	
	F	M	F	M	X	exp(-rx)	F	M
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
0	1000	1080	993	1065	0.3	0.992958	986	1058
1	992	1063	3967	4244	2.6	0.940592	3731	3992
5	991	1059	4949	5290	7.5	0.838056	4148	4433
10	988	1056	1939	5274	12.5	0.744941	1444	3929
15	988	1053	4935	5257	17.5	0.662172	3268	3481
20	986	1050	4924	5234	22.5	0.588599	2898	3081
25	983	1043	4902	5196	27.5	0.523201	2565	2719
30	977	1035	4876	5146	32.5	0.465069	2268	2393
35	973	1024	4849	5084	37.5	0.413396	2005	2102
40	967	1010	4810	4985	42.5	0.367464	1768	1832
45	957	984	4757	4864	47.5	0.326636	1554	1589
50	946	961	4677	4690	52.5	0.290344	1358	1362
55	925	915	4560	4420	57.5	0.258084	1177	1141
60	899	854	4387	4037	62.5	0.229409	1006	926
65	856	761	4112	3498	67.5	0.203919	839	713
70	789	638	3623	2811	72.5	0.181262	657	510
75	660	486	2899	2024	77.5	0.161123	467	326
80	499	323	1948	1209	82.5	0.14322	279	173
85+	280	160	1692	782	87.5	0.127307	215	100
Total							32640	35867

5.5 Relations Between NRR and Natural Increase R

To calculate both the intrinsic growth rate and the net reproduction rate, we use the same data the fertility and mortality schedules m(a) and p(a) respectively for a specific population. One serves as indicators of long-term growth potential i.e. the intrinsic growth rate represents the annual growth rate that would apply if current rates remain constant, while the net reproduction rate indicates the growth factor from one generation to the next. Therefore, it is expected that these two measures are closely related analytically.

We know that for a stable population

$$\int_0^{\beta} e^{-ra} m(a)\, p(a) da = 1$$

And the net reproduction rate NRR is given as

$$N.R.R. = \int_{\alpha}^{\beta} m(a)\, p(a)\, da$$

comparing the above two equations, we see that

if N.R.R. = 1 , then r should be zero

if N.R.R. > 1, then r should be +ve

if N.R.R. < 1, then r should be -ve

This relation in NRR and r indicates that if p(a) and m(a) schedules are such that if a population is growing in each generation (for N.R.R. >1), then it would also be growing annually and vice-versa.

Note: One should not confuse with limits of the integrand of the fundamental equation. Here

$$\int_0^{\infty} \varphi(a) da = \int_0^{\beta} \varphi(a) da = \int_{\alpha}^{\beta} \varphi(a) da$$

Where $\varphi(a) = e^{-ra}\, p(a)\, m(a)da$ since below α and above β, $m(a) = 0$

5.6 Mean Length of Generation in a Stable Population

The average number of daughters per woman subject so the given schedules of fertility and mortality is a measure of proportionate increase per generation in the stable population. However, generation is not a specific time interval, that is it is not same for all populations. Thus, the mean length of generation is the number of years required to multiply the stable population by N.R.R. i.e.

$$e^{rT} = \text{N.R.R.} \tag{5.9}$$

The above equation provides a basis for estimation the value of r if the mean length of generation is known. Thus, T depends on the ages at which child bearing is concentrated. It would be greater in a population where child bearing occurs late as compared to those where child bearing is early. Theoretically T represents the amount of time (in years) required for a population growing at a rate r to increase by a factor of the given by NRR. For example, if r=.015 and NRR=1.5, T will be 27.03 indicating that the stable population will need 27.73 years to expand by the amount specified by the Net Reproductive Rate (NRR).

Let us consider, briefly, the following mean ages of fertility which are used in the computation of T. These are

1) $\bar{m}$, the mean age of the fertility schedule defined as

$$\bar{m} = \frac{\int_a^\beta am(a)da}{\int_\alpha^\beta m(a)da} = \frac{\int_a^\beta am(a)da}{GRR}$$

2) μ_1, the mean age of net fertility schedule as

$$\mu_1 = \frac{\int_a^\beta am(a)p(a)da}{\int_\alpha^\beta m(a)p(a)da} = \frac{\int_a^\beta am(a)p(a)da}{NRR} \tag{5.10}$$

3) $\bar{A}$, the mean age of child bearing in the stable population is defined as

$$\bar{A} = \frac{\int_a^\beta ae^{-ra}m(a)p(a)da}{\int_\alpha^\beta e^{-ra}p(a)m(a)da} = \int_a^\beta ae^{-ra}m(a)p(a)da \tag{5.11}$$

Now let us derive some relationships among these expressions under stability conditions.

Let us derive a relationship between a pair of simpler measures namely G.R.R. and N.R.R. which deal with increase in the population. Here we have

$$G.R.R. = \int_{\alpha}^{\beta} m(a)\, da$$

And

$$N.R.R. = \int_{\alpha}^{\beta} m(a)\, p(a)\, da$$

Since p (a) in most of the populations follows nearly a linear pattern between α to β. Assuming such a linear pattern in p(a), the N.R.R. is approximated as N.R.R. = p($\bar{m}$) G.R.R., where $\bar{m}$ is the mean age of fertility schedule which is defined as

$$\bar{m} = \frac{\int_{\alpha}^{\beta} am(a)da}{\int_{\alpha}^{\beta} m(a)da}$$

To prove this, let p(a) = u – va, where u and v are constants.

Then

$$N.R.R. = \int_{\alpha}^{\beta} (u - va)\, m(a)\, p(a)\, da$$

$$= u.G.R.R. - v\bar{m}.G.R.R.$$

$$= (u - v\bar{m})G.R.R.$$

$$= p(\bar{m})G.R.R.$$

Thus, N.R.R. can be computed with the help of only fertility schedule, if related data on mortality schedules are not available or are unreliable.

As mention above, it is seen that p(a) follows approximately a linear pattern from α to β, hence if we take p(a) approximately a straight line as p(a) = u – va, then, one has the following relation

$$\mu_1 = \bar{m} - \frac{v\sigma^2}{p(\bar{m})} \tag{5.12}$$

Where σ^2 is the age variance of fertility schedule.

To prove this relation, we have

$$\mu_1 = \frac{\int_\alpha^\beta am(a)p(a)da}{\int_\alpha^\beta m(a)p(a)da}$$

On substituting the linear approximation for p(a) in the above expression, we get

$$\mu_1 = \frac{\int_\alpha^\beta am(a)(u - va)da}{\int_\alpha^\beta m(a)(u - va)da}$$

$$= \frac{u\int_\alpha^\beta am(a)da - v\int_\alpha^\beta a^2m(a)da}{u\int_\alpha^\beta m(a)da - v\int_\alpha^\beta am(a)da}$$

$$= \frac{u.\bar{m}.G.R.R. - v(\sigma^2 + \bar{m}^2)G.R.R.}{u.G.R.R. - v.\bar{m}.G.R.R.}$$

$$= \frac{\bar{m}(u - v\bar{m}) - v\sigma^2}{u - v\bar{m}}$$

Thus, we get,

$$\mu_1 = \bar{m} - \frac{v\sigma^2}{p(\bar{m})} \tag{5.13}$$

Since we know that the chance of surviving in the interval a to $(a + da)$ is $1 - \mu(a)d(a)$ which is approximately equal to $e^{-\mu(a)da}$. The chance of surviving to a sequence of small intervals $(a_0, a_1), (a_1, a_2), (a_2, a_3), \ldots, (a_{n-1}, a_n)$ is approximately equal to, $e^{-\mu(a_0)(a_1 - a_0)}$, $e^{-\mu(a_1)(a_2 - a1)}$, ..., $e^{-\mu(an_1 - 1)(a_n - a_{n-1})}$. As the intervals are small it is permissible to replace the summation by the integral so that the probability of survival from a_0 to a_n is $e^{-\int_{a_0}^{a_n}\mu(x)dx}$.

Thus, we can write, p(a) = $e^{\int_0^a \mu(x)dx}$

Taking logarithm of both the sides, in

$$log\, p(a) = \int_0^a \mu(x)dx$$

which can be rewritten as

$$\frac{d}{da} log\, p\,(a) = -\mu(a) \tag{5.14}$$

we are having p(a) = u – va, which gives

$$\frac{d}{da} p(a) = -v$$

Or we can write it as

$$\frac{1}{p(a)} \frac{d}{da} p(a) = \frac{-v}{p(a)} \tag{5.15}$$

Comparison of the above two equations (14) and (15) give us

$$\frac{-v}{p(a)} = -\mu(a) \quad or \quad \frac{v}{p(a)} = \mu(a)$$

Or we can write

$$\frac{v}{p(\bar{m})} = \mu(\bar{m})$$

Thus, from the equations (13),we get

$$\mu_1 = \bar{m} - \sigma^2 \mu(\bar{m}) \tag{5.16}$$

Further, when the expression Ā is expanded, it is found that (after ignoring the terms higher than second order in r, taking r as small)

$$\bar{A} = \mu_1 - \mu_2 r + \mu_3 \frac{r^2}{2} \tag{5.17}$$

Where μ_n is the nth cumulant of net fertility function which is obtained from the relationship

$$\left[\frac{d^n}{dr^n}(\bar{A})\right]_{r=0} = (-1)^n \mu_{n+1} \tag{5.18}$$

5.6.1 Mean Length of Generation using Mean Age of Child Bearing in the Stable Population:

Now let us derive the expression for T i.e. the mean length of generation given by (Coale 1972), as

$$T = \frac{\int \bar{A}(r)dr}{r}$$

Suppose the fundamental equation for stable population is expressed as

$$y(s) = \int_{\alpha}^{\beta} e^{-sa} m(a)p(a)da$$

Differentiating the above equation with respect to *s* we get-

$$\frac{d}{ds} y(s) = -\int_{\alpha}^{\beta} a e^{-sa} p(a)m(a)da$$

Since we know that $\bar{A}(s)$ is the mean age of child bearing in stable population, can be written as

$$\bar{A}(s) = \frac{\int_{\alpha}^{\beta} a e^{-sa} p(a)m(a)da}{\int_{\alpha}^{\beta} e^{-sa} p(a)m(a)da}$$

And hence we can write,

$$\frac{d}{ds} y(s) = -y(s)\bar{A}(s)$$

or $$\frac{1}{y(s)}\frac{d}{ds}y(s) = -\bar{A}(s)$$

or $$\frac{d}{ds}log\,y(s) = -\bar{A}(s)$$

Integrating the above, we get

$$log\,y\,(s) = \int \bar{A}(s)ds + c \tag{5.19}$$

The value of c is determined by putting the value of s = 0 in the equation (5.19) as

$$log\,y(0) = -(\int \bar{A}(s)ds)_{s=0} + c$$

Here $\int \bar{A}(s)ds$ for $s = 0$ is zero

$c = log\,y(0)$

$$log\,y(s) = -\int \bar{A}(s)ds + log\,y(0)$$

thus $$y(s) = y(0)e^{-\int \bar{A}(s)ds}$$

Now $$y(0) = \int_{\alpha}^{\beta} p(a)m(a)da = N.R.R.$$

and hence $$y(s) = N.R.R.e^{-\int \bar{A}(s)ds}$$

if r represents the rate of increase in the stable population, then,

y (r) = 1

and then we get

$$1 = N.R.R.e^{-\int \bar{A}(r)dr} \;\; i.e. \;\; N.R.R. = e^{\int \bar{A}(r)dr}$$

But we know that N.R.R. = e^{rT}

hence $$rT = \int \bar{A}(r)dr$$

or $$T = \frac{\int \bar{A}(r)dr}{r} \qquad (5.20)$$

Further after expanding, it is

$$\frac{\int \bar{A}(r)dr}{r} = \frac{\mu_1 r - \mu_2 \frac{r^2}{2} + \mu_3 \frac{r^3}{3}}{r}$$

Thus, we can write

$$T = \mu_1 - \mu_2 \frac{r}{2} + \mu_3 \frac{r^3}{3}$$

If ‘r’is very small such that the terms involving r^2, r^3, …etc. are neglected then we have the relationship

$$T = \mu_1 - \mu_2 \frac{r}{2} \qquad (5.21)$$

where μ_2 is the age variance of net fertility schedule.

We can also right from equation (5.17) on neglecting the higher order term for r as

$$\bar{A} = \mu_1 - \mu_2 r \qquad (5.22)$$

On comparing the above two equations, we can write

$$T = \frac{\mu_1 + \bar{A}}{2}$$

We know that σ^2 is the age variance of fertility schedule and μ_2 is the age variance of net fertility schedule. If we assume that μ_2 and σ^2 are approximately equal then we have a further approximation for given r as

$$T = \bar{m} - \sigma^2(\mu(\bar{m}) + \frac{r}{2}) \qquad (5.23)$$

The intrinsic rate of increase can also be determined by the following relationship

$$\frac{r}{2} \approx \frac{[\log G.R.R + \log p(\overline{m})]}{2\overline{m}} \tag{5.24}$$

Since we know that e^{rT} = N.R.R.

and $N.R.R. = p(\bar{m}).\ G.R.R.$

which can be written as-

$$log\, N.R.R. = log\, p(\bar{m}) + log\, G.R.R..$$

$$\Rightarrow \qquad T = \frac{log\, p(\bar{m}) + log\, G.R.R.}{r}$$

The above equation gives approximation for T which is purely a function of the fertility schedule.

For a more practical understanding, some other terminologies useful in the applications of stable population theory are given below:

5.7 Reproductive Value

Reproductive value is a concept introduced by Ronald A. Fisher in1930 as part of his work on population genetics. In a stable population, the reproductive value of an individual at a given age represents the expected contribution of that individual to the future population, considering both their current and future reproductive potential.

In the context of stable population, it can be interpreted as the discounted sum of future reproductive output, where the discounting reflects both the probability of surviving to each future age and the growth rate of the population.

Keyfitz in his book has given this notion as loan and repayment interpretation of the characteristic equation by suggesting to calculate how much of the debt is outstanding by the time to the girl that has reached age x< β or say the expected number of subsequent children discounted back to age x.

Fisher wrote for reproductive value,

$$v(x) = \frac{1}{e^{-rx}p(x)} \int_x^{\beta} e^{-rx} p(a)m(a)\, da$$

Where $\frac{p(a)}{p(x)} m(a)$ is the expected number of births in the interval a to a+da, a>x and the above equation represents the debt outstanding or births outstanding at age x when the expected births are discounted back a-x years.

v(x) is called as reproductive value at age x. In particular, for v(0)=1, and for x> β, v(x)=0. In a stable population, younger individuals generally have higher reproductive values as they have more potential future reproductive years ahead of them.

5.8 The Stable Equivalent Population

Every population at a time has specific fertility and mortality schedules say m(a) and p(a) respectively and if those rates are to continue indefinitely, it will lead to a stable population. This theoretical stable population is often called the "stable equivalent" population. The resulting growth rate, birth rate, death rate, and age distribution of this model population are referred to as "intrinsic." These intrinsic characteristics are determined by the m(a) and p(a) schedules and are not affected by the actual, irregular age distribution of the population in the beginning. The ergodic property of population dynamics ensures that the model stable population will be equivalent to the projected population over time, no matter the starting point.

A population will be stable if its age-specific vital rates have remained constant over a long period and in such cases, the "stable equivalent" population will match the actual population in terms of crude birth rate, death rate, growth rate, and age structure. However, if fertility or mortality rates have been variant, the actual population will likely differ from the stable equivalent population in some or all of these characteristics. To determine if a population is stable, it can be verified by comparing the age distributions from two censuses if age-specific growth rates are constant with age. Constant age-specific growth rates imply that the age composition, or proportionate age distribution, of the population remains unchanged.

5.9 Replacement Level Fertility

Replacement Fertility Level is a particular value of Total Fertility Rate (TFR) which implies that a generation will exactly replace itself assuming the population closed to migration. According to United Nations for Replacement level fertility the value of TFR is ideally taken as 2.1. It is the fertility rate at which a population exactly replaces itself from one generation to the next. It ensures that each new generation is large enough to replace the previous one, accounting for fertility and mortality rates and ensuring the population remains stable in the long term. It can also be explained as the level of fertility at which a cohort of women on an average are having only enough daughters to 'replace' themselves in the population. For exact replacement the value of NRR should be equal to 1, meaning each woman, on average, replaces herself with exactly one daughter who survives to the age of childbearing. Considering the importance of female and male births for reproduction and balance of sex ratio, TFR therefore is taken to be 2.1 for exact replacement of a complete generation.

5.10 Population Momentum

Keyfitz (1971) has provided a significant application of the stable population model by focusing on population size rather than age composition or vital rates. He studied the effect of immediate reduction in fertility rates to replacement level (NRR=1.00) and maintaining it till the new stable equilibrium is attained. In order to achieve replacement-level fertility, all age-specific fertility rates need to be reduced proportionally by a factor 1/NRR, where NRR stands for the net reproduction rate before the decline. Even after making this adjustment, populations would still continue to grow, often significantly. This finding had a significant impact on policy discussions by highlighting how challenging it would be to control population growth even after fertility rates have dropped to replacement levels. This tendency for population growth to continue beyond the time at which replacement level fertility had been achieved due to a relatively high concentration of people in the childbearing years is known as population momentum or momentum of population growth. Populations with a high proportion of individuals in childbearing ages will have higher momentum, meaning they will continue to grow even if fertility rates decrease and populations with higher proportion of older individuals or already low fertility rates will exhibit lower momentum.

The relative amount of momentum is usually measured by the ratio of the size of the long-run stationary population to that of the population when replacement fertility is first achieved (Nathan Keyfitz, 1971). Using Lotka’s integral equation for a population with constant vital rates, Keyfitz showed that if replacement fertility rates, $m^*(a)=m(a)/NRR$ are applied to a closed population with an age distribution c(a) and survival function p(a), the annual number of births in the resulting stationary population will be

$$B_s = \frac{\int_0^\beta N(a) \int_\alpha^\beta \frac{p(y)}{p(a)} m^*(y) dy\, da}{A^*}$$

Where A^* is the mean age at birth in stationary population.

The above equation can be written as

$$B_s = \int_0^\beta \frac{N(a)}{p(a)} w(a) da$$

Where,

$$w(a) = \frac{\int_\alpha^\beta p(y) m^*(y) dy}{A^*}$$

Then we can write the eventual size of the stationary population as

$$N_s = B_s . e_0^0 = \int_0^\beta \frac{N(a)}{p(a)} w(a) da . e_0^0$$

Preston and Guillot (1997) gave the expression for population momentum by dividing the eventual size of the stationary population N_s by the initial population size when replace level fertility was achieved, which can be written as

$$M = \frac{N_s}{N} = \int_0^\beta \frac{N(a)}{N} \frac{e_0^0}{p(a)} w(a) da$$

or

$$M = \int_0^\beta \frac{c(a)}{c_s(a)} w(a) da$$

Where $c_s(a)$ is the age structure at age a in stationary population.

The above equation consists of three distributions, each summing to 1.000 across age intervals. The first one is c(a) which represents the proportionate age distribution of the population at the time replacement-level fertility is introduced. The second distribution $c_s(a)$ represents the proportionate age distribution of the stationary population that will eventually form after many years of replacement-level fertility and the third distribution w(a) represents the expected number of lifetime births that will occur from age a onwards under the replacement-level fertility regime is similar across different populations. It starts high at young ages and gradually drops to zero by age 50.

The equation helps us understand population momentum, which is how a population's growth continues even after fertility rates drop to replacement levels. If the current age distribution has more young people compared to the future stable population, momentum will be greater than 1.00, meaning the population will keep growing. If there are fewer young people, momentum will be less than 1.00, and the population will shrink. If the current age distribution is already stable, momentum will be exactly 1.00. This explains the reason for developing countries having a momentum factor above 1.00 as they have a higher proportion of young people, leading to continued population growth. On the other hand, stable populations with negative growth rates will have fewer young people compared to the future stable population.

Derivation of Momentum of Population Growth as given by Keyfitz

The momentum of population growth refers to the tendency of a population to continue growing even if fertility rates decline, due to the existing age structure

If the age specific fertility rate m(x) be replaced by $m(x)/R_0$, such that new NRR, say R_0^*, becomes unity, Keyfitz (1971) has derived a formula for the ultimate population size as :

$$\frac{B e_0}{\mu r}\left(1-\frac{1}{R_0}\right) \qquad (5.25)$$

where B : initial annual number of births at present.

e_0 : expectation of life at birth.

μ : Mean age of child bearing in the stationary population.

R_0 : net reproduction rate at present.

and r : annual rate of natural increase.

In brief, derivation given by Keyfitz is as follows :

Keyfitz (1971) has shown that the coefficient in the exponential birth function is

$$Q_j = \frac{\int_0^{\alpha} \exp(-r_j t)\, G(t)\, dt}{\int_a^{\beta} x \exp(-r_j x)\, \varphi(x)\, dx} \qquad j = 0, 1 \ldots \qquad (5.26)$$

Q_0 is called the stable equivalent of the observed births. It is the number of births that would become $Q_0 \exp(r_0 t)$ at time t, growing intrinsically at the rate of r_0. The number of women aged x to (x + dx) in the stable population is the number of survivors of those born x years ago is

$$Q_0 \exp(-rx)\, p(x)\, dx \qquad \text{(taking } r \text{ for } r_0\text{)}$$

Let $$Q = \int_0^{\infty} Q_0 \exp(-rx)\, p(x)\, dx \qquad (5.27)$$

Q is an index of future growth of population. Let us discuss the amount of its stationary equivalent. That is the number of births in the stationary equivalent that would be obtained when r is 0, and $\varphi(x)$ is replaced by $\frac{\varphi(x)}{R_0}$ so that new net reproduction rate may R_0^x become unity. The number of women of the stationary equivalent is obtained after multiplying by the survival function $\frac{l_x}{l_0}$ of the life-table, and integrating (26) over (0, ω), we get

$$Q = \frac{e_0 \int_0^{\beta} G(t)\, dt}{\int_{\alpha}^{\beta} x \frac{\varphi(x)}{R_0}\, dx} \qquad (5.28)$$

G(t) does not depend on r_0 as it is the number of births at time t, to those who were already alive at the beginning. In fact, using the arguments given earlier, we can see that

$$G(t)= \int_{\alpha-t}^{\beta-t} l(x)\, \frac{p(x+t)}{p(x)}\, \frac{m(x+t)}{R_0}\, dx \qquad (5.29)$$

Where p(x) is the number of women at age x.

Substituting (5.29) in (5.28) and solving the double integral, we get

$$Q' = \frac{e_0}{\mu} \int_0^{\beta} \int_{\alpha-t}^{\beta-t} l(x)\, \frac{p(x+t)}{p(x)}\, \frac{m(x+t)}{R_0}\, dx \qquad (5.30)$$

Since the denominator of (5.26) is the mean age of the stationary equivalent denoted by μ, in order to differentiate it from g.

Now $$l(x) = B\exp(-rx)\, p(x) \qquad (5.31)$$

Then $$Q = \frac{B}{R_0} \frac{e_0}{\mu} \int_0^{\infty} \int_0^{\infty} \exp(-rx)\, p(x+t)\, m(x+t)\, dx\, dt \qquad (5.32)$$

The change in limits do not affect the value as m(x) is zero outside (α, β).

Put x = x and x+t = y so that dx = dx and dt = dy.

The limits of integration now change to (0, ∞) and (x, ∞).

Now equation (5.30) becomes

$$Q' = \frac{B}{R_0} \frac{e_0}{\mu} \left[\int_0^{\infty} \int_x^{\infty} \exp(-rx)\, p(y)\, m(y)\, dy\, dx \right]$$

$$= \frac{B}{R_0} \frac{e_0}{\mu} \int_0^{\infty} \int_x^{\infty} \exp(-rx)\, p(y)\, m(y)\, dy\, dx - \int_0^{\infty} \int_0^{y} \exp(-rx)\, p(y)\, m(y)\, dy\, dx \;]$$

$$= \frac{B}{R_0}\frac{e_0}{\mu}\left[\frac{R_0}{r} - \int_0^\infty \exp(-rx)\,\varphi(x)\,dx\right]$$

where $\frac{d\varphi(x)}{dx}$ p(y) f(y) .

Integrating by parts the second integral, we have

$$\int_0^\infty \varphi(x)\frac{d\exp(-rx)}{-r} = \frac{\varphi(x)}{-r}\exp(-rx)\Big|_0^\infty + \frac{1}{r}\Big|\int_0^\infty \exp(-rx)\varphi(x)\,d(x) = 0 + \frac{1}{r_0}$$

$$(\because \int_0^\infty e - rx\,\phi(x)dx = 1)$$

Therefore, $$Q' = \frac{B}{R_0}\frac{e_0}{\mu}\left(\frac{R_0 - 1}{r}\right) \qquad (5.33)$$

If P is the size of the present population, we have an expression for the momentum as

$$\frac{Q'}{P} = \frac{b}{r}\frac{e_0}{\mu}\left(\frac{R_0 - 1}{R_0}\right) \qquad (5.34)$$

where $b = \frac{B}{P}$, the birth rate.

Keeping this into mind Ryder (1975) has given an adjustment to Keyfitz formula (5.34) and according to unrealistic due population momentum.

$$\frac{b e_0}{\mu r}\left(1 - \frac{1}{R_0}\right) e^{rA/2}$$

where A is the time taken by the existing population to come at the replacement level.

5.8 Summary

The present chapter explains estimation of the natural growth rate of a stable population through different methods as per the conditions and availability of data. It also gives a clear understanding of generational differences in populations through the mean length of generation and relationships among different averages or mean ages in stable population. Some other important concepts such as reproductive value, population momentum etc. and their applicability in real life conditions have also been explained for further understanding of concepts related to stable population or population growth in general. The concept of Replacement fertility level has been introduced to understand TFR and NRR clearly.

5.9 Self-Assessment Exercises

1. Derive the expression for intrinsic growth rate in Lotka's fundamental equation using method of moments and length of generation method. Also explain the relation between NRR and intrinsic growth rate r.
2. Give derivation for the intrinsic rate of growth r using Functional Iteration method. Also give an indirect estimation for *r* when the age composition data from one census is available.
3. Explain the concept of Mean Length of generation T. Derive the expression for T using mean age of child bearing in Stable population.
4. Define Reproductive value in Stable population and write the expression for Fisher's Reproductive value.
5. What do you understand by Stable Equivalent population and Replacement level fertility?
6. What is Population Momentum in a stable population. write expression for population momentum as Given by Preston(1977).

5.10 References

- Samuel H. Preston, Patrick Heuveline, Michel Guillot (2001). Measuring and Modeling Population Processes.
- United Nations, (1968) *The Concept of a Stable Population- Application to the study of Population of countries with Incomplete demographic statistics,* Population studies, No.39, New York.

5.11 Further Readings

- Nathan Keyfitz, Hal Caswell (1977). Applied Mathematical Demography.Samuel H. Preston, Patrick Heuveline, Michel Guillot (2001). Measuring and Modeling Population Processes.

- Lopez, A., (1961) Problems in stable population Theory, Office of the population Resarch, Princenton University, Princenton.

U.P. Rajarshi Tandon Open
University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N Demography

Block: 3 Fertility and Fertility Models

Unit – 6 : Fertility and its Measures

Unit – 7 : Cohort Measures and Indirect Estimation of Fertility

Unit – 8 : Fertility Models

Course Design Committee

Dr. Ashutosh Gupta **Chairman**
Director, School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

Prof. Anup Chaturvedi **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. S. Lalitha **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. Himanshu Pandey **Member**
Department of Statistics
D. D. U. Gorakhpur University, Gorakhpur.

Prof. Shruti **Member-Secretary**
Professor, School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Course Preparation Committee

Dr. Anupma Singh **Writer**
Department of Statistics
University of Lucknow, Lucknow

Prof. Shruti **Editor**
School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Prof. Shruti **Course Coordinator**
School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N **DEMOGRAPHY**

First Edition: ***July 2024***
ISBN: 978-93-48987-49-5

Printed and Published by Col. Vinay Kumar, Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2024.

Block & Units Introduction

The present SLM on ***Demography*** consists of Ten units with Four blocks. This is the third block of this SLM.

Block 3- ***Fertility and Fertility Models*** is the third block of said SLM and it is divided into three units.

Unit 6: ***Fertility and its Measures****,* is the first unit of the Self Learning Material which explains the concept of fertility in detail. It also covers different period measures of fertility and reproduction such as, Crude birth rate (CBR), General fertility rate (GFR) Age- specific fertility rate (ASFR), Total fertility rate (TFR), Gross reproduction rate (GRR), etc.

Unit 7: ***Cohort Measures and Indirect Estimation of Fertility,*** is the second unit of the Self Learning Material which focuses on different cohort measures of fertility, Use of birth order statistics, child women ratio, own-children method, children ever born (CEB) data and with data on current fertility, Brass P/F ratio for adjusting fertility rates.

Unit 8: ***Fertility Models,*** *is the third and the last unit* of the Self Learning Material which explains indirect estimation of fertility. This unit talks about modelling of fertility and studying some important probability models on time of first birth/conception and number of births/conception n specified time, birth interval models, study of fertility through birth interval analysis.

At the end of every unit the summary, self-assessment questions and further readings are given.

UNIT- 6 FERTILITY AND ITS MEASURES

Structure

6.1 Introduction

Fertility is one of the major components of demography which is responsible for change in population size and the structure. In this chapter, we are going to study the basic concepts, relative definitions and various measures used to study fertility and reproduction. The measures used to study fertility and migration are crucial for biological replacement of populations and for the survival of the human establishment. Population growth is largely an outcome of a surplus or deficit of births over deaths. Theory of regulation says that a society balances fertility according to the mortality in a population. The rapid growth of populations of developing nations since mid-19^{th} century is a result of decline in mortality. In order to reduce population growth it is important to reduce fertility and mortality both but mortality can only be reduced at a certain level as no one will ever suggest an increase in the death rate in order to develop. Therefore, a sensible approach to control a rapidly growing population is to reduce or regulate the fertility as per desire.

The term **fertility** refers to the actual birth performance of a couple. It is affected by a number of socio-economic, cultural and psychological factors on one hand and number of biological factors on the other and the mechanism through which they influence fertility is quite complex. There are some factors that relate to family formation and planning such as intended family size and age of first child, access to contraception, unwanted fertility, gender preferences, success with previous offspring and other factors known as the proximate determinants of fertility (Bongaarts 1978). It is also true that socio-economic and other factors regulate fertility through the biological factors. The main biological factors are: (i) fecundability, (ii) incidence of sterility, (iii) chances of conception being terminated in a foetal death or in a live birth and (iv) the period of temporary sterility comprising the duration of gestation and the post-partum amenorrhoea, associated with a pregnancy termination. The first two components are influenced the physiological and other characteristics of both partners, while the rest are concerned with female only. These factors may vary from female to female in a population and for a female, they may depend on time, age, duration of marriage, number of pregnancies or live births already experienced and such other maternal characteristics.

6.1.1 Conceptual Definitions

Fecundability- The fecundability is defined as the probability of conception in one unit of time of exposure when both the partners are biologically fit. Generally, the unit is taken as one lunar month which is the length of a menstrual cycle. Since the physiological status of both the partners as well as frequency of intercourse determines the value of this parameter, it is obviously considered as being couple dependent.

Fecundity- The physiological capacity of a woman (or a couple) to reproduce a child. It is

Sterility- The term sterility relates to the inability of a couple to procreate. A sterile couple is necessarily infertile but the converse is not always true. Even when a couple is fecund, it is possible that there is no conception, especially if fecundability is low. **There are two types of sterility: primary and secondary.** A female is primarily sterile if and only if she cannot change her status from parity zero to one. (the term 'parity' will be explained in the next unit.)She is said to be secondary sterile if she becomes sterile after one or more births. Sterility may be attributed to the husband or to the wife or to both of them. Thus like fecundability, the sterility is also couple dependent.

Foetal Wastage- A conception may not always result in a live birth. The outcome of a pregnancy that ends in a spontaneous or induced abortion or a still birth is known as foetal wastage.

Non-Susceptible Period- After getting conception the menstruation cycle of women stops and this conception results into a live or still birth, generally it is of nine months and known as gestation period. After termination of pregnancy women waits for some more times to resumption of the menstruation, this duration is known as post-partum amenorrhoea period. The sum of gestation and post-partum amenorrhoea period is called as non-susceptible period. In this period women cannot conceive further. It is the duration of gestation and the interval after its termination and before the resumption of ovulation which is known as post-partum amenorrhoea period.

Parity and Birth Order- The number of children ever born to a woman is referred to her parity. i.e. two women are those who have had two children ever born and zero-parity women who have had no children. When the total number of births at a particular point of time is classified according to their occurrence or order, the rank or order of the birth is called birth order of the particular child. Birth order one means first child and birth order three means third child. Parity starts from zero and used for women only however birth order status from one and used for child only.

Fertility Schedule - Fertility refers to the actual number of live births that women have. It differs from fecundity as mentioned above the physiological capacity to bear children which is normally begins after the onset of menarche provided that the female get married and stops with the episode of menopause. This period is known as reproductive period or the childbearing period.

Thus, fertility starts at about age 15 years (following menarche), attains a broad maximum at about age 20-25 and after that falls slightly by age about 30 years and then decline rapidly after age 35. A typical mean age at birth of last child among married women not practicing birth control is about 40 years. Only a small minority can bear children after age 45 and practically none after age 50 (following menopause). Therefore, commonly the age group 15-49 is considered as reproductive period.

The actual curve of child bearing by a female is a function of their age. Of course, this depends not only on the bearing capacity to conceive and bear children, but on variations with age in exposure to intercourse with a fertile partner, and whether or not measures are taken in different degree at different ages to prevent conception or to cause an early termination of pregnancies. The rise of fertility with age is strongly affected by taboos and customs that determine when women enter fertile unions. Among traditional societies in which marriage is ordinarily a pre-requisite for fruitful intercourse, there are wide variations in mean age at marriage. It is seen from even less than 15 years to nearly 30-35 years. The decline in fertility by age is influenced by the rising incidence of widowhood, divorce, and perhaps abstinence, as well as by declining fecundability. The practice of contraception and abortion could in principle reduce fertility at any age in the fertile span and thus produce as essentially arbitrary modification of the age structure of fertility. The combine effect of variation in fecundability, cohabitation rates, pregnancy wastage and contraception with age is to produce age schedule of child bearing. Thus, age schedule of fertility in human population has a number of general features in common. All rise smoothly from zero to an age in the teens to a single peak in the twenties or thirties and then fall continuously to near zero in the forties and to zero near fifties. Thus, **natural fertility** refers to the fertility which exists without any hindrance in the reproduction period i.e. in the absence of deliberate birth control while controlled fertility is the fertility which is obtained after some deliberate use of birth control.

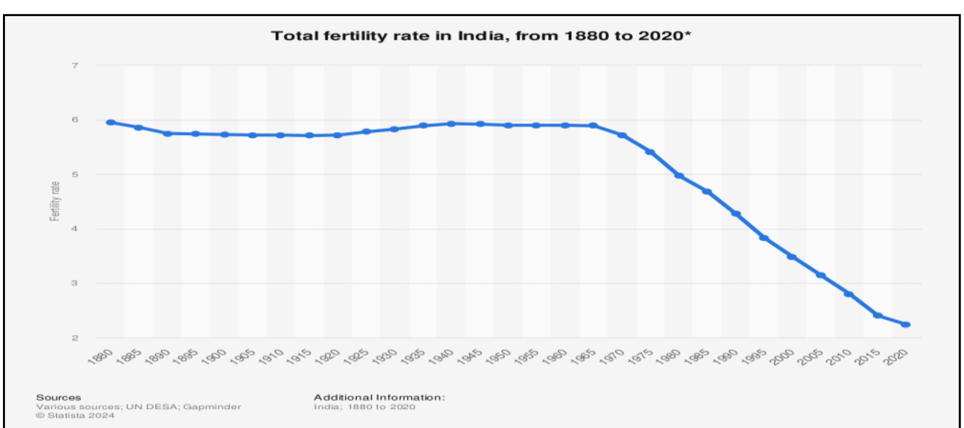
6.1.2 Factors Affecting Human Fertility

Human fertility is affected by a number of socio- economic and cultural factors on one side and a number of biological factors on the other side. It is also true that various socio- economic factors affect human fertility through biological factors. Among sociological factors we have the age at marriage, the coital pregnancy depending upon the level of various religious taboos and practices of abstinence. Coming to the biological factors we have fecundability which is the probability of a conception in one menstrual to a healthy and fecund woman leading a married life without use of a contraception. The other factors are the probability of a conception terminating in a foetal loss or live birth and the postpartum non-susceptible period associated with a live birth or a foetal loss and the probability of becoming sterile at various ages. Most of these factors vary from female to female and these may again vary with the age and/or parity. Various determinants of fertility and their interrelationships has been discussed in earlier Units.

6.1.3 Levels of Fertility in India

Global fertility trends have seen significant changes over the past century. Historically, most countries had high fertility rates, with women having several children. However, in recent decades, there has been a widespread decline, especially in developed nations, due to factors like improved healthcare, education, and economic growth. Many countries now have fertility rates below the replacement level of 2.1 children per woman, leading to concerns about aging populations. Conversely, some developing regions still experience higher fertility rates, although these are gradually decreasing as living standards improve. Fertility levels in India have undergone significant changes over the past century. In India, from 1880 to 1970, the fertility rate was quite high and consistent, with women typically having between 5.7 and six children over their lifetimes (figure 1.1). However, since the 1950s, there has been a steady decline, driven by factors such as increased access to education, improved healthcare, and family planning initiatives. By the early 2000s, the fertility rate had dropped to around three children per woman, and in recent years, it has approached the replacement level of 2.1. This decline reflects broader socio-economic developments and is contributing to changes in India's demographic structure.

Figure 1.1: Levels of Fertility in India (1880-2020)



6.2 Objectives

After completion of this unit, you should be able to:

- Understand the concept of fertility and reproduction, about the differences between fertility and reproduction
- Understand various measures of fertility calculated for different type of age groups and populations.
- Calculate measures of fertility and reproduction for different population

6.3 Measures of Fertility

Analysis of fertility requires good quality data on births and population. These data are conventionally obtained from vital registration systems, population censuses and sample surveys. The crude birth rate and the specific birth rates (generally known as the fertility rates specific for age of woman, marital status of woman, parity of woman, birth order of child etc.) are simple measures of fertility. Based on the specific fertility rates, the measures of reproduction are calculated. Some measures of fertility are based on the census or survey data alone. Fertility studies involving effect of contraception, post-partum abstinence, and breastfeeding on fertility, and birth interval dynamics make use of many other population measures. Fertility data is analysed in two ways: Period fertility and cohort fertility.

i. **Period Measures of Fertility-** In this approach, fertility is computed related to a particular period based on data on the number of births for the referred period. Period measures are usually calculated for one calendar year.

ii. **Cohort Measures of Fertility-** In this approach; fertility is computed indirectly on basis of data on the number of children ever born to women up to a certain point of time in the reproductive age group. Cohort measures are derived from the experiences of women born or married in a particular calendar year.

Various measures of fertility discussed in this unit are period measures, as period measures are more commonly used measures to represent fertility of a population. They are also easier to define and understand. Different measures of fertility based on period approach are discussed below:

6.3.1 Crude Birth Rate

The crude birth rate is calculated by dividing the number of registered births in a year by the mid-year population for the same year. The rate is expressed per 1,000 population.

$$\text{Crude Birth Rate} = \frac{\textit{Live births registered in a year}}{\textit{Mid} - \textit{year total population for the same year}} X\ 1000$$

Sometimes, CBR is computed covering data for two or three years to account variability in the rates as follows:

Year	Y_{t-1}	Y_t	Y_{t+1}
Birth	B_{t-1}	B_t	B_{t+1}
Population	P_{t-1}	P_t	P_{t+1}

Then the CBR can be computed in two ways.

i. CBR = Average of all these three CBRs

$$\frac{1}{3}\left[\frac{B_{t-1}}{P_{t-1}} + \frac{B_t}{P_t} + \frac{B_{t+1}}{P_{t+1}}\right] * K$$

or

ii. CBR = $\left[\frac{B_{t-1} + B_t + B_{t+1}}{3P_t}\right] * K$

Where, P_t represents the mid-year population

Estimation of Mid-Year Population from Census Data can be computed as:

i. $P = P_1 + \frac{n}{N}(P_2 - P_1)$

ii. If the interval between the two censuses is wide, then a compound rate of growth is used to compute $P_2 = P_1(1+r)^t$.

Where,

P is the estimated population;

P_1 is the population at the first census;

P_2 is the population at second census;

N is the number of years between two censuses and,

n is the number of years between the date of first census and the date of estimation.

Example 1.1 Let for a population of 1,000,000 in 2024 the total number of births is 15,000. The crude birth rate (CBR) is calculated using the formula:

$$CBR = \frac{\textbf{Total number of births}}{\textit{Total mid year population}} X1000$$

For the given data

CBR=(15,000/1,000,000)×1000

Therefore, CBR for the given population is 15 births per 1,000 people.

This Crude Birth Rate is, however, not an adequate measure of fertility, as it is calculated without any consideration of the age and sex composition of the community.

For one thing, it cannot be called a probability rate, since the whole population cannot be supposed to be at the risk of experiencing the particular type of vital event we are considering here. Only females and only those between certain ages are really liable to the risk of giving birth to a child. Among such females, again, the risk varies from one age-group to another: a woman of 25 is certainly under a greater risk than a woman of 40.

The crude birth rate (CBR) is often regarded as an wage measure of overall population changes due to births, and it is straightforward in both concept and calculation. It’s important to note that each birth is counted directly as an addition to the population. This causes the denominator to increase, meaning the rate doesn’t accurately reflect the true likelihood of childbearing. Due to this bias, the CBR tends to understate changes in fertility. The term "crude" is used because the denominator includes individuals of all ages and both sexes, even though only women within certain age ranges can have children. Therefore, when calculating the CBR, we don't specifically account for the population at risk of giving birth. This measure is significantly influenced by the population's age and sex composition, as well as other characteristics. Comparing two populations using this measure can be misleading, as their age-sex compositions might differ substantially.

6.3.2 General Fertility Rate

The general fertility rate is calculated by diving the number of registered births in a year by the mid-year population of females aged 15-44 years for the same year. The rate is expressed as per 1,000 female’s population in reproductive ages.

$$General\ fertility\ rate\ = \frac{Live\ births\ registered\ in\ a\ year}{Mid\ year\ female\ population\ ageing\ 15-49}\ X\ 1000$$

The formula for the GFR is, thus,

$$i = 1000 * \frac{B}{\sum_{w_1}^{w_2} fP_x}$$

Where i = general fertility rate per 1,000 females in child-bearing ages;

B = number of live births in the given region during the given period;

${}^{f}P_x$ = number of females of age x, 1.b.d. in the given region during the given period; and

w_1, w_2 = lower and upper limits of the female reproductive span.

The computation of the GFR requires that a decision be taken before hand as to which years of life of a woman should be included in the child-bearing (or reproductive) span. Although the practice varies in this respect, the generally adopted method is to take $w_2 = 49$. Births to mothers under 15 and above 49 are so rare that they are not recorded separately but are included in the age-groups 15 and 49, respectively.

However, the disadvantages of GFR are that all unmarried women of all ages are included in the denominator and each woman has equal risk of conception at all ages during the reproductive period.

GFR is further refined by taking only legitimate births to married women. It is then called General Marital Fertility Rate (GMFR) defined as:

$$GMFR = \frac{Legitimate\ births}{Married\ \ female\ \ population} * K$$

Example 2: Let us assume a hypothetical population given in the table 1.1

Table 1.1 (Distribution of Births and Female Population of a Population)

Age Group (1)	**Number of Births (2)**	**Female Population (3)**
15-19	36700	1499930
20-24	224678	1543524

25-29	198567	1479863
30-34	65345	1387594
35-39	15087	1318762
40-44	2341	1023456
45-49	198	990735
Total	**542916**	**9243864**
Total Population		**31845398**

The CBR of the population is given by the ratio of the total number of births and Total population multiplied by 1000.

Therefore, CBR is (Total (2)/31845398)*1000 = 17.1 per 1000 people

GFR= (Total number of births/ Total female population in reproductive span)*1000

= (542916/9243864)*1000

= 58.7 per 1000 women

6.3.3 Age Specific Fertility Rate

The age-specific fertility rates are calculated from births and female population both specific to each age (or age group) or woman. Thus,

$$Age-specific\ fertility\, rate = \frac{Live\, births\ women\, aged\ (x, x+n)}{Mid-year\ female\ population\, aged\ (x, x+n)} *1000$$

Where `*x*' is the age of the female and '*n*' the class interval of age.

Thus, the specific fertility rate for the age-group x to $x + n - 1$ is given as-

$$ {}_n i_x = \frac{{}_n B_x}{f_n P_x} * 1000 $$

Where ${}_nB_x$ indicates number of births to women of age x to $x + n$ in the given region during the period and

f_nP_x indicates the number of women of age x to $x + n$ in the region during same given period.

In the case of an annual age specific fertility rate, $n = 1$ and here one write simply

$$ i_x = \frac{B_x}{fP_x} * 1000 $$

All the three rates can be made 'more' specific to various groups of women. Usually, the numerator is made 'more' specific while keeping the denominator unchanged. Thus, births could be classified by sex, birth order of child, legitimacy status of the child – the idea being that the total of 'more' specific fertility rates is the specific fertility rate. If data permit, both the numerator and the denominator of the rates are made 'more' specific. Age-parity specific fertility rates (APSFR) qualify this category. Thus,

$$ APSFR \text{ for parity } 1 \text{ of women at age } x = \frac{\text{Number of births occured in a year to women of parity } 1 \text{ at age } x}{\text{Mid year population of parity } 1 \text{ at age } x} * 1000 $$

The quotient is multiplied by 1,000. Women are classified by parity to indicate their past number of children ever born before the birth of the current child. Children are classified by the birth order.

Example 3: Consider the data given in table 1.1. To calculate the ASFR for the given population we can simply take the ratio of column no. 2 and column no. 3 and multiply by 1000, as follows:

Table 1.2 (Age Specific Fertility Rate per 1000 women)

Age Group (1)	**Number of Births**	**Female Population**	**ASFR (4)**

	(2)	(3)	(per 1000 women) (2/3*1000)
15-19	36700	1499930	24.5
20-24	224678	1543524	145.6
25-29	198567	1479863	134.2
30-34	65345	1387594	47.1
35-39	15087	1318762	11.4
40-44	2341	1023456	2.3
45-49	198	990735	0.2
Total	542916	9243864	365.2

Column no. (4) of Table no.1.2 represents the age specific fertility rate of the population. It is clear from the table that the highest fertility for the given population is seen in the age group 20-24 and 25-29, which, in general is considered the most fertile window of the female reproductive span. Also, we can see that the number of births are decreasing with age, is observed to be the least in the age group 45-49 which is almost the menopausal age.

The Crude Birth Rate, General Fertility Rates and the Specific Fertility Rates are also called as Cross-sectional Measures of fertility as they measure the fertility horizontally with respect to time.

6.3.4 Total Fertility Rate

Mean or median age of the fertility schedule (a set of the age-specific fertility rates from the minimum age to the upper age of the woman's reproductive life span) is calculated from the age-specific fertility rates at single- or five-year age groups of women. Note that these are better measures of the average age of mother giving a birth as these take into account the age distribution of the female population in the year.

Total fertility rate is a summary measure of the age-specific fertility rates over the reproductive span, i.e. 15 to 49 or 15 to 44. It is calculated as the sum of the age-specific fertility if ages rates are in single years of age. If the age intervals in the calculation of ASFRs is greater than one, then they represent the average of the interval and it is required that they should be multiplied by the width of the interval. Therefore, if the rates are for each five year age group of woman, the sum of the rates is multiplied by 5, on the assumption that the five year age group rate will hold at each single age within the age interval of five years. The total fertility rate can be calculated for each year (or period) or for birth cohorts. The formulae are as follows:

Total fertility rate = $\sum_x f_x$ *where, f_x is the age specific fertility rate at age x*

Or Total fertility rate = $5*\sum_x f_x$; if the width of the age interval is taken at 5.

In general, we can write Total Fertility Rate as:

$TFR = n\sum_x f_x$; where n is the width of the age interval.

The value of x will go from 15 to 44 or 15 to 49.

If S_x is proportion of females not married or widow i.e. not susceptible to the risk of conception during the reproductive period, then Total Marital Fertility Rate (TMFR) is approximated as-

$$TMFR = TFR + B_x \frac{S_x}{P_x}$$

If S_x is same at all ages during the reproductive period then TMFR becomes,

$$TMFR=(1 + S)\ TFR$$

But S_x follows U-shaped distribution i.e. at beginning and end of the reproductive period it remains high due unmarried and menopausal reasons respectively.

Total Fertility Rate is the most popular measure of fertility used by the demographers. It is generally expressed as the number of children per woman. It implies the number of children ever born to a woman if she has survived through her reproductive span subjected to fertility schedule described earlier in this chapter. TFR is an important measure as it is a single figure and is independent of the age structure. However, the computation of TFR requires a lot of data, like, number of births by age of the mother and age of the women.

Note: If ASFR is used to calculate the TFR, then the rate comes out to be per thousand, therefore it must be divided by 1000 to be represented in a proper way.

Example 4: For the data given in table no. 1.1TFR can easily be calculated by calculating the ASFRs. Table no. 1.2 shows the ASFRs for respective age groups. Hence, TFR can be given as ($5*\sum_x f_x$)/1000 , since the age interval is 5.

TFR= (5*365.2)/1000

= 1.8 per woman

This shows that, for the given population a female will give birth two approximately 1.8 children in her life.

6.3.5 Child-Women Ratio (CWR)

The ratio of children under age 5 to women of child-bearing age is taken as an index of fertility derived from the age-sex distribution of population. Commonly, we have the following CWR:

$$\text{CWR}= \frac{C(0-4)}{W(15-44)} \times 1000,$$

if lower and upper limits of reproductive period be taken as 15 and 44 respectively.

where C stands for number of children and W for number of women.

The great advantage of the CWR is that it does not require a special question

However, CWR has the following weaknesses

(i) It is directly affected by the under-enumeration of young children.

(ii) Mortality affects both women of child

In analysis of fertility, the cohort approach has gained additional relevance from the fact the total fertility is a cumulative process and that a woman's past birth history may affect her future fertility. Various cohort measures of fertility will be discussed in unit II in details.

6.4 Reproduction

The Total Fertility Rate gives the estimate of lifetime fertility. Generally, the TFR is expressed as the average number of births per woman. The total fertility rate, calculated per woman, gives a measure of reproductivity of both sexes of children. A total fertility rate of 1.8 children per woman implies that if women continue to reproduce at the current level of the age-specific rates (which gives a total fertility rate of 1.8), she will give birth to 1.8 children. As the sex ratio is usually in favour of males (it varies from population to population, let’s assume that it can be taken as 105 male births to 100 female births), only 0.88 female babies would be born. Thus, one female will be replaced by 0.88 female in the 'long run'. Therefore, to understand the growth of the population better, study of reproduction and its measures is required.

6.5 Measures of Reproduction

As we have stated earlier, it is important to study the part of population which is majorly responsible in the change of growth, i.e. females. In reproduction studies we study the reproductive measures which explore the rates of female births in particular. Reproduction is measured by two vital rates namely Gross Reproduction Rate and Net Reproduction Rate.

6.5.1 Gross Reproduction Rate

For a proper measure of population growth, it is necessary to take into account the age-sex composition of the population. It is also appropriate that we should take female births alone, since it is mainly females through which a population increases. In this case age-specific fertility rates will only be calculated for female births and can be written as-

$$f_{i_x} = \frac{f_{B_x}}{fP_x}$$

where f_{B_x} is the number of female births to women of age x during the given period in the given community.

Summing over these rates for all ages in the reproductive span, we get a measure of population growth, known a Gross Reproduction Rate (GRR) as,

$$GRR = \sum_{w_1}^{w_2} f_{i_x}$$

Therefore, the gross reproduction rate is the average number of daughters that would be born alive to a hypothetical cohort of women if they lived to the end of their reproductive years and if they experienced the same age-specific fertility throughout their lives that women in each age group experience in a given year or period of years.

By definition, GRR can be written as-

$$GRR = \frac{Total\ number\ of\ female\ births\ occuring\ to\ females\ in\ their\ reproductive\ span}{Total\ number\ of\ females\ in\ the\ reproductive\ span(15-49)} Xk$$

The gross reproduction rate is not affected by the age structure of the population because it is, in effect, an age-standardized fertility rate with each age given a weight of one. In addition, it can also be thought of as the ratio between female births in two successive generations assuming that there are no deaths before the end of the reproductive period, or it may be considered as the ratio between the number of females in one generation at a given age and the number of their daughters at the same age, assuming that there is no mortality during the child-bearing years.

GRR indicates the number of daughters who would be born, on the average, to each of a group of females beginning life together, assuming that none of them dies before reaching the end of the child-bearing period, if they experienced throughout this period the current level of fertility as represented by the series of rates

If the fertility rates are in 5 yearly age-group, viz.

$$f_5 i_x = \frac{f_5 B_x}{f_5 P_x}$$

then the GRR will be given by

$$GRR = 5 \times \sum_{w_1}^{w_2} f_5 i_x$$

In some cases, births may not be available according to mother's age and according to sex. Here the simple formula cannot be applied, however, an approximate value of the GRR can still be obtained if sex ratio at birth, i.e. the ratio of the number of male births to the number of female births is available. Here we shall have, approximately,

$\frac{f_{Bx}}{B_x} =$ a constant (say k) where f_{Bx} is females births and B_x is totalirths to females of age x

Then, $$k = \frac{\sum_{w_1}^{w_2} fB_x}{\sum_{w_1}^{w_2} B_x} = \frac{f_B}{B}$$

so that $$fB_x = B_x \times \frac{f_B}{B} \; and \qquad f_{i_x} = i_x \times \frac{f_B}{B}$$

Therefore, as estimate of the GRR will be

$$\frac{f_B}{B} \sum_{w_1}^{w_2} i_x$$

$\sum_{w_1}^{w_2} i_x$ is just the TFR except for the usual multiplier 1000

Thus, GRR is calculated as follows:

GRR = Total fertility rate*sex-ratio at births

or = $TFR * \frac{1}{1 + s_b}$

Where S_b is the sex-ratio at birth i.e. number of males per female.

Example 1.5: For some population, if, the sex ratio at birth is taken to be 105 males to 100 females. And for a year for which the TFR is approximately 3.54, the GRR will be-

$$\text{GRR}= TFR * \frac{1}{1+s_b}$$

$$= 3.54 \times \frac{100}{205} = 1.73 \text{ per woman}$$

1.73 GRR indicates that a women will give birth to 1.73 females or 2 females approximately throughout her reproductive life span. A gross reproduction rate of less than unity indicates a declining population in the future.

One limitation of the gross reproduction rate is that it assumes that all women who enter the reproductive age (say 15 years) will live to the end of their reproductive life (to age 49 for example). This assumption is not valid as some women die might not survive throughout their reproductive live span. The value of GRR will always come out to be great than TFR as TFR considers the births with respect to both sexes whereas GRR only considers female births.

6.5.2 Net Reproduction Rate

As we have mentioned above, Gross Reproduction Rate does not take cognizance of the fact that some of the females who are assumed to begin life together may die before reaching age 15, some may die between age 15 and 16, and so on. Or simply, a female can die any time during their reproductive span. That is the GRR takes into account current fertility but ignores current mortality at the same time.

Taking this fact into account the factor of mortality in measuring population growth, a life table for females on the basis of the observed age-specific death rates for females, f_{m_x} is constructed. The values in the L_x column of the table (denoted by f_{L_x} in this case) give the mean size of the cohort of f_{l_0} females in the age interval x to $x+1$ for varying x. Hence,

$$f_{l_o} \times f_{L_x}$$

gives the number of female children that would be born to the cohort at age ‘x’ l.b.d.

Summing the above expression over complete reproductive span, i.e, w_1 and w_2, which gives the lower limit and upper limit of the reproductive span respectively.

we get,

$$\sum_{w_1}^{w_2} f_{l_o} \times f_{L_x}$$

as the total number of female children that are expected to be born to the f_{l_0} female during their life time. Thus, a new measure of population growth is

$$\frac{1}{f_{l_0}} \sum_{w_1}^{w_2} fi_x \times f_{L_x}$$

The above expression is called as the **Net Reproduction Rate (NRR).** The NRR shows how many females would be born, on the average, per member of a group of females beginning life together, if they were observed rates of mortality and fertility throughout their life time.

In simple language NRR can be expressed as the following formula-

$$\frac{Total\ number\ of\ female\ births}{Total\ number\ females\ in\ their\ reproductive\ span(15-49\ years\ of\ age)}\ X\ Life\ table\ survival\ Ratio$$

Obviously, the NRR cannot be greater than the GRR. Thus, NRR indicates how many future mothers would be born to present mothers according to the current levels of fertility and mortality.

If the NRR = 1, then it may be said that current fertility and mortality are such that a group of newly born females will exactly replace itself in the next generation. In such the population may be said to have a tendency to remain constant in size. A net reproduction rate of one is roughly equivalent to a two-child family.

When the value of NRR >1, it implies that the population shows a tendency to increase, for in that case a group of females is expected to be replaced by a larger number of females in the next generation.

Similarly, a value of NRR <1, is an indicator of a population showing a tendency to increase. In such scenario, a group of females is expected to be replaced by a smaller number of females in the next generation.

In this way the NRR be looked upon as a good index of population growth.

The net reproduction rate is the average number of daughters that would be born alive to a hypothetical cohort of women if they experienced the same age-specific fertility throughout their lives that women in each age group experienced in a given year, or period of years, and if they were also subjected to the mortality rates of the same year or period of years. Although the net reproduction rate purports to describe the fertility and mortality experience of a generation of women, the rates presented in this table are actually based on the fertility and mortality reported or estimated for a given reference period, usually a single year or a five-year period. Net Reproduction rate is not affected by the age structure of the population, similar to GRR. However, it does differ from the GRR because it takes mortality into account.

The NRR can also be thought of as the ratio between female births in two successive generations taking mortality into account, or it may be considered as the ratio between the number of females in one generation at a given age and the number of their daughters at the same age, again taking mortality into account.

Example 1.6: Determination of Gross and Net Reproduction Rates for the given data

(1) Age in years	**(2)** Age-specific fertility rate	**(3)** Female life-table stationary population	**(4)** col.(2) x col. (3)
15-19	0.086	4285	368.51
20-24	0.237	4153	984.26
25-29	0.179	4060	726.74
30-34	0.118	4003	472.35
35-39	0.069	3934	271.45
40-44	0.035	3861	135.14
45-49	0.011	3783	41.61
Total	0.735	-	3000.06

The sex-ratio at birth for the country may be supposed to be 205 males to 100 females. Since GRR and NRR both only consider female births and the data given is ASFR for both the sexes, we multiply it with the sex ration at birth. Hence, from the above table, we get

$$GRR = 5 \times 0.735 \times \frac{100}{205} = 1.793 \qquad and \qquad NRR = \frac{3000.06}{1000} \times \frac{100}{205} = 1.46$$

Total fertility rate, Gross reproduction rates and Net reproduction rates are also called as synthetic measures of fertility as they provide estimates of the lifetime fertility of the females, under certain set of assumptions. Given a hypothetical group of 1,000 females on their 15th birthday, and assuming:

1. the group is closed to international migration
2. the group is subjected to the cross-sectional fertility rates at each age based on births occurring during a particular year during their reproductive period.
3. no female dies before completing her reproductive period at age 45.

The gross and net reproduction rates give an indication of the replacement of the current female population in future. NRR of less than 1 indicates the female population is not replacing itself. In other words, the size of the daughters' generation is smaller than the size of the mothers' generation.

Two points are worth noting:

- TFR > GRR > NRR
- GRR is approximately half the TFR.

ADDITIONAL EXAMPLES:

Example 1.7: In 2011, a city had a total of 507000 live births, while its total population was 27512000 and total female population in the age group 15-49 was 7576000. Obtain the crude birth rate and the general fertility rate.

Solution:

Crude Birth Rate= $\frac{507000}{27512000} \times 1000 = 8.4 \textit{ per thousand population}$

General fertility Rate= $\frac{507000}{7576000} \times 1000 = 66.9\ per\ thousand\ population$

Example 1.8: In a population there were 1560 and 1135 females in the age groups 20-24 years and 25-29 years respectively in the year 2007. There were 310 and 170 births to females of above two age groups respectively. Compute age specific fertility rates for the two age groups for the year 2007.

Solution:

Age Specific Fertility Rate for the age group 20-24

$= \frac{310}{1560} \times 1000 = 198.72\ per\ thousand\ females$

Age Specific Fertility Rate for the age group 25-29

$= \frac{170}{1135} \times 1000 = 149.78\ per\ thousand\ females$

Example 1.9: Methods for computing GRR and NRR.

Age Group	**Female Population**	**Live births**	**ASFRs (per women)**	**Female live births**	**Age-specific maternity rates** (f_x^w)	**Age-specific survival rates *** (e_0^0 **=68)** ($_5L_x/l_0$)	**Expected Female births pr women**
	(1)	**(2)**	**(3) = (2)/(1)**	**(4)**	**(5)=(4)/(1)**	**(6)**	**(7)= (5)x(6)**
15-19	77860	8888	0.114	4320	0.055	0.905	0.050
20-24	63512	18520	0.292	8623	0.136	0.900	0.122

25-29	50720	14456	0.285	7219	0.142	0.895	0.127
30-34	45610	11020	0.242	4833	0.106	0.888	0.094
35-39	38705	7645	0.198	3686	0.095	0.880	0.084
40-44	31947	2833	0.089	1476	0.046	0.870	0.040
45-49	25844	423	0.016	193	0.007	0.857	0.006
Total	-	63785	1.235	30350	0.588	-	0.524
TFR	-	-	6.175	-	-	-	-
GRR**	-	-	-	-	2.94	-	-
NRR	-	-	-	-	-	-	2.62

Probability of surviving from birth to the mid-point of the age group.

$$GRR = TFR \times \frac{1}{1+ Sex\,Ratio\,at\,birth}$$

$$= 6.175 \times 0.4758 = 2.94$$

or

or $$GRR = \sum_{x=15}^{45} 5\, f_x^{\,w} = (0.588)*5 = 2.94$$

$$NRR = \sum_{x=15}^{45} 5\, f_x^{\,w} \times \frac{{}_5L_x^{\,w}}{l_0} = 0.542*5 = 2.62$$

Now, NRR = 1 implies exact replacement,

\> 1 implies more than replacement

< 1 implies not replacing itself and a negative growth rate in the long run.

6.6 Self-Assessment Exercises

Exercise 1.1: The number of births occurring in a population in 2020 is shown here classified according to age of mother, together with the female population in each age-group of the child bearing period:

Age	Female population (in thousands)	Number of births to mothers in the age group
15-19	200	4000
20-24	173	26000
25-29	161	32000
30-34	160	23000
35-39	155	11000
40-44	125	2000
45-49	87	125
Total		

The total population was 4000500.

Determine:

(a) the crude birth rate,

(b) the general fertility rate,

(c) the age specific rates and

(d) the total fertility rate for 1988.

Exercise 1.2: Write whether the following statements are true or false:

(a) CBR is not very sensitive to small fertility change.

(b) Although CBR is affected by age composition of population and level of fertility but not by the age pattern of fertility.

(c) GFR is a more acceptable measure of fertility level because it controls the variations in age composition within the reproductive age range.

(d) ASFR's are widely affected by variations in population composition.

Ans.: (a) true, (b) false, (c) false (d) false

Exercise 1.3: The fertility rates (computed on the basis of female births alone) for Eastern Uttar Pradesh 2001 is shown in the following table, together with the survival factor for each 5 year age-group. Compute GRR and NRR on the basis of the given data.

Age-group	Fertility Rate (female births)	Survival Ratio
15-19	0.0108	0.968
20-24	0.0666	0.968
25-29	0.0675	0.964
30-34	0.0433	0.952
35-39	0.0226	0.951
40-44	0.0072	0.943
45-49	0.0009	0.928

Exercise 1.4: Fill up the blanks

(a) Computation of TFR is based on births of both sexes, whereas GRR is based on …………………only.

(b) Like TFR, GRR also assumes that women in reproductive age groups ………… till the end of their reproductive period.

(c) GRR in a population is 1.8 means if 100 mothers follow the current schedule of fertility, they will be replaced by……….. daughters.

(d) NRR is nothing but a refinement over GRR where the……… is introduced.

(e) Reproductive survival ratio is defined as the ratio of……………

(f) Fertility of replacement level corresponds to the value of N.R.R. = …………………..

(g) Gross reproduction rate cannot be ………. .net reproduction rate.

(h) General fertility rate is not able to pinpoint for………………

Ans.: *(a) female births, (b) survives (c) 180 (d) mortality (e)* ${}_5L_x^w / l_0$ *(f) 1 (g) less than (h) family planning*

Exercise 1.5: Explain GRR and NRR and show that NRR ≤ GRR. When GRR will be equal to NRR?

Exercise 1.6 For a country in particular year the total fertility rate is found to be 2.893 and the net reproduction rate is 1.123. Explain what exactly is meant by these figures.

Exercise 1.7: From the following table calculate (i) total fertility rate (ii) gross reproduction rate (iii) net reproduction rate. If it is known that the ratio of females to total births is 0.48.

Age-group	15-20	20-25	25-30	30-35	35-40	40-45	45-49
Birth rates per 1000 females	23.6	114.9	145.1	122.6	79.6	35.9	3.3
Year lived by 1000 live born female	4482	4426	4359	4288	4202	4100	3979

Exercise 1.8: Write whether the following statements are true or false :

(a) CBR is not very sensitive to small fertility change.

(b) Although CR is affected by age composition of population and level of fertility but not by the age pattern of fertility.

(c) The ratio of total number of yearly births to the female population in the reproductive age groups is called general fertility rate.

(d) GFR is a more acceptable measure of fertility level because it controls the variations in age composition within the reproductive age range.

(e) Distribution of fertility levels in the childbearing ages is best raveled by the computation of ASFRs.

(f) ASFRs are widely affected by variations in population composition.

(g) Computation of TFR of a hypothetical cohort accounts attrition due to death occurred within the reproductive period.

(h) TFR is also affected the by age structure of women under study.

Answers: (a)True,(b) False (c) True (d) False (e) True (f) False (g) False (h) False

Exercise 1.9: Fill in the blanks:

(a) A married woman having no live born child is said to be of parity ________________________________.

(b) Parity starts with zero, while, the birth order starts with__________________________________.

(c) The child-woman ratio always underestimates the current level of fertility because___________________________________.

Answer: (a) Zero (b) one (c) The survival rate is higher among women

Exercise 1.10: Differentiate between fertility and reproduction? Explain various measures of fertility?

6.7 Summary

After completion of this unit, one will be able to explain the phenomenon of fertility and reproduction. Fertility is one of the most important components of population growth, other than mortality and migration. But a detailed knowledge of measures of fertility is required to know which of the measures should be used in which condition. TFR is the most popular and readily used measure of fertility, whereas NRR is used to understand the current and future growth of the population. This chapter enables us to calculate various measures of fertility and reproduction for different population.

6.8 References

- Bogue, D. (1969). *Principles of Demography*, New York: John Wiley and Sons.
- Yusuf, F., Martins, J. M., Swanson, D. A. (2014) Techniques of demographic analysis, Springer
- Srinivasan, K. (2015). Training Manuals on Demographic Techniques, UNFPA, Delhi.
- UN DESA and STATISTICA 2024.

6.9 Further Readings

- Keyfiitz, N. (1968). *Introduction to the Mathematics of Population*, Reading, Mass: Addison-Wesley.
- Pathak, K. B. and Ram, F. (1998) Techniques of Demographic Analysis, Himalaya Publishing House
- Kanitkar, T. and Bhinde, A. A. (2011) Principles of Population Studies, Himalaya Publishing House

UNIT- 7 COHORT MEASURES AND INDIRECT ESTIMATION OF FERTILITY

Structure

7.1 Introduction

Fertility analysis is primarily conducted in two ways: through a period, perspective and a cohort perspective. The period perspective examines events within a specific timeframe (calendar years) in relation to the population's exposure during that time. In contrast, the cohort perspective focuses on events and exposure durations for specific cohorts as they progress over time. A 'cohort' refers to a group of individuals who share a common experience at the same time. In demography, two main types of cohorts are typically used: birth cohorts, consisting of those born in the same year or period, and marriage cohorts, consisting of those married in the same year or period. Period measures generally assess fertility rates in a cross-sectional manner, while cohort measures take a longitudinal approach. Period measures are usually calculated for one calendar year, whereas cohort measures are derived from the experiences of women born or married in a particular calendar year. Therefore, period measures reflect the combined experience of different cohorts of women in any given year. Understanding the relationship between cohort and period fertility can be complex, particularly since both perspectives are rarely analysed using the same data source. Period measures are simpler and more commonly used, which is why they are typically discussed first. We have discussed the Period measures of fertility in Unit one already. In this chapter we are going to discuss various measures of fertility of cohort data.

7.2 Objectives

After completion of this chapter, one will be able to:

- Understand the difference between period measures and cohort measures of fertility.
- Calculate various cohort measures defined for fertility for any given population.
- Understand the concept of birth interval.
- Understand the estimation of fertility through indirect methods.

7.3 Parity and Birth Order

Parity: The number of children ever born to a woman is referred to her **parity**. i.e. two-parity women are those who have had two children ever born in their life span and zero-parity women who have had no children.

Birth Order: When the total number of births at a particular point of time is classified according to their occurrence or order, the rank or order of the birth is called **birth order** of the particular child. Birth order one means first child and birth order three means third child. Parity starts from zero and used for women only however birth order status from one and used for child only.

A cohort's total fertility rate can be readily estimated through a census of survey question about parity, the number of live births a woman has had. The mean parity, or mean number of children ever born, of a cohort of women who have completed childbearing, is equal to the cohort's total fertility rate if reporting is accurate and if there are no differentials in mortality or migration by parity. The fertility process can be represented not only through a woman's movement from one age to the next but also by her movement from one parity to the next. This latter movement can be represented by parity progression ratios, introduced by Henry in 1953.

7.4 Cohort Measures of Fertility

Cumulative fertility is the one of the most common cohort measures of fertility, which shows the childbearing experience of a cohort from the beginning of exposure to the risk of conception to some later date. Cohort fertility is studied to find how many children a particular cohort (birth or marriage) contributes in its life-time. When all the members of a cohort have gone through the child bearing period or reproductive span then the cumulative fertility becomes completed generation fertility or family size. Every cohort experiences a different fertility pattern than the other. It is not easy or always possible to follow the number of births occurring to each woman at different ages of the reproductive span, but it is possible to find cohort measures based on current and period fertility rates available for past years.

7.4.1 Parity Progression Ratio (PPR)

The term parity progression ratio indicates the extent to which a woman progresses from one parity to the next. In a probabilistic sense, we can say that the parity progression ratio indicates the chance for a woman of parity "i" to go for higher parity (to have additional child) in a year or births.

The Parity Progression Ratio is defined as-

$$PPR_{(i,i+1)} = \frac{Number\ of\ women\ at\ parity\ i+1\ or\ more}{Number\ of\ women\ at\ parity\ i\ or\ more} = \frac{P_{i+1}}{P_i}$$

This cohort measure is usually calculated only for cohorts who have completed their childbearing. The cohort total fertility rate is retrospectively estimated as the total number of births amount to women in the cohort divided by the number of women in the cohort. If we denote P_i as the number of women at parity 'i' or more and W, as the total number of women, then the number of first births will equal P_1, of second births P_2, etc., and therefore TFR can be given as-

$$TFR^c = \frac{P_1}{W} + \frac{P_2}{W} + \frac{P_3}{W} + \ldots\ldots$$

$$= \frac{P_1}{W} + \frac{P_1}{W}.\frac{P_2}{P_1} + \frac{P_1}{W} + \frac{P_1}{W}.\frac{P_2}{P_1}.\frac{P_3}{P_2} + \ldots\ldots\ldots$$

$$= PPR_{(0\text{-}1)} + PPR_{(0,1)}.PPR_{(1.2)} + PPR_{(0.1)}\ .\ PPR_{(1.2)}\ .\ PPR_{(2.3)} + \ldots$$

Parity progression ratios are especially useful in studying the patterns of fertility-limiting behaviour in a particular population, which are often keyed to the number of children a woman has already born (Henry, 1961a; Feeney and Feng, 1993). Fertility limiting behaviour actually means the pattern of birth spacing and reduction of the number of births at various birth order.

The percentage proportion of women for a specific parity (one and over) can be obtained by taking difference between cumulative fertility rates for successive birth orders and dividing the difference by 10. The cumulative fertility rates for successive order will provide the number of women with parity n per 1000 women.

Therefore, we can write,

Percent of women at parity n

$$= \frac{Cumulative\ fertility\ rate\ of\ order\ n - Cumulative\ fertility\ rate\ of\ order\ n+1}{10}$$

The percent of women in the last parity is estimated by dividing the cumulative rate for the lowest parity included in the past parity group by 10.

7.4.2 Order Specific Fertility Rates

Different age specific cohort experiences different fertility pattern. Each cohort has a different pattern of spacing between births. Analysis of order specific birth rates for birth cohort can be used to study spacing between the births by taking differences in the median ages for the successive order of births.

It is defined as the number of births of the order "i" per 1000 women of child-bearing ages,

$$\frac{B_i}{W_{15-49}} \times 1000$$

Where W_{15-49} be the total number of women in the age group 15-49. It should be noted that the sum of OSFRs over all orders equals the GFR, i.e.,

$$\text{GFR} = \sum_{i=1}^{n} \left(\frac{B_i}{W_{15-49}} \times 1000 \right)$$

with "n" being the maximum order of birth.

7.4.3 Age Order-Specific Fertility Rate (AOSFR)

This is the number of births of order "i" to women of age "x" per 1000 women of age x, i.e.

$$\text{AOSFR}_{i(x)} = \frac{B_{i(x)}}{W_x} * 1000,$$

Where,

$B_{i(x)}$= be the number of births of the order "i" to women of age x,

W_x = the number of women of age x.

Note that

$$\sum_{i} AOSFR_{i(x)} = \sum_{i(x)} \left[\frac{B_{i(x)}}{W_x}\right] X\ 1000 = ASFR_x$$

7.4.4 Age-Parity-Specific Fertility Rate

This cohort fertility rate shows the relation between birth order and parity

The APSFR for the i^{th} parity and age "x" is defined as:

$$APSFR_x = \frac{B_{i(x)}}{W_{i-1(x)}} * 10000.$$

$ASFR_x$ may be expressed as a weighted sum of $APSFR_x$ and the proportion of women at each parity.

$$ASFR_x = \frac{B_x}{W_x} = \frac{B_{1(x)} + B_{2(x)} + \text{..........} + B_{i(x)}}{W_{(x)}}$$

$$= \frac{B_{1(x)}}{W_{0(x)}} * \frac{W_{0(x)}}{W_{(x)}} + \frac{B_{2(x)}}{W_{1(x)}} * \frac{W_{1(x)}}{W_{(x)}} + \text{......} + \frac{B_{i(x)}}{W_{i-1(x)}} * \frac{W_{i-1(x)}}{W_{(x)}}$$

7.4.5 Birth Intervals

Parity Progression Ratio shows the extent of the birth order among women. But fertility rates are also affected by the timing of the birth. The spacing between births affects the intrinsic growth rate as well as mean generational length of a population. The birth intervals provide insights into fertility behaviour of a population by disaggregating the reproductive process in to different stages. These stages start with marriage, then first birth, second birth and so on. The birth interval between two successive births of women of specified birth order, aggregated over a group of women in a specific population can be used to study the fertility patterns in the population.

The details of the type of birth intervals and their use in estimating fertility is explained in the unit III of this block.

7.5 Indirect Methods of Estimation of Fertility

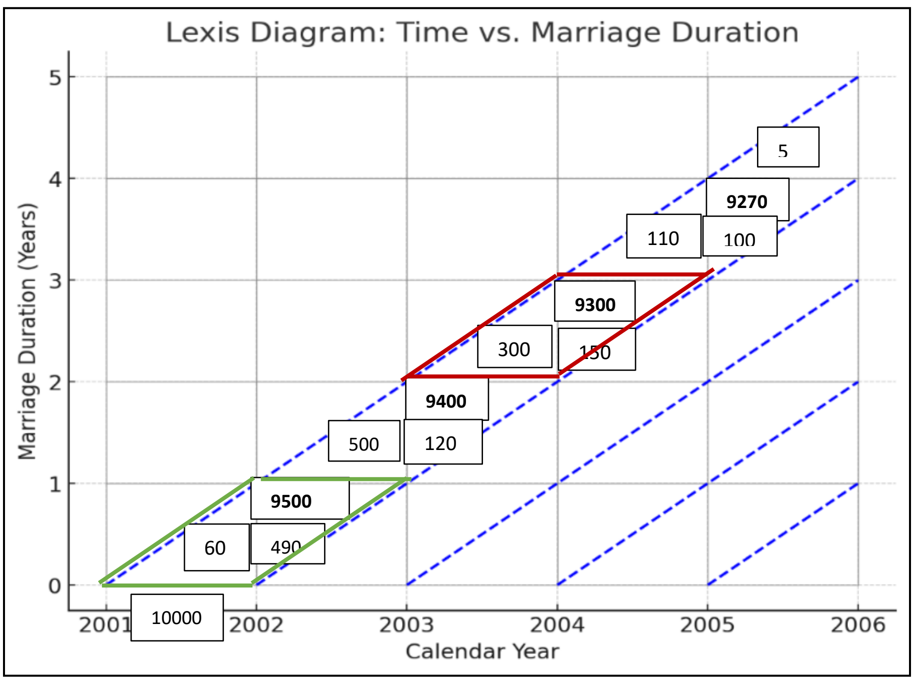
Over the decades, it has been seen that even if there is a subsequent growth in the quality and availability of data on various demographic characteristics, we still don't get sufficient enough data to understand and estimate the characteristics. A lot of countries still lack in proper and timely registration of births and conventional measurement of fertility from vital registration therefore, become too difficult or impossible. Fertility rates can be estimated indirectly using data on socio-economic characteristics from censuses and surveys. These methods are useful in the situations where vital registration is virtually incomplete. So far, we have discussed various period and cohort measures of fertility. These measures are the direct measures, by which, we mean that they can be calculated easily when the appropriate and reliable data on births and other required variables are available. In the situation where there is shortage of reliable data, other methodologies are used to estimate fertility of a population. Some of the indirect methods developed in demography are given below out of which we are going to discuss some in this unit.

- Lexis Diagram
- Reverse Survival method;
- Own-children method.
- Children ever born method.
- Brass's P/F ratio method;
- Application of stable population theory (discussed in block II)

7.5.1 Lexis Diagram

A **Lexis's diagram or Lexis Diagram** is a two-dimensional graphical representation used to analyse the relationship between age, period, and cohort in demography. It was developed by the German statistician, Wilhem Lexis in 1875. It allows visualization of population dynamics over time and helps distinguish between the effects of age, time period, and generational (cohort) differences. X-Axis (Time/Calendar Period): Represents chronological time, such as years or decades. Y-Axis (Age): Represents the age of individuals or cohorts at different time periods. The grid formed by plotting time on the x-axis and age on the y-axis forms a set of rectangular cells or squares. Each cell in this grid represents a combination of age and calendar year. A diagonal line, cohort line in the Lexis diagram represents a cohort. It runs at a 45-degree angle, where each additional year on the time axis corresponds to an additional year of age. Events such as births, deaths, or other life-course milestones (e.g., marriage, employment) can be represented as points or segments along these diagonal cohort lines. Lexis's Diagram shows the effects related to how a cohort of a certain characteristic behave or experience an event. These events may be births, deaths, marriage divorce etc.

<u>Fig: 2.1 The Lexis Diagram</u>



To understand the use of lexis diagram to represent cohort fertility, let us take an hypothetical cohort of women married in year 2001 in a geographical area. Let 10,000 be the size of initial cohort of women married followed till the end of 2005. For the sake of simplicity, we have taken five years of marriage starting from 0 till 5. That is the maximum duration of the marriage is considered to be 5 years, above with the marriage is assumed to be terminated by different reasons such as, divorce, separation, death of the husband etc. The x axis represents the consecutive calendar years and y axis represents the years of marriage. We have to determine how many marriages survived without any break or attrition due to death of a marriage partner or divorce or separation and how many births took place in these marriages in different periods of time. The survivors of the cohort 2001 will be the ones who have completed one year of marriage in 2002, two years in 2003 and so on.

The final table to show the fertility behaviour of the marriage cohort of 2001 till the end of 2005 for the given hypothetical population is given below.

Table 2.1: Fertility behaviour (Number of births) and survivors of the marriage cohort of 2001 till the end of 2005

Births and number of women (Marriage Cohort 2001)					
Marriage Duration (years)	**Calendar Years**				
	2001	**2002**	**2003**	**2004**	**2005**
0	**10,000**				
	(60)	(490)			
1		**9500**			
		(500)	(120)		
2			**9400**		
			(300)	(150)	
3				**9300**	
				(110)	(100)
4					**9270**
					(50)

By figure 2.1, it is seen that the size of the initial cohort is 10,000, the number of marriages continued in the married state in 2002 will be 9500 at the end of one year. The parallelogram represents the number of marriage survived. Similarly, the number of marriages surviving by the end of two years is 9400, and so on.

Now, the total number of births to the initial cohort of size 10,000 surviving till the end of 2005 will be = (60+490+500+120+300+150+110+100+50) = 1880.

Based on the lexis diagram for marriage duration and number of births, three types of fertility rates can be derived, namely, cohort marriage duration specific fertility rate, marriage duration period specific fertility rate and cohort period specific fertility rates.

7.5.1.1 Cohort Marriage Duration Specific Fertility Rate

The cohort marriage duration specific fertility rate is the fertility rate specific to the duration (x to x+n) of marriage and is given as-

$$= \frac{No.of\ births\ between\ marital\ duration\ x\ to\ x{+}n\ \ of\ the\ given\ cohort}{Person{-}years\ lived\ between\ exact\ agesx\ to\ x{+}n\ of\ the\ same\ cohort} \text{ X1000}$$

Example 2.1: For the data given by figure 2.1 calculate cohort marriage duration specific fertility rate for the duration 2 and 3 years, i.e. fertility of those whose marriage lasted for two and three years after marriage.

It will be the total number of births in the parallelogram 2-3-2005-2004-2004 (highlighted by red)

Married women surviving at two years of marriage= 9400

Married women surviving at three years of marriage= 9300

Person years of married life lived between duration two and three years after marriage

= (9400+9300)/2 = 9350

Births for this cohort during this period = 270 (with 120 in and 2003 and 150 in 2004)

Therefore,

Cohort Marriage Duration Specific Fertility Rate = (270/9350)*1000

= 28.87 per thousand women

7.5.1.2 Marriage Duration Period Specific Fertility Rate

It is given as:

$$= \frac{\begin{array}{c} No.of\ births\ between\ marital\ duration\ x\ to\ x{+}n\ \ of\ the\ given\ cohort\ for\ a \\ given\ calenadar\ year/years \end{array}}{Person{-}years\ lived\ between\ exact\ agesx\ to\ x{+}n\ of\ the\ same\ cohort} \text{X1000}$$

Example 2.2: For the data given in figure 2.1, find Marriage Duration Period Specific Fertility Rate for the marital period 2 years of the cohort 2001 for calendar year 2003

Total number of births for the marital period 2 years of the cohort 2001 calendar year 2003= 60+490+500+120+300=

Total number of marriage sustained till the end of 2003= 9400

Therefore,

Cohort Period Specific Fertility Rate = (1470/ 9400)*1000

= 156.4 per thousand women

7.5.1.3 Cohort Period Specific Fertility Rate

This is the same as Cohort Marriage Duration Specific Fertility Rate but the calendar year has to be specified.

Example 2.3: For the data given in figure 2.1, find Cohort Period Specific Fertility Rate for the marital period 2002.

Total number of births in the calendar year 2002= 490+500 = 990

Total number of marriage sustained in 2002= 9500

Therefore,

Cohort Period Specific Fertility Rate = (990/ 9500)*1000

= 104.2 per thousand women

Therefore, it is evident that the three rates can differ due to variations in the size of the cohorts and changes in their fertility performance over time.

7.5.2 Reverse Survival Method

The reverse survival method is a demographic technique used to estimate fertility rates in populations where reliable birth records may be lacking or incomplete. This method relies on data from age distributions within a population, often obtained through censuses or surveys, and applies survival rates to estimate the number of births in previous years. By analysing the current age structure of the population and working backwards, demographers can estimate how many individuals were born in earlier periods.

This approach is particularly useful in developing regions where vital registration systems are insufficient, making direct measures of fertility difficult. It also allows for fertility estimation over time without the need for historical birth records. The method assumes that the population has experienced relatively consistent patterns of mortality and migration, which can affect the accuracy of estimates if these factors vary significantly.

The reverse survival method works by adjusting the age-specific population counts based on survival probabilities from standard life tables. By applying survival ratios to older age groups, demographers can estimate how many births must have occurred to produce the observed population at younger ages. Although it is an indirect method, it offers valuable insights into fertility trends when more precise data is unavailable.

The Population aged "0-4" enumerated in the census or surveys is the number of survivors of births occurred to a population within the last 5 years period from the census date. Think how to convert these survivors in ages 0, 1, 2, 3 & 4 years into the number of births to which these are survivors. We use some "a priori" knowledge on mortality. If we know the level of infant mortality, we can survive back the population aged "0" to get the number of births. If we express population aged "0" as P_0 and survivorship probability from birth to age 0 as P_0, which is the same as (L_0/l_0) in the life table notation, we can calculate births in the last one year as below :

$$P_0 = \mathrm{B}_{(t,t-1)} \left(\frac{L_0}{l_0}\right)$$

$$\text{Or } \mathrm{B}_{(t,t-1)} = P_0 \left(\frac{L_0}{l_0}\right)$$

where, L_0/l_0 = Survivorship probabilities from birth to age zero.

Assuming that l_0=1 (radix of the life table) we have survivorship probability ($P_{B-0)}$ ~L_0.

$B_{(t, t-1)}$ = the number of births in the preceding one year from time t.

This procedure of calculating the births in the preceding one year is called the reverse survival method. Since we are using population to calculate births indirectly rather than the directly observed births, we call the procedure `indirect method'.

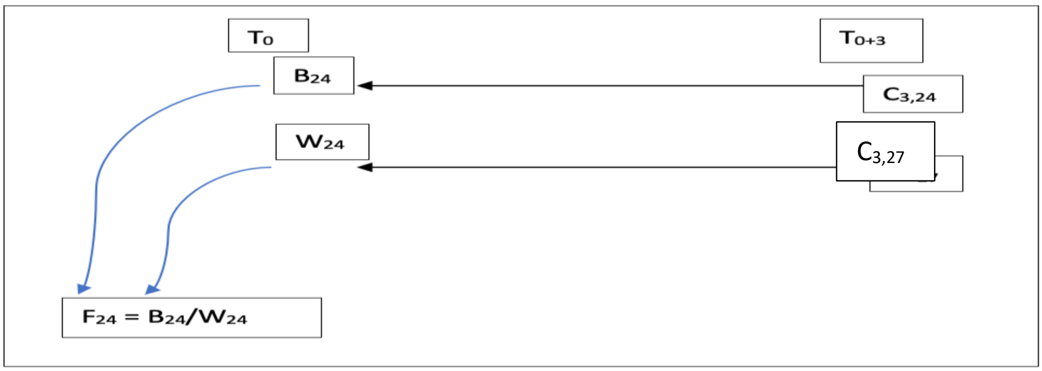
It is obvious that fertility analysis is more complex than mortality analysis is several respects. Human fertility involves two individuals of opposite sexes.

7.5.3 Own Children Method

The own-children method of fertility estimation originated in a series of special tabulations of young children by age of mother from the 1910 and 1940 censuses of the United States (U.S. Bureau of the Census 1947), which were used to generate estimates of differential fertility. Employing such tabulations, Wilson H. Grabill and Lee-Jay Cho developed the basic ideas of the method in the early 1960s (Grabill and Cho1965). Further work occurred in a study of differential fertility in the United States (Cho, Grabill, and Bogue 1970). Cho (1971c) subsequently refined the approach to obtain annual fertility estimates for each of the ten years immediately preceding a census. The 1970 round of censuses ushered in applications of the method in countries around the world, including, Bolivia, Brazil, Colombia, Costa Rica, Guatemala, Hungary, Indonesia, Japan, the Republic of Korea, Malaysia, Pakistan, Paraguay, the Philippines, and Thailand.

The own children method of estimating fertility is a reverse survival technique for estimating age-specific fertility rates for the time previous to a census or survey. Firstly, from the available household records, enumerated children are matched to the mothers within the household. The matched (i.e. own) children classified by their respective ages and mother’s age are then reverse survived to estimate number of births by age of mother in previous years. After the adjustments are made for mis enumeration (undercount or misreporting) and unmatched (i.e. non-own) children, ASFRs are calculated by taking the ratio of the number of reverse survived births by the number of reverse survived women. These estimates are, in general, computed for 15 years or years before the census. The reversal period is taken for 15 years because a large number of women (mothers) whose residence has changed in fifteen years cannot be matched. The estimation is done initially by single years of age and time. Estimates for grouped ages or calendar years are obtained by aggregating single year numerator (births) and denominator (women) and then dividing the aggregated numerator by aggregated denominator. This method minimises the effect of mis enumeration of fertility estimates.

Figure 2.2 Reverse Survival in Own Children Method



Let there be data on number of births from census at time $T=T_{0+3}$. $C_{3,27}$ denotes the number of children aged 3 who are matched to women aged 27 at the time of census. Assume that $C_{3,27}$ is already adjusted for mis-enumeration and the presence of non-own (unmatched) children. W_{27} denoted the number of women ageing 27 enumerated in the census, adjusted for mis-enumeration. Now, three years ago, at $T=T_0$, these $C_{3,27}$ children were births i.e. B_{24} to women ageing 24 i.e. W_{24}. Considering the deaths in the time period in between, the numbers B_{24} and W_{24} will be somewhat larger than the current numbers. The ASFR for three years back can easily be calculated by taking the ratio B_{24}/W_{24}.

For simplicity, we assume that the data is already adjusted for mis-enumeration and the presence of unmatched (non-own) children, and that mortality has been constant over the period of estimation. Then, if t is the time of the census

$$B_{a-x}(t-x) = C_{x,a}\,\frac{l_0}{l_x}$$

$$W_{a-x}(t-x) = W_a\,\frac{l^f_{a-x}}{l^f_a}$$

The Age-Specific Fertility Rate can therefore be given as-

$$f_{a-x}(t-x) = \frac{B_{a-x}(t-x)}{W_{a-x}(t-x)}$$

Where,

$f_a(t)$: Instantaneous age specific birth rate for women aged 'a' at time 't'.

$C_{x,a}$: Number of own children aged 'x' of mothers aged 'a' enumerated in the census.

$B_a(t)$: Number of births to women aged 'a' at time 't'.

$W_a(t)$: Number of women aged 'a' at time 't'.

l_a: Probability of surviving from birth to age 'a'.

l_a^f: Probability of surviving from birth to age 'a' for women.

The values of l_a *and* l_a^f are obtained from suitable life tables for the population. In the continuous formulation, ages and times are conceptualised as exact ages and times rather than single- year age groups or time periods.

The General Fertility Rate based on own-children method can be written as-

$$GFR(t-x) = \frac{\frac{C_x^*(t)}{L_x}}{\sum_a W_a(t) \frac{L_{a-x-0.5}^f}{L_{a-0.5}^f}}$$

Where,

$C_x{*}(t)$ = children aged '*x*' in completed years at time '*t*'.

L_x = survivors aged '*x*' in the life table for the cohort.

The own-children method is useful in less developed countries where vital registration often is seriously deficient. In these countries, censuses as well as vital registration tend to suffer from undercount; but censuses, to which the own-children method can be applied, tend to be more accurate. Moreover, if omissions in censuses tend to be of entire households, then age-specific child-woman ratios tend to be biased comparatively little by undercount. Then own-children estimates of age-specific birth rates, which can be viewed as mortality-adjusted, age-specific child-woman ratios, also tend to be biased comparatively little by undercount.

Own-children estimates of fertility can be tabulated by whatever characteristics are recorded in the census or household survey. Therefore, it can be used to tabulate fertility estimates by various socio-economic variables such as education, occupation, language, and religion, etc. Since vital registration systems do not normally collect information on these variables, fertility estimates derived from vital registration cannot be so tabulated.

This method is cost-effective since it does not require new data collection, and is usually applied to data from censuses or household surveys originally conducted for other purposes, such as labour force or income surveys. The key requirement is that children can be matched to their mothers, and basic demographic information like age and sex is available. The own-children method is often applied to large census or survey samples, making it more useful than fertility surveys with limited maternity histories. It allows for the calculation of age-specific birth rates for each year in the 15-year period preceding the census or survey, without age truncation. This is because children are matched to women up to age 65, enabling estimates for women up to age 50 in the years before enumeration. In contrast, fertility estimates based on maternity histories often suffer from age truncation.

The major limitation of the own-children method is that both the age pattern of fertility and the estimated trend of fertility can be severely distorted by age misreporting. Such distortions can be lessened but usually not eliminated by aggregating estimates over several calendar years. However, if the method is applied to two or more successive censuses, the estimated fertility trends for single calendar years partly overlap, since each census yields estimates for a fifteen-year period. For example, own-children fertility estimates derived from two censuses ten years apart yield trends that overlap during the first five years preceding the first census. The degree to which these two trends coincide during this five-year period may indicate the nature of systematic biases due to age misreporting, and the trend accordingly adjusted. The own-children estimates may be biased by migration, and if migration rates are high and if migrants are a highly selected group by virtue of their age-specific fertility behaviour, then bias in the fertility estimates from this source can also be serious. Although this is not usually much of a problem since migrants are normally a small proportion of a population.

7.5.4 Children Ever Born Method

The Children Ever Born (CEB) method is an indirect technique to estimate fertility in populations where complete birth records are unavailable. This method relies on survey or census data, where women of reproductive age (usually aged 15–49) are asked about the total number of children they have ever given birth to. Like the own-children method, CEB method is also useful in developing nations where vital registration data are incomplete or unreliable.

The CEB method calculates fertility estimates based on the mean number of children born to women in different age groups. Women are grouped by age, and the mean number of children ever born in each age group is computed. Mean number of children ever born is usually computed as the ratio of the number of live births to women in particular age group and the the number of women. When the total number of live births to women in age group is not available but a tabulation of the distribution of women by age group and number of children ever born is given, then the mean number of children ever born to women in age group is calculated. This information is further used to estimate Age-Specific Fertility Rates (ASFRs) and The Total Fertility Rate (TFR). The basic assumption of CEB method is that the fertility performance of women currently aged 45-49 reflects the completed fertility of women at the end of their reproductive lives.

Let:

$C_x =$ Average number of children ever born to women in age group $'\mathrm{x}'$.

$W_x =$ Number of women in age group $'\mathrm{x}'$.

$N_x =$ Total number of children ever born to women in age group $'\mathrm{x}'$.

$N_x = C_x \, X \, W_x$

$\mathrm{ASFR}_x =$ Age $-$ specific fertility rate for age group x

TFR = Total fertility rate

The mean children ever born for age group x is given as:

$$C_x = \frac{N_x}{W_x}$$

ASFRs can be derived indirectly from CEB by adjusting for survivorship and fertility patterns. The Total Fertility Rate (TFR) is derived by calculating the age-specific fertility rates and summing them across all age groups.

In practice, because the ASFRs are estimated indirectly, the TFR can be approximated using the cumulative CEB values of older women (typically women aged 45-49) to estimate completed fertility.

Example 2.4: Let us assume a hypothetical population with the data provided on children ever born for five age groups:

Table: 2.2

Age Group (1)	**Number of Women At age 'x' (W_x) (2)**	**Total Children Ever Born (N_x) (3)**	**Mean Children Ever Born (C_x) (4) = (2)/(3)**
15-19	200	50	0.25
20-24	180	270	1.50
25-29	160	480	3.00
30-34	140	560	4.00
35-39	120	540	4.50
40-44	100	450	4.50
45-49	80	360	4.50

The mean number of children ever born (CEB) increases with age, reflecting the accumulated fertility of older women. The TFR can be estimated using the CEB for the oldest group (45-49) as a proxy for completed fertility. Therefore:

$$TFR \approx 4.50$$

This indicates that, on average, women in this population are expected to have about 4.5 children by the end of their reproductive lives.

The Children Ever Born (CEB) method is a simple and cost-effective way of estimating fertility, especially in populations where reliable birth records are unavailable. While it provides useful fertility estimates, it relies heavily on the assumptions about the similarity of fertility patterns across cohorts and the accuracy of reported data, such as age and number of children. Data on children ever born to successive age groups of post-reproductive women (aged 45 and above) can offer insights into trends in completed cohort fertility, assuming that these women's fertility behaviour is largely representative of their original birth cohort. This is valid as long as the effects of mortality and migration are minimal, and the omission of deceased children does not significantly impact the estimates.

7.5.5 Brass's P/F Ratio Method

The P/F ratio method was proposed by Brass in 1964. The basis of this method is that if the fertility of a given population has been constant over time than cohort and period measures of fertility will be identical. Brass stated that P being the average parity (children ever born)of a cohort of women till a given age and F are closely related to cumulated current (period) fertility up to that same age.

Brass developed this method to obtain the level of current fertility by comparing the average CEB and CFR based on births during last one year or so. The assumptions made by Brass are given as:

i. Fertility has remained constant in the past.

ii. The reference period error is independent of age. It means that the reported age pattern of fertility is correct but not the level.

iii. The data on CEB reported by the younger women is accurate.

Average parity of women belonging to any age group does not necessarily relate to the cumulated fertility of the women of age at the mid-point of the age group. If the age distribution of the women within the age group is uniform and fertility is constant at each age within that age group, then

If F_i is the Cumulative Age Specific Fertility and f_i is the age specific fertility rate for the women of age group 'i' starting with 15 - 19 age group.

Then, $F_{i-1} = 5 * (f_1 + f_2 + \cdots + f_{i-1)})$, if the age specific fertility rate at is given at each age.

The comparison between P_i and F_i can provide the Age Specific Fertility Rate (ASFR).

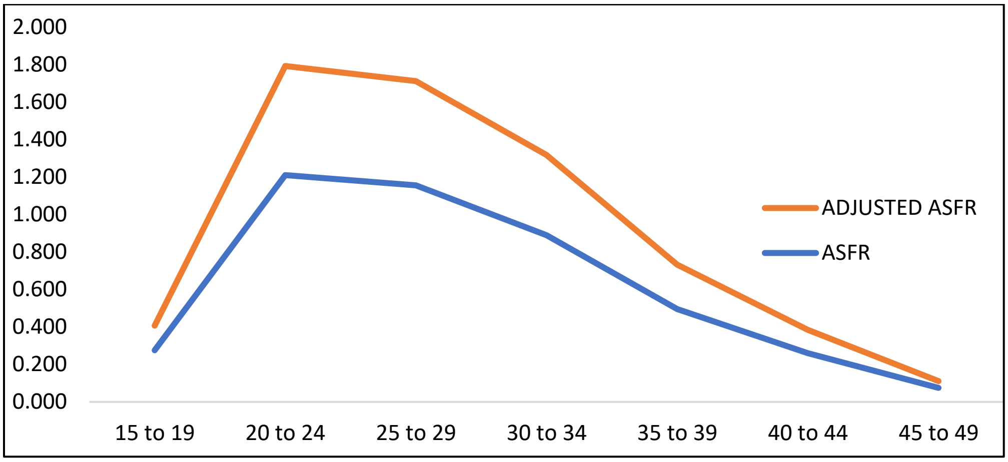
The P/F Ratio, defined as P_i/F_i gives the estimate of Age Specific Fertility Rate. The estimate of the ASFR is for women ½ year younger than the present age. P/F Ratios for women ages 20-24 and 25-29 are used to adjust the whole set of ASFRs upwards. This gives the adjusted ASFRs and hence TFR can also be calculated by using the relationship between ASFR and TFR.

Example 2.5: For the hypothetical population for year 1998-2000, data on the Age Specific Fertility Rate (ASFR) and Children Ever Born (CEB) is given. Calculate the adjusted ASFR and hence TFR by P/F Ratio Method.

Table: 2.3

Age (1)	**ASFR (2)**	**Estimated Cumulative Fertility F_i (3)**	**CEB P_i (4)**	**P/F RATIO (5) = (4)/(3)**	**Adjusted ASFR (6)=(2)/1.481**
15 to 19	0.275	0.275	1.470	5.345	0.407
20 to 24	1.210	1.485	3.500	2.357	1.792
25 to 29	1.155	2.640	3.910	1.481	1.711
30 to 34	0.890	3.530	5.040	1.428	1.318
35 to 39	0.495	4.025	5.590	1.389	0.733
40 to 44	0.260	4.285	6.210	1.449	0.385
45 to 49	0.075	4.360	6.450	1.479	0.111
TFR	4.360	*		TFR	6.457

Figure: 2.3 Age Group Wise ASFR and Adjusted ASFR



ASFR given in column (2) are given for the year 1998-2000. Using these we calculate the Cumulative ASFR. The P/F ratio in column number (5) shows the relation between ASFRs and CEB or mean parity.

The examination of the P/F ratio series is crucial. For example, if the data are correct and fertility has been declining in recent years, the P/F ratios may display an upward trend across older age groups of women. This happens because older women had most of their children during a period of higher fertility, while younger women gave birth more recently, when fertility had already fallen. This pattern is particularly evident when an increase in the age at marriage results in reduced fertility among the youngest women. In such cases, the technique's results do not reflect the current cumulative fertility pattern. This method remains unaffected if some women fail to report the number of children ever born, as long as those who did not report have the same number of children as those who did. Mis-enumeration produces a bias in the estimates from this method. Misreporting of age can lead to unpredictable biases. This could obviously influence the overall level of the total fertility rate.

7.6 Bongaarts Proximate Determinants of Fertility

Davis and Blake in 1956 came up with eleven intermediate variables that affect fertility directly or indirectly. Bongaarts (1978) refined the eleven intermediate variables suggested by Davis and Blake. He determined a list of eight variables known as the Proximate Determinants of Fertility. These factors were exposure factors, Marital Control Factors and Natural Marital Control

Factors. The eight variables are: marriage, contraception, induced abortion, lactational infecundability, frequency of intercourse, spontaneous intrauterine mortality, sterility and duration of fertile period.

Bongaarts then reduced these eight variables into four indices also known as Bongaarts Model. The four principle proximate variables i.e. marriage, contraception, lactational infecundability and induced abortion are considered inhibitors of fertility because fertility is lower than its maximum value as a result of delayed marriage (and marital disruption), the use of contraception and induced abortion, and lactational infecundability induced by breast feeding or abstinence.

The proportion of childbearing women will be the women who are sexually active and living in a stable sexual union i.e. marriage. The higher the number of marriages higher will be the fertility. **Contraception,** like abstention and sterilization is a parity dependent practice followed to reduce the risk of conception. The higher the contraception use, lower will be the fertility in the population. **Lactational Infecundability** means the duration of infecundability after the termination of gestation period of a woman. This lasts until the normal ovulation and menstruation is restored and it consists of duration of lactation. Prolonged lactation, as stated by Bongaarts, affects ovulation and hence increases the birth interval and reduces natural fertility. **Induced Abortion** is an act of deliberately interrupting the natural gestation period by terminating it. The increase in amount of abortion reduces natural fertility of a woman.

The fertility inhibiting effects of the four principal variables are measured in the model by four indices which are defined as follows:

C_m = Index of Marriage
(= 1, if all women of reproductive age are married and
= 0, in the absence of marriage).

C_c = Index of Contraception
(=1, in the absence of contraception and
= 0, if all fecund women use 100% effective contraception).

C_a = Index of induced Abortion
(=1, in the absence of induced abortion and

= 0, fall pregnancies are aborted).

C_i = Index of lactational infecundability
(=1, in the absence of lactation and postpartum abstinence and
= 0, if the duration of infecundability is infinite).

Obviously, the indices take only two values, 0 and 1. When there is no fertility-inhibiting effect of a given intermediate variable, the corresponding index equals 1 and if the fertility inhibition is complete, the index equals 0.

Four different types of cumulated fertility levels are identified from which the impact of proximate variables can be derived.

Notably, Total Fertility Rate (TFR), Total Marital Fertility (TM) and Total Natural Marital Fertility may vary widely among populations. However, the TFs of most populations fall within the range 13 to 17 births per women with an average near 15. Actually, the TF is relatively less variant because the three proximate factors (natural fecundability, spontaneous intrauterine mortality and permanent sterility) which determine Total Fecundity (TF), usually cause only modest changes in fertility.

Thus, the basic relations between the indices and the cumulative fertility measures are

$$C_m = \frac{TFR}{TM} \qquad (7.1)$$

$$C_c * C_a = \frac{TM}{TNM} \qquad (7.2)$$

$$C_i = \frac{TN}{TF} \qquad (7.3)$$

The above equations can be rearranged as :

$$TFR = Cm * TM \qquad (7.4)$$

$$TM = C_1 * C_a * TNM \qquad (7.5)$$

$$TNM = C_i * TF \qquad (7.6)$$

From these equations, if follows that :

TFR= $C= * C_c * C_a * C_i * TF$ (7.7)

TM= $C_c * C_a * C_i * TF$ (7.8)

TNM= $C_i * TF$ (7.9)

Equations (7,8,9) summarize the basic structure of the '**Bongaarts Model'** by relating the cumulative fertility measures to the proximate determinants.

Now let us see how we estimate the indices C_m, C_c, C_a and C_i from the measures of proximate variables in the following section.

Estimation of Indices:

The four indices developed in Bongaarts Model have been estimated in the following manner.

Index of Marriage (C_m):

$$C_m = \frac{\sum m(a) * g(a)}{\sum g(a)} = \frac{TFR}{TM} \tag{7.10}$$

where,

m(a) = Age Specific Proportions of married females

g(a) = Age Specific Marital Fertility Rates

TM denotes the Total Marital Fertility and

Index of Contraception (C_c):

Cc = 1-1.08 * u * e, (7.11)

where,

u = proportion of married women using modern contraception; and

e = average use-effectiveness of contraception.

The coefficient 1.08 in the above equation presents an adjustment for the fact that women (couples) do not use contraception if they know or believe that they are sterile. Thus,

$$u = \sum U_i \quad and \quad e = \frac{\sum e_i U_i}{\sum U_i} \qquad (7.12)$$

then, we have: TFR = C_m * C_c * TNM (7.13)

Index of Abortion (C_a):

$$C_a = \frac{TFR}{TFR + A} \qquad (7.14)$$

where, "A" is the reduction in fertility associated with a given level of abortion rate, and it is calculated as-

A = b x TA, (7.15)

where, TA is the total abortion rate, Thus, we have,

$$C_a = \frac{TFR}{TFR + 0.4(1+u)TA} \qquad (7.16)$$

and TFR =C_m * C_c * C_a * TNM (7.17)

Index of Infecundability (C_i):

The index C_i equals the ratio of the TNM in the presence and absence of postpartum infecundability caused by breast-feeding and/or abstinence.

The index C_i is, therefore, estimated as:

$$C_i = \frac{20}{18.5 + i} \qquad (7.18)$$

where, *i* = average duration of postpartum infecundability caused by breast-feeding or postpartum abstinence.

Thus, we can have:

$$TNM = C_i * TF \quad (7.19)$$

and $$TFR = C_m * C_a * C_i * TF \quad (7.20)$$

7.7 SELF ASSESSMENT EXERCISES

Q.1 Explain the difference between period and cohort measures of fertility.

Q.2 For the given data on ASFRs and Children ever born for year 2005-2007, Estimates the Adjusted ASFRs based on P/F Ratio method.

AGE	**ASFR**	**CEB**
15 to 19	0.234	2.321
20 to 24	1.267	3.210
25 to 29	1.198	2.822
30 to 34	0.987	3.901
35 to 39	0.543	4.601
40 to 44	0.008	5.876
45 to 49	0.001	6.564
TFR		

Q.3 Write short notes on the method of Children Ever Born and Own Children Method with example.

7.8 SUMMARY

This unit has focused on discussing various direct and indirect measure of fertility. Apart from period measures, discussed in unit one, we have discussed new cohort measures of fertility which are applicable in various form of population. Also, in the absence of complete and reliable data, direct measures cannot be used to explain the fertility of a population which makes it essential to discuss some other methods to estimate measures fertility such as age specific fertility and total fertility. Bongaarts proximate determinants of fertility have also been discussed in this unit after observing that various other biological and social factors such as marriage, contraception, abortion, infecundability, etc also lay major impact on fertility of a women.

7.9 REFERENCES

- Cho, L. J., Retherford, R.D., Choe, M.K (1986). The Own Children Method of Fertility Estimation. Honolulu, HI : Population Institute
- Bogue, D. (1969). *Principles of Demography*, New York: John Wiley and Sons.
- Pathak, K. B. and Ram, F. (1998). Techniques of Demographic Analysis, Himalaya Publishing House
- Srinivasan, K. (2015). Training Manuals on Demographic Techniques, UNFPA, Delhi.
- https://demographicestimation.iussp.org/content/overview-fertility-estimation-methods-based-pf-ratio

7.10 FURTHER READINGS

- Keyfiitz, N. (1968). *Introduction to the Mathematics of Population*, Reading, Mass: Addison-Wesley.

- Pathak, K. B. and Ram, F. (1998). Techniques of Demographic Analysis, Himalaya Publishing House

UNIT - 8 FERTILITY MODELS

Structure

8.1 Introduction

One of the most interesting developments in the study of human fertility is the application of complex models. A model is a symbolic and simplified representation of reality which aids in the patterning of observed behaviour. A model may be deterministic or stochastic. Deterministic models assume a functional relation between their input and output variables. In stochastic models, the input variables are treated as probabilistic distributions and the relations between input and output variables are left to probability.

Over the time probability models have become a popular and powerful tool in studying demographic phenomenon. One of the difficulties in demographic research is its non-experimental nature. In demography, it is not feasible to control a number of factors and repeat experiments under identical condition but at the same time it is important to know what would happen by changing a factor if every other factor is kept constant. A lot of methodologies such as Regression methods and OLS have been used very frequently in the researches to study interrelationships of different factors but the probability approach is very different and has proven to be very useful to various populations. At the later stage, such models may be utilized to estimate some of the parameters of the model by obtaining a fit to the observed data. In some specific situations some results may appear apparently inconsistent but when the problem is tackled with the help of the model the apparent inconsistency may be explained.

8.2 Objectives

After completion of this unit, you should be able to:

- Understand various type of birth intervals.
- Understand the probability models for different type of birth intervals.
- To explain the probability model for total number of conceptions in fixed time.

8.3 Birth Interval

Study of birth interval in human population is important not only because it reflects the reproductive behaviour of a population but also due to their influence on health status of mothers as well as children. Birth intervals are also related with the population growth. Birth intervals can be categorised as follows:

i. First Birth Interval
ii. Closed (Or Inter-Live) Birth Interval

iii. Open Birth Interval

iv. Forward Birth Interval

v. Stradling Birth Interval

Marriage 1st birth 2nd birth ....(i-1)th birth ith birth Survey Date (i+1)th birth

-----*--------*--------------*-------------------*-------------*-------------*--------------*---------

T_o T_1 **T_i** **U_i F_i**

---------SBI-----------

Where,

T_o: The interval between marriage to the first live birth called first birth interval;

T_i: The closed (or inter-live) birth interval between i^{th} and $(i+1)^{th}$ birth (i= 1,2,3,…);

U_i: The open birth interval for women of i^{th} parity;

Fi: The forward birth interval for i^{th} parity women giving $(i+1)^{th}$ birth after survey date;

$U_i + F_i$: The **Straddling Birth Interval (SBI)**;

The above figure explains the construct of birth intervals in reproductive span of a woman. It should be noted that except first the birth interval, all birth intervals possess a period of non-susceptible period known as Postpartum amenorrhoea (PPA) period associated with the result of a conception, in which a woman cannot conceive. This result is usually end of the gestation period, termination of pregnancy due to different reasons.

8.4 Estimation of Fecundability for Different Birth Intervals

Estimation of fecundability of couples in a population is an important component of any study of human reproduction. It is basic to the investigation of differences in the child bearing patterns in different societies and essential for evaluating the efficiency of contraceptive method. Since it cannot directly be observed, statistical methods of estimation are necessary for this purpose. There are various approaches to the estimation of fecundability for various birth patterns, some of them are discussed in this unit.

8.4.1 Probability Model based on the Proportion of Women Conceiving First Time in a Particular Month after marriage (First Birth Interval)

Gini proposed method of estimating this probability that would be unaffected by the duration of exposure to the risk of pregnancy based on the assumption that fecundability of a married woman remains constant from month to month before the first conception. He derived an estimate of the mean fecundability of those women who conceive for the first time in months x to x + y from marriage.

If the proportion of females conceiving in the first month after marriage is p, in the second month is q.p (q=1-p), in the third month is q.q.p and in the x^{th} month is q^{x-1} p.

Hence, the proportion conceiving in the x^{th} month can be written as-

$$c_x = \int p(1-p)^{x-1} f(p)\mathrm{dp} \qquad (8.1)$$

And therefore, we can write that the proportion conceiving in the $(x+1)^{th}$ month is-

$$c_{x+1} = \int p(1-p)^{x} f(p)dp \qquad (8.2)$$

And similarly, the proportion conceiving in the $(x+y)^{th}$ month is-

$$c_{x+y} = \int p(1-p)^{x+y-1} f(p)dp \qquad (8.3)$$

So mean fecundability of conceiving in x^{th} month is –

$$p = \frac{\int p.p(1-p)^{x-1} f(p)dp}{\int p(1-p)^{x-1} f(p)dp}$$

$$= \frac{C_x - C_{x+1}}{C_x} \qquad (8.4)$$

And the mean fecundability of conceiving in (x+1)th month is-

$$p = \frac{\int p.p(1-p)^x f(p)dp}{\int p(1-p)^x f(p)dp}$$

$$= \frac{C_{x+1} - C_{x+2}}{C_{x+1}} \qquad (8.5)$$

Similarly, we can write the mean fecundability of conceiving in (x+y)th month is given by-

$$p = \frac{c_{x+y} - c_{x+y+1}}{c_{x+y}} \qquad (8.6)$$

Since proportion conceiving in xth, (x+1)th,……….(x+y)th month is c_x, c_{x+1}, c_{x+2},……..,c_{x+y}, so the mean fecundability of conceiving in x to (x+y) month is given by-

$$\frac{c_x \dfrac{c_x - c_{x+1}}{c_x} + c_{x+1} \dfrac{c_{x+1} - c_{x+2}}{c_{x+1}} + \dots\dots\dots + c_{x+y} \dfrac{c_{x+y} - c_{x+y+1}}{c_{x+y}}}{c_x + c_{x+1} + c_{x+2} + \dots\dots\dots + c_{x+y}}$$

$$= \frac{c_x - c_{x+y+1}}{c_x + c_{x+1} + c_{x+2} + \dots\dots + c_{x+y}} \qquad (8.7)$$

This model (8.7) gives the average probability of conceiving for the first time in (x+y)th month after marriage.

8.4.2 Probability model for the time of first conception (or time of first birth) after marriage (Homogeneous group of females)

If it is assumed that fecundability is constant from month to month till the occurrence of first conception, then, let the random variable X represent the time of first conception. Therefore, X will follow a geometric distribution and is given as

$$p(x = r) = pq^{r-1} \quad r=1,2,3,\ldots;\ q=1-p \tag{8.8}$$

where p is fecundability.

The mean of x is 1/p. So that 'p' can be estimated as the reciprocal of mean time of the first conception.

i.e. $\hat{p} = \frac{1}{\bar{x}}$; $\bar{x} =$ Average time of 1st conception.

If it is assumed that every conception results in live birth then time of first birth (say y) is nine months larger than time of first conception 'X' i.e. Y= X+9. So that fecundability can also be estimated by utilizing the data on first birth.

One point should be kept in mind i.e. the data on the time of first birth or first conception should be utilized only for such females where the marital duration is large. So that the probability of giving the first birth in this interval is almost one. Now,

$$E(x) = \int \frac{1}{p} f(p)dp = \frac{1}{p_H} \tag{8.9}$$

Where p_H is the harmonic mean fecundability. If fecundability varies in the population then

$$E(x) = \frac{1}{p_H} \tag{8.10}$$

Thus, the reciprocal of the average time of the first conception gives an estimate of harmonic mean fecundability. Obviously harmonic mean fecundability would be smaller than the arithmetic mean fecundability.

8.4.3 Probability model for the time of first conception (considering a heterogeneous group of females)

I. (Taking Time as Discrete):

Probability model for the time of first conception is derived on the basis of following assumptions:

1. Fecundability of a female remains constant from month to month till the occurrence of first conception.

2. Fecundability varies among the females and it has β -distribution given as

$$f(p) = \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)}$$

Where, $0 < p < 1, a > 0, b > 0$ (8.11)

Let 'x' denote the time of first conception after marriage. Thus for a female with fecundability p the distribution of x is geometric i.e. for a heterogeneous group of females the probability distribution of x is given as-

$$p[X = j] = p(1-p)^{j-1}$$

$; j = 1, 2$ (8.12)

So that for heterogeneous group of females the probability distribution of X is obtained as

$$p[X = j] = \int_0^1 \frac{p^{a-1}(1-p)^{b-1}\, p(1-p)^{j-1}}{\beta(a,b)} dp$$

$$p[X = j] = \frac{\beta(a+1, b+j+1)}{\beta(a,b)}$$

$; j = 1, 2, 3 \ldots$

(8.13)

Since for given p, the expected value of X is given as-

$$E(X) = \frac{1}{p}$$

(8.14)

So for the heterogeneous group of females the mean is obtain as-

$$E(X) = \int_0^1 \frac{1}{p} \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)} dp = \frac{\beta(a-1,b)}{\beta(a,b)} = \frac{\frac{\sqrt{a-1}\sqrt{b}}{\sqrt{a+b-1}}}{\frac{\sqrt{a}\sqrt{b}}{\sqrt{a+b}}}$$

(8.15)

Since,

$$= \frac{a+b-1}{a-1} \equiv \frac{a+b-1}{a-1} \cong \frac{a+b}{a}$$

(8.16)

$$\frac{1}{p} = \frac{a+b}{a}$$

$$\hat{p} = \frac{\hat{a}}{\hat{a}+\hat{b}}$$

Similarly, for given p

$$E(X^2) = \frac{2}{p^2} - \frac{1}{p}$$

(8.17)

Therefore, for a heterogeneous group of females.

$$E(x^2) = \int_0^1 \frac{2}{p^2} - \frac{1}{p} \frac{p^{a-1}(1-p)^{b-1}}{ab} dp$$

$$E(x^2) = \frac{(a+b-1)(a+2b-2)}{(a-1)(a-2)}$$

(8.18)

$$V(x) = E(x^2) - [E(x)]^2$$

$$= \frac{(a+b-1)\ (a+2b-2)}{(a-1)\ (a-2)} - \frac{(a+b-1)^2}{(a-1)^2}$$

$$= \frac{(a+b-1)}{(a-1)} \left[\frac{a+2b-2}{a-2} - \frac{a+b-1}{a-1} \right]$$

$$= \frac{(a+b-1)}{(a-1)^2\,(a-2)} [a^2 + 2ab + -2a - a - 2b + 2 - a^2 - ab + a + 2a + 2b - 2]$$

$$V(x) = \frac{(a+b-1).\,ab}{(a-1)^2\ (a-2)}$$

(8.19)

Thus, if we have data on the time of first conception after marriage and all the females are exposed to the risk of conception for a long time, then the theoretical expressions for mean and variance of the time of first conception can be equated with the observed mean and variance and the two parameters a and b can be estimated following the method of moments.

Knowing the values of the estimates of parameters a and b say $\hat{a}\ and\ \hat{b}$, the mean fecundability is estimated as $\frac{\hat{a}}{\hat{a}+\hat{b}}$ because theoretically the mean fecundability is

$$\int_0^1 \frac{p\ p^{a-1}(1-p)^{b-1}}{p(a,b)}\,dp = \frac{\beta(a+1\,b)}{\beta\ (a,\ )} = \frac{a}{a+b}$$ (8.20)

If we further make an assumption that every conception results in live birth, then the time of first birth say Y = X + 9 where 9 is the length of gestation period.

Thus, mean fecundability can also be estimated with a knowledge of the data on time of first birth.

Further, we also assume that not all females would be fertile during the risk period, therefore let (1-β) be the proportion of married women are primarily sterile and the cohort is observed up to T* months, we have

$$\text{Prob }[T=1] = \beta\left[1-\alpha+\alpha\frac{B(a-1,b)}{B(a,b)}\right] \qquad (8.21)$$

$$\text{Prob }[T=t] = \alpha\beta\frac{B(a+1,b+t-1)}{B(a,b)}, \; t = 2, 3 \ldots T^* \qquad (8.22)$$

$$\text{Prob }[T>T^*] = 1-\beta+\alpha\beta\frac{B(a+1,\ b+T^*)}{B(a,b)}. \qquad (8.23)$$

Application of this model has been illustrated by Pathak and Sastry (1984).

ii.) Taking Time as Continuous:

Without discretising the time parameter, we can take time as a continuous random variable as well. The role of the geometric distribution can then be taken by the negative exponential distribution for describing the waiting time to conceive after marriage. Thus, if T denotes the time of first conception, then its density function, say, f(t) is given by

$$f(t) = \lambda e^{-\lambda t}, \; t > 0, \lambda > 0 \qquad (8.24)$$

where, λ is instantaneous fecundability. A number of modifications of the above simple distribution exists to describe realistic situations as discussed in the earlier section on discrete models. As for instance, we assume that λ follows a type III Pearsonian distribution as considered by Singh (1964) and we have,

$$f(t) = \int f\left(\frac{t}{\lambda}\right) dF(\lambda)$$

$$= \int_0^\infty 1_{\lambda e} - \frac{b^a e-^{\lambda b} \lambda^{a-1}}{a!} d\lambda$$

$$= \int_0^\infty 1_{\lambda e} - \frac{(\lambda b)^a e-^{\lambda(b+1)}}{a!} d\lambda$$

$$= \int_0^\infty \frac{b^0 \{\lambda^a (b+t)^{a+1}\} e-\lambda^{(b+t)}}{a!} d\lambda$$

$$= \frac{ba}{(b+t)\, a+1} \text{ x } \frac{(a+1)!}{a!}$$

$$= \frac{ab^a}{(b+t)^{a+1}} \qquad (8.25)$$

The distribution function corresponding to f(t), say, F(t) is then obtained as

$$F(t) = \text{Prob.} \left[x < t / \lambda\right]$$

$$= \int [1 - e^{-\lambda t}] dF(\lambda)$$

$$= 1 - \frac{b^a}{(b+1)^a} \qquad (8.26)$$

and so the survival function, S(t), is given by

$$S(t) = \int e-\lambda t \, dF(\lambda)$$

$$= \frac{b^a}{(b+t)^a} \qquad (8.27)$$

And so, the mean and variance corresponding to above density, function is given by,

$$E(T) = \frac{b}{a-1} \tag{8.28}$$

$$\text{and} \quad V(T) = \frac{ab^2}{(a-1)^2(a-2)} \tag{8.29}$$

8.4.4 Probability Model to Estimate Fecundability for Closed (Inter-live) Birth Interval

We have studied the estimation of fecundability using the data on time of first conception or first birth. Naturally this method will give the estimate of fecundability near the age of marriage or fecundability level prevailing between marriage and first conception. It is obvious, that fecundability vary over age and parity and hence if one wants to estimate the fecundability level prevailing between say i^{th} and $(i+1)^{th}$ birth, then the probability model for the first birth interval will not be very useful. This interval will naturally be somewhat different than the interval of first birth as it involves a period of postpartum infecundability associated with a birth. Thus if we assume that every conception results in a live birth then a closed birth interval is composed of the following components.

A_i: the postpartum non-susceptible period after i^{th} birth;

W_i : the waiting time of the next conception after the end of A_i; and

g: the gestation period.

Usually g is taken as a constant (9 months) whereas A_i and W_i are random variables. Hence to find out the distribution of closed birth interval one has to derive the distribution of $A_i + W_i$. Naturally this distribution will involve the parameter fecundability which can be estimated using appropriate methods.

Let Xi (i = 1, 2, …) be the closed birth interval between the i^{th} and $(i + 1)^{th}$ live births,

Then,

$$X_i = A_i + W_i + g \tag{8.30}$$

If there are ‘r’ foetal wastages between i^{th} and $(i+1)^{th}$ live births then the closed birth interval with ‘r’ foetal losses within the interval can be written as follows:

$$X_i = A_i + A_{i1} + A_{i2} + \ldots + A_r + W_{i1} + W_{i2} + W_{i3} + \ldots + W_{ir} + g \qquad (8.31)$$

where,

A_i is the duration of post-partum amenorrhea following the birth of the i^{th} child,

W_{ij} (j =1, 2 …., r +1) is the waiting time for the j^{th} conception and

A_{ij} (j = 1, 2, …n) is the period of non-susceptibility due to j^{th} conception that ends in a foetal loss.

The distribution of X_i have been derived by several authors assuming-

i. For a woman, the occurrence of foetal losses is independent and the number of foetal losses between the i^{th} and $(i + 1)^{th}$ live births, say, R follows a geometric distribution as-

$$\text{Prob}\,[R = r] = (1\text{-}\theta)r\ \theta,\ \ r = 0, 1\ \ldots,$$

where θ is the probability that a conception ends in a live birth.

ii. The random variable A i.e. the duration of post-partum amenorrhea following i^{th} birth has pdf a’(w); (suffix i is omitted).

iii. W_j’s and A_j’s are i.i.d random variables with density functions f’(w) and f’(a) respectively, and

iv. A, W_j and A_j are independent and have the Laplace transforms as 1’(s), 2’(s) and 3’(s), then the Laplace transform ..’(s) of the density function $f_i(x)$ for a fixed ‘h’ is given.

Once the form of ’(s) is known, we can obtain the density of X_i by taking the inverse of the Laplace transform corresponding to ’(s). A number of models has been derived for some specified form of the distribution of A, W_j and A_j,.

Now, if there are no foetal losses between i^{th} and $(i+1)^{th}$ births i.e.$(1 - \theta) = 0$, and h be the value of non-susceptible period associated with a live birth (A+g), then the time of first conception after marriage or the start of the process of human reproduction follows an exponential distribution, and that the time between ith and (i+1)th conception, i≥1, follows a displaced exponential distribution with the intensity parameter 'λ', i.e.

$$f_0(i) = \lambda e^{-\lambda t} \quad ; \quad \lambda > 0, i > 0 \tag{8.32}$$

and

$$f_i(i) = \lambda e^{-\lambda(t-h)} \quad ; \quad \lambda > 0, i > h, i \geq 1 \tag{8.33}$$

Thus the average length of X_i, the interval between the ith and (i + 1)th live births, with i>0 for given h may be obtained as h + (1/λ) if the time of reproduction is infinite.

8.4.4.1 Probability Model for Order Specific Closed Birth Interval (OSCBI)

Let $X_i(T)$ be the random variable denoting the length of the interval between ith and (i +1)th births to a woman with marital duration T and at least (i +1) births before T, then from Sheps et al. (1970), we have the density function of $X_i(T)$, say, $h_i(x/T)$, as follows.

$$h_i(x/T) = \frac{G_{i-1}(T-x) f_1(x)}{G_i(T)} \tag{8.34}$$

where $G_i(T)$ is the waiting time distribution function for the (i+1)th birth since the start of the process and $f_i(x)$ is the pdf of the interval between ith and (i+1)th conceptions.

Under the assumptions given earlier, the expression for $G_i(T)$ can be derived and the expression for $h_i(x/T)$ then is given as-

$$h_0(x/T) = \frac{\lambda^{\,e-\lambda x}}{G_0(T)}, \qquad 0 \leq x < T\_g \tag{8.35}$$

and,

$$h_i(x/T) = \frac{G_{i-1}(T-x)\ e^{-\lambda(x-g)}}{G_i(T)}$$

$$h \leq x, T-g-(i-1)h, \quad i \leq 1 \tag{8.36}$$

where,

$$G_{i-1}(T\text{-}x) = \left[1 - e^{-\lambda(T-g-(i-1)h-x)} \sum_{S=0}^{i-1} \frac{[T-g-(i-1)h-x)^{s-1}}{s!}\right]. \qquad (8.37)$$

Obviously, $h_0(x/T)$ represents the probability density function of first birth interval for the women with marriage duration T years. At the time of survey, there is always possibility that some women are not pregnant because of the limited period of exposure available to them and various models have been developed for such condition. If we treat time as continuous, we obtain the model $h_0(x/T)$.

Similarly, for the higher order of births, i.e., the interval between any two successive births, a number of probability models have been proposed over time. But, these models are not applicable to women of shorter marital durations. The expressions for the first two moments corresponding to the above model are obtained as follows:

$$E(X_0) = \frac{1}{G_0(T)}\left[\frac{1}{\lambda} - \frac{1+\lambda(T-g)}{\lambda} e^{-\lambda(T-g)}\right] \qquad (8.38)$$

E(Xi) =

$$\frac{1}{G_i T}\left[\frac{1+\lambda h}{\lambda} - \frac{1+\lambda T'}{\lambda} e^{-\lambda(T-h)} - e^{-\lambda(T-h)} \sum_{s=0}^{t-1}\left\{\frac{h\lambda^{s+1}(T'-h)^{s+1}}{(s+1)!} + \frac{\lambda^{s+1}(T'-h)^{s+2}}{(s+2)!}\right\}\right] \qquad (8.39)$$

$$E(X_0^2) = \frac{1}{e^{-\lambda(T-g)}}\left[\frac{2}{\lambda^2} - \frac{1+[1+\lambda(T-g)]^2}{\lambda^2} e^{-\lambda(T-g)}\right] \qquad (8.40)$$

$$E(X_i^2) = \frac{1}{Gi(T)}\left[\frac{1+(1+\lambda h)^2}{\lambda} - \frac{1+(1+\lambda T')^2}{\lambda} e^{-\lambda(T-h)} - e^{-\lambda(T-h)}\right.$$

$$\left.\sum_{s=0}^{i-1}\left\{\frac{h^2\lambda^{s+1}(T'-h)^{s+1}}{(s+1)!} + \frac{2h\lambda^{s+1}(T'-h)^{s+2}}{s+2)!} + \frac{2\lambda^{s+1}(T'-h)^{s+3}}{(s+3)!}\right\}\right] \qquad (8.41)$$

Note that, in the above model, risk parameter λ is taken to be constant for each parity. If λ is a function of parity and λ_i (i=0,1…) be the risk of conception associated with i^{th} birth, then Singh et al. (1983) derived the p.d.f. of xi(t) as under:

$$h_i^*(x)/T = \frac{\lambda_0 e - \lambda_0 x}{G_0 * (T)}, \quad 0 < x < T\text{-}g \tag{8.42}$$

and

$$h_i^*(x/T) = \frac{\lambda_i e^{-\lambda i(x-h)} G^*_{i-1}(T-x)}{G_i^*(T)}, \quad h \le xT < T-g-(i-1)h, \tag{8.43}$$

where,

$$Gi^*(T) = \sum_{s=0}^{i} \frac{\left(\prod_{\substack{k=0\\k\neq s}}^{i} \lambda_k\right)\left\{1 - e - \lambda_s (T-g-ih)\right\}}{\prod_{\substack{k=0\\k\neq s}}^{i} (\lambda_k - \lambda_s)}$$

with

$$G_0^*(T) = 1 - e^{-\lambda_0 (T-g)} \tag{8.44}$$

8.4.5 Probability Model to Estimate Fecundability for Open Birth Interval

Open birth interval is defined as the time between the last birth and the survey point. It is easy to understand that this interval depends upon the level of fecundability, the order of birth, marriage duration and the postpartum non-susceptible period. A number of models have been proposed by various authors to describe the distribution of the open birth interval and different methods have been proposed to estimate the parameters involve in the model. This is a very different approach of estimating fecundability. One of the important advantages of open birth interval is that it relates to an event of the recent past and hence the data relating to this may be more reliable. One of the important applications of open birth interval is to estimate Parity Progression Ratio of the females of a particular parity going or not going to the next parity following the survey point. At the survey point a female may be in postpartum amenorrhea or in gestation or in free period stage or she might have reached menopause (secondary sterility).

A number of models have been proposed by various authors to describe the distribution of the open birth interval and different methods have been proposed to estimate the parameters involve in the model.

One model is discussed as follows:

i. As considered in the sections, let T_i's (the interval between the i^{th} and $(i+1)^{th}$ births, i>0) be mutually independent random variables and $f_i(t)$ be the corresponding probability density function.

Let G_{i+1} (T) be the distribution function of the random variable ($T_0 + T_1 + T_2 + +\ldots.+ T_i$) (the time from marriage to the $(i+1)^{th}$ births) is given by-

$$G_{i+1}(T) = P\,(T+T_1+\ldots.+T_i \;\leq t) \qquad (8.45)$$

If we denote the pdf of the random variable ($T_0 + T_1 +\ldots + T_i$) by $g_{i+1}(t)$, then it can be written as

$$G_{i+1}(t) = \int_0^t G_{i+1}(x)\, dx \qquad (8.46)$$

and probability of exactly 'i' births, in the time interval (0, T) can thus be given by-

$$P_i\,(T) = G_i\,(T) - G_{i+1}\,(T). \qquad (8.47)$$

Let us consider that there is a woman with parity "t" who will have another birth after the survey point T. Obviously, the interval from the date of last birth to the survey point, say, $U_i(T)$ for such woman is the left part of her i^{th} closed birth interval in which the survey point lies. Hence, the distribution of $U_i(T)$ is a conditional type of distribution which is given by-

$$h(u/T) = \sum_{i=1}^{n} P_i(T)\, h_i(u/T)$$

$$= \sum_{i=1}^{n} P_i(T) \frac{g_i(T-u)\, Q_i(u)}{P(T_i)}$$

$$= Q(u) \sum_{i=1}^{n} g_i(T-u) \qquad (8.48)$$

Where,

$\sum_{i=1}^{n} h_i(t)$ is the birth rate at time 't' regardless of parity and corresponds to the renewal density. Under well-known results that the renewal density tends proportionally to the inverse of the mean interval between the successive renewals (Cox, 1962), we have for large T,

$$h(u) = \frac{Q(u)}{\mu} \qquad (8.49)$$

where, μ is the mean inter-live birth interval and Q(u) is the probability of zero birth in open interval of length U. The r^{th} raw moment of 'U' (the open birth interval regardless of parity and large T) can be obtained as follows:

$$E(u') = \frac{1}{\mu}\int_0^{\infty} u^r [1 - f(u)]\, du$$

$$= \frac{E(T^r)}{(r+1)\, E(T)} \qquad (8.50)$$

where μ = E(T) with T being the interval between any two successive births. This result has been obtained by Leridon (1969), Sheps et al. (1970) and Srinivasan (1968) under the stationary conditions.

If we assume that a female has-

(i) A constant risk of conception is "λ",

(ii) A constant period of non-susceptibility "h" is associated with each delivery, and

(iii) There is a one-to-one correspondence between a conception and a live birth, we may have:

$$f(t) = \lambda\, e^{-\lambda(t-h)},\ t > h,\ \lambda > 0. \qquad (8.51)$$

Under these assumptions, the pdf of $U_i(T)$ for women with parity i and marriage duration T at the time of survey, say, h(u/T) with i = 1, 2, …., n, can easily be obtained as follows-

$$h_i\,(u/T) = h_{i1}\,(u/T), \quad 0 \le u < h.$$

$$= h_{i2}\,(u/T),\ h \le u < T - g - (i-1)\,h \qquad (8.52)$$

where,

$$h_{i1}(u/T) = \frac{g_i(T-u)}{P_i(T)} \qquad (8.53)$$

$$h_{i2}(u/T) = \frac{g_i(T-u)\,Q_i(u)}{P_i(T)} \qquad (8.54)$$

$$Q_i(u) = e^{-\lambda(u-h)} \qquad (8.55)$$

$$G_{i+1}(T) = 1 - e^{-\lambda(T-g-ih)} \sum_{s=0}^{i} \frac{\{\lambda(T-g-ih-u\}^2}{s!}$$

$$\text{for} \quad t > g + ih, \qquad (8.56)$$

$G_0(T) = 1$, and n is the maximum number of live births within time (0, T) such that n= [(T-g)/h] + 1 with [(T-g)/h] is the greatest integer not exceeding (t-g)/h.

This is a very distant-approach of estimating fecundability. One of the important advantages of open birth interval is that it relates to an event of the recent past and hence the data relating to this may be more reliable.

8.5 A Probability Distribution for Number of Conception (number of births) in a fixed period of time

Fecundability can be estimated by using the data on the number of conceptions (or birth) to a group of females within a specified period of time.

Suppose the number of females are observed for a period of T months from marriage. Then in this time period, numbers of conceptions to a female is a random variable. A number of mathematical models have been proposed to describe the distribution of this random variable. In this section, we will consider one such distributions

Let x denote the number of conceptions to a female in an interval (0, T). We derive the distribution of x on the basis of following assumptions:

1. The number of coituses in the interval (0, T) of length T follows a Poisson distribution with parameter (λ_1, T)

2. The coituses are mutually independent.

3. The probability p_1 that a coitus results in a conception is constant.

4. If there is a conception at a particular period of time then there is no possibility of other conception in the subsequent h parameter. h is the duration of time from a conception to the start of the next menstrual cycle following delivery. For a particular female the variation of h is small. Therefore, h can be assumed to be constant, which is reasonable as a first approximation.

5. Either the female is exposed to the risk of a conception throughout the interval (0,T) or she is not exposed to the risk of conception throughout the interval (0,T) with probabilities $\alpha \; and \; (1-\alpha)$ respectively.

6. The total number of conceptions during the time interval [0,T] cannot exceed n, where n=[T/h]+1 and [T/h] stands for the greatest integer not exceeding T/h.

7. The female is exposed to the risk of conception at the start of the interval.

Form assumptions 1, 2 and 3, it is easily seen that the number of conceptions follows or Poisson distribution with parameter $\lambda T = \lambda_1 p_1 T$. These assumptions are strong, but in the absence of any empirical evidence on the distribution of the number of coituses, Poisson distribution is applied because of its simplicity and range of variability.

Under the given assumptions, distribution of the number of conceptions during [0,T] is given as-

$$P[X = 0] = e^{-\lambda T} \tag{8.57}$$

$$P[X = i] = \sum_{m=0}^{i} e^{-(T-ih)} \frac{[\lambda(T - ih)]^m}{m!} - \sum_{m=0}^{i-1} e^{-(T-ih+h)} \frac{[\lambda(T - ih + h)]^m}{m!}$$

For i=0,1,2,…, n-1 (8.58)

$$P[X = n] = 1 - P(X \leq n - 1) \tag{8.59}$$

Under the assumptions 1, 2, 3 and 4, the probability expressions (3.30), (3.31) and (3.32) are given as-

$$P[X = 0] = 1 - \alpha - \alpha e^{-\lambda T} \tag{8.60}$$

$$P[X = i] = \alpha \left[\sum_{m=0}^{i} e^{-(T-ih)} \frac{[\lambda(T-ih)]^m}{m!} - \sum_{m=0}^{i-1} e^{-(T-ih+h)} \frac{[\lambda(T-ih+h)]^m}{m!}\right]$$

For i=0,1,2,…, n-1 (8.61)

And therefore,

$$P[X = n] = 1 - P(X \leq n - 1) \tag{8.62}$$

Where, n is the total number of conceptions during the time interval [0,T] and n=[T/h]+1 and [T/h] stands for the greatest integer not exceeding T/h.

This methodology proposed by Singh (1963,68) is based on the interrelationship between Poisson and Exponential distribution.

As per the assumptions made in 1, 2 and 3, we know that the number of conceptions on [0,T] follows a Poisson distribution with parameter λT.

Under the assumptions, the intervals between consecutive conceptions will follow exponential distribution which will be distributed i.i.d. with pdf-

$$f(t) = \lambda e^{-\lambda t}, \; \lambda > 0, \; t > 0 \tag{8.63}$$

If T_0 is the time between marriage to first conception, T_1 as the time between first and second conception, T_2 as the time between second and third conception and T_i is the time interval between i^{th} and $(i+1)^{th}$ conceptions and so on, then from assumption 2, we can say that T_i follows exponential distribution with pdf-

$$f(t) = \lambda e^{-\lambda(t-h)}, \; \lambda > 0, \; t > 0 \tag{8.64}$$

And T_0 follows exponential with pdf-

$$f(t) = \lambda e^{-\lambda t}, \; \lambda > 0, \; t > 0 \tag{8.65}$$

Now, if

$X_0 = T_0$

$X_i = T_i - h$

Then, $X_0, X_1, X_2, \ldots$ are i.i.d random variable with pdf-

$$f(x) = \lambda e^{-\lambda x}, \; \lambda > 0, \; x > 0 \tag{8.66}$$

$Z_i = X_0 + X_1 + X_2 + \ldots + X_i$ will be the sum of (i+1) i.i.d exponential distributions which will follow Gamma distribution with pdf-

$$f(z_i) = \lambda^{i+1} e^{-\lambda z_i} \frac{z_i^i}{i!} \quad i = 0, 1, 2, \ldots; \quad \lambda > 0, \; z_i > 0 \tag{8.67}$$

Clearly, $T_0 + T_1 + T_2 + \ldots\ldots + T_i < T$ is equivalent to the event that at least (i+1) conceptions occur in [0,T]. Similarly, the event $T_0 + T_1 + T_2 + \ldots\ldots + T_i < T$ is equivalent to the event $X_0 + X_1 + X_2 + \ldots + X_i < T - ih$

Thus,

$$P[T_0 + T_1 + T_2 + \ldots\ldots + T_i < T] = P[X_0 + X_1 + X_2 + \ldots + X_i < T - ih]$$

$$= P[Z_i < T - ih]$$

$$= \int_0^{T-ih} \lambda^{i+1} e^{-\lambda z_i} \frac{z_i^i}{i!} \, dz_i \tag{8.68}$$

By integration by parts

$$P[Z_i < T - ih] = 1 - e^{-\lambda(T-ih)} \sum_{m=0}^{i} \frac{\lambda^m (T-ih)^m}{m!} \tag{8.69}$$

Similarly,

$$P[Z_{i-1} < T - (i-1)h] = 1 - e^{-\lambda[T-(i-1h)]} \sum_{m=0}^{i-1} \frac{\lambda^m [T-(i-1)h]^m}{m!} \tag{8.70}$$

The event $Z_{i-1} < T$ is nothing but the probability that at least i conceptions occur in [0,T]. Thus, the probability that exactly i conceptions occur in [0,T] is given as-

$$P[X = i] = \left[1 - e^{-\lambda[T-(i-1h)]} \sum_{m=0}^{i-1} \frac{\lambda^m [T-(i-1)h]^m}{m!} - 1 - e^{-\lambda(T-ih)} \sum_{m=0}^{i} \frac{\lambda^m (T-ih)^m}{m!} \right]$$

$$= \left[e^{-\lambda(T-ih)} \sum_{m=0}^{i} \frac{\lambda^m (T-ih)^m}{m!} \quad e^{-\lambda[T-(i-1h)]} \sum_{m=0}^{i-1} \frac{\lambda^m [T-(i-1)h]^m}{m!} \right]$$

i= 0, 1, 2, …n-1(3.71)

Also,

$$P[X = n] = 1 - P(X \leq n - 1)$$

$$P[X = n] = 1 - \left[1 - e^{-\lambda[T-(i-1h)]} \sum_{m=0}^{i-1} \frac{\lambda^m [T-(i-1)h]^m}{m!} \right] \quad (8.72)$$

Such that,

$$P[X = 0] + P[X = 1] + \cdots + P[X = n] = 1$$

Incorporating assumption 3,

$$P[X = 0] = 1 - \alpha - \alpha e^{-\lambda T} \quad (8.73)$$

$$P[X = i] = \alpha \left[\sum_{m=0}^{i} e^{-\lambda(T-ih)} \frac{[\lambda(T-ih)]^m}{m!} - \sum_{m=0}^{i-1} e^{-\lambda[T-(i-1h)]} \frac{\lambda^m [T-(i-1)h]^m}{m!} \right] \quad (8.74)$$

8.6 Self-Assessment Exercises

1. Collect a data from the secondary data sources such as NFHS or AHS and fit the probability models given in the section 3.4 and 3.5 to check whether the models fits to the current fertility behaviour or not.

8.7 Summary

Statistical modelling is an important part of demography and it is an important tool to understand and project demographic characteristics. This unit focuses on the study of probability models developed to estimate fecundability for different type of birth intervals. The probability model for the number of births is also discussed here. These models can be applied for the type of data available on the births histories. If a reliable and accurate set of data is available on number of births, parity and spacing then these models can provide accurate estimates for fertility of the population.

8.8 References

- Singh, S. N. (1963). Probability models for the variation in the number of births per couple, The Journal of the American Statistical Association, Vol. 58, pp. 721-727.
- Pathak, K. B. (1966). A Probability Distribution for the Number of Conceptions. Sankhyā: The Indian Journal of Statistics, Series B (1960-2002), Dec., 1966, Vol. 28, No. 3/4 (Dec., 1966), pp. 213-218
- Brass W. (1975). Methods for Estimating Fertility and Mortality from Limited and Defective Data. Population Lab, University of North Carolina at Chapel Hill, USA.
- Yadava, R.C. and Rai, P.K. A Revisit to "Chance Mechanism of the Variation in the Number of Births Per Couple" Demography India Vol. 53, No. 1 (2024).

8.9 Further Readings

- Sheps, M.C. and Menken, J.A. (1973) Mathematical models of conception and birth. University of Chicago Press, Chicago.
- Brass W. (1975). Methods for Estimating Fertility and Mortality from Limited and Defective Data. Population Lab, University of North Carolina at Chapel Hill, USA.
- Retherford, R., Choe, M. (1993). Statistical models for causal analysis. *John Wiley and Sons Inc.*, New York.

U.P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N Demography

Block:4 Mortality and Life Table

Unit – 9 : Mortality and its Measures

Unit – 10: Life Tables

Course Design Committee

Dr. Ashutosh Gupta **Chairman**
Director, School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

Prof. Anup Chaturvedi **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. S. Lalitha **Member**
Ex. Head, Department of Statistics
University of Allahabad, Prayagraj

Prof. Himanshu Pandey **Member**
Department of Statistics
D. D. U. Gorakhpur University, Gorakhpur.

Prof. Shruti **Member-Secretary**
Professor, School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Course Preparation Committee

Dr. Anupma Singh **Writer**
Department of Statistics
University of Lucknow, Lucknow

Prof. Shruti **Editor**
School of Sciences
U.P. Rajarshi Tandon Open University, Prayagraj

Prof. Shruti **Course Coordinator**
School of Sciences
U. P. Rajarshi Tandon Open University, Prayagraj

MScSTAT – 401N/ MASTAT – 401N **DEMOGRAPHY**
©UPRTOU
First Edition: ***July 2024***
ISBN: 978-93-48987-49-5

Printed and Published by Col. Vinay Kumar, Registrar, Uttar Pradesh Rajarshi Tandon Open University, 2024.

Block & Units Introduction

The present SLM on ***Demography*** consists of Ten units with Four blocks. This is the last block of this SLM.

Block 4 - ***Mortality and Life Table,*** is the fourth block of the said SLM which is divided into two units.

Unit 9- ***Mortality and its Measures,*** is the first unit of the Self Learning Material, which explains the basic concepts and definitions of mortality. It also explains the methods of standardization of death rates in the situations where the mortality condition of varying populations has to be compared.

Unit 10- ***Life Table,*** is the second unit of the Self Learning Material, which focuses on the concept and development of Life Table. It explains the construction of complete and abridged life tables and also the approximations used for different life table function. This unit explains the use of life table in reality.

At the end of unit, the summary, self-assessment questions and further readings are given.

UNIT - 9 MORTALITY AND ITS MEASURES

Structure

9.1 Introduction

Mortality is one of the three components of population change and other two being fertility and migration. Mortality plays an important role in determining the growth of population and important cause in response to variation in mortality. The developing countries, which are undergoing a typical demographic transition, have also been affected initially by the fall in the death rates. In fact, the single most important contribution of demography has been the revelation of the fact that sharp declines in mortality rates, rather than any rise in the fertility rates, have been responsible for bringing about a rapid growth of population. The study of mortality is useful for analyzing current demographic condition as well as for determining the prospect of potential changes in mortality condition of the future.

Study of mortality is possible by several approaches, for various biological, social, economic, and cultural factors affect the health of an individual and consequently the mortality rate in society. These factors differ from society to society and population to population. Whereas, some of the factors which affect the mortality in every population that is, age and sex. All of these factors are broadly known as differentials and determinants of mortality. When, mortality is viewed from the demographic point of view, it is studied to determine changes in population size composition and structure, rather than from the medical angle. The demographic study of mortality, therefore, does not usually take into consideration the genetic factor. On the other hand, constitutional and the environmental factors provide the basis of a demographic analysis of mortality. Of the various constitutional factors including the physical, physiological, anatomical and psychological characteristics of man- the most important for the demographic study of mortality include the natural physical surrounding of the individual as well as his/her social and economic environment and personal habits.

Mortality analysis begins with good quality data on deaths and population. These data are conventionally obtained from vital registration systems and population censuses respectively. In India, various health surveys, such as, National Family Health Survey, Annual Health survey etc. are conducted on a large scale over time to understand the levels and patterns of mortality across various regions and states of the country. The crude death rate and the specific death rates (age, sex, age-sex, age-sex-cause of death specific) are simple measures of mortality. The other measures are based on the life tables. The method of its construction is dealt with in this chapter. The life table methodology has been used in many other applications in demography, such as in the analysis of marriage patterns (nuptiality tables), labour force (working life tables), school enrolment etc. The life tables are used in population estimation, population projections, and to show the impact of disability, cause of death elimination etc. on the survival of the population.

9.1.1 Mortality in India

Mortality trends are crucial indicators of a country's overall health, socioeconomic development, and the effectiveness of its public health interventions. In India, mortality rates have undergone significant changes over the past century, reflecting the nation's journey from high levels of mortality in the early 20th century to notable improvements in the 21st century. In the early 20th century, India experienced extremely high mortality rates, particularly due to infectious diseases, malnutrition, and inadequate healthcare infrastructure. Life expectancy at birth was alarmingly low, with averages of around 25-30 years in the first few decades. The prevalence of diseases such as malaria, tuberculosis, cholera, and plague were rampant, and the lack of effective treatments or vaccines resulted in high death tolls. Child and infant mortality rates were particularly high, with many children dying from preventable diseases due to poor sanitation, limited access to clean water, and insufficient maternal care.

One of the most significant mortality events during this period was the Spanish flu pandemic of 1918-1919, which claimed millions of lives across India. Famine and malnutrition also contributed to high mortality rates, with the Bengal famine of 1943 being a particularly devastating event, leading to millions of deaths due to starvation and related diseases.

Post –Independence Mortality Trend in India :

After gaining independence in 1947, India began to make substantial efforts to improve public health and reduce mortality rates. The establishment of the National Health Policy, expansion of healthcare infrastructure, and the launch of mass vaccination programs were pivotal in reducing mortality rates across the country. By the 1970s, significant improvements had been made, with life expectancy rising to around 50-55 years. The successful eradication of smallpox in 1977 marked a significant milestone in public health, contributing to a decline in mortality from infectious diseases.

During this period, infant and child mortality rates also began to decline due to better maternal and child healthcare services, improved nutrition, and public health campaigns promoting vaccination and hygiene. The introduction of family planning programs also played a role in reducing birth rates, which indirectly impacted mortality trends by improving the health of mothers and children.

Current Levels of Mortality:

India has continued to see improvements in mortality rates, with life expectancy at birth reaching approximately 70 years by 2020. This increase in life expectancy is due to the progress made in healthcare, economic development, and living standards. However, the decline in mortality from infectious diseases has been accompanied by a rise in non-communicable diseases (NCDs), such as cardiovascular diseases, diabetes, and cancer, which have become the leading causes of death in India today.

The infant mortality rate (IMR) has significantly decreased over the past few decades, from over 80 deaths per 1,000 live births in the 1980s to around 30 deaths per 1,000 live births in recent years. Similarly, maternal mortality has seen a decline, although challenges remain, particularly in rural and economically disadvantaged regions. Despite the overall improvement in mortality rates, India faces significant challenges in addressing disparities across different regions and population groups. Rural areas, particularly in states such as Uttar Pradesh, Bihar, and Madhya Pradesh, continue to have higher mortality rates compared to urban areas. These disparities are often linked to differences in access to healthcare, education, and economic opportunities. The levels and trends of mortality in India have seen dramatic improvements over the past century, reflecting the country's progress in public health, healthcare infrastructure, and socio-economic development. However, challenges remain in addressing regional disparities and the rising burden of non-communicable diseases. Continued investment in healthcare, targeted public health interventions, and efforts to reduce health inequalities are essential to sustaining and further improving mortality trends in India.

9.2 Objectives

After completion of this unit, one will be able to-

- Understand the meaning and relevance of mortality in demographic studies and in a population.
- Calculate various measures of mortality for a given population.
- Apply the method of direct and indirect standardization for a given set of population to make rational inference about their mortality situations.
- Understand the factors affecting mortality with respect to India.

9.3 Measures of Mortality

Mortality is a phenomenon that varies with age, sex and population. We often see that the mortality in early ages is higher as compared to the mortality observed in the middle age groups. Or the mortality in an under developed population is seen to be higher than the developed ones as we know that the developed nations have better health facilities and hence lesser deaths by diseases. Therefore, it becomes a need to develop and study different types of measures of mortality. We have discussed some important measures of mortality in the next section.

9.3.1 Crude Death Rate

As the name says, Crude Death Rate is considered to be a crude measure of mortality of a population. It gives the rough picture of the deaths occurring in any population in a specified time period. Crude Death Rate (CDR) is calculated by dividing the number of registered deaths in a year by the mid-year population for the same year. We take the population as the mid-year population because the distribution of the deaths is assumed to be uniform throughout the given year or the time period. The rate is expressed as per 1,000 populations.

$$\mathrm{CDR} = \frac{\mathrm{D}}{\mathrm{P}} \times 1000$$

Where D and P represents deaths registered in a year and P is the mid-year population.

This rate has a simple interpretation, for it gives the number of deaths that occur, on the average, per 1,000 people in the community.

Example:

Total number of deaths during 2010 for a region (1 Jan 2010 to 31 Dec 2010), D =65,020

Total Population at the middle of the year (1 July 2010), P =7,21,200

Then the CDR for region for the year 2010 is computed as,

$$\text{CDR} = \frac{\text{D}}{\text{P}} \times 1000 = \frac{65020}{721200} \times 1000 = 90.15 \text{ per } 1000 \text{ population}$$

The CDR of 90.15 suggests that, in that region for year 2010, 90.15 person die per 1000 population.

Merits

1. It is relatively easy to compute, requiring only the total population size and the total number of deaths in the year or specified time period.
2. Besides, it is a probability rate in the true sense of the term. It represents an estimate of the chance of dying for a person belonging to the given population, because the whole population may be supposed to be exposed to the risk of dying of something or the other
3. CDR is perhaps the most widely used of any vital statistics rates. As an index of mortality, it is used in numerous demographic and public health problems.

Demerits

1. Crude Death Rate also has some serious drawbacks. In using the CDR, we ignore the fact that it completely ignores the age and sex distribution of the population because the chance of dying is not the same for the young and the old or for males and females. Children in the early ages of their life, and older generation are exposed to higher risk of mortality as compared to younger people and the fact that it may also vary with respect to race, occupation or locality of dwelling.
2. C.D.R is not suitable for comparing the mortality in two places or same place in two period unless-

(a) The population of the places being compared have more or less then the same age and sex distribution

(b) Two period are not too distant, since in a stable large community, age-sex structure of the population shows very little change.

Remarks:

We can compute the C.D.R for males and females separately.

C.D.R for males $=\frac{\text{Mele death}}{\text{Male population}} \times 1000$

C.D.R for females$=\frac{\text{female death}}{\text{female population}} \times 1000$

Usually C.D.R lies between 8 and 30 per thousand and female C.D.R is generally less than male C.D.R.

9.3.2 Specific Death Rate

The crude death rates for specific causes of death are calculated in a similar way by selecting deaths due to specific cause as the numerator and mid-year population as the denominator. Thus,

$$\text{Cause} - \text{Specific Death Rate} = \frac{D^c}{P} \times 1000$$

Where,

'c' represents a specific cause.

The rates could be made specific to sex by selecting the numerator and the denominator for each sex of the population.

9.3.3 Age Specific Death Rates (ASDR)

As we have already mentioned in the introduction of this chapter, mortality is affected highly by the age. Therefore, it is important that mortality rates of different sections of the ages is calculated separately. The Age Specific Death Rates are calculated from deaths and population both specific to each age (or age group) of the population. Thus,

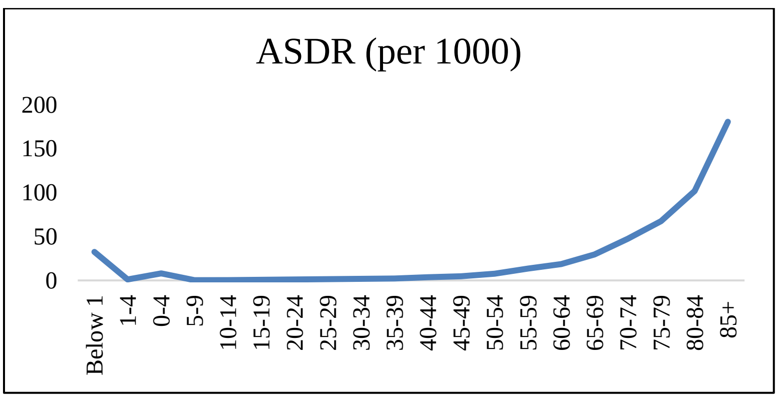
$$\text{ASDR} = \frac{nD_x}{nP_x} \times 1000$$

Where, 'x' indicates the age and 'n' indicates the class interval of age.

The ASDR is a type of central death rate, that means, a rate relating to the events in a given category during a year to the average population of the category. In a high mortality situation, the death rates by age, that is, the age specific death rates form a U-shaped curve indicating a high mortality in early and old ages. At low levels of mortality, the pattern of ASDR changes to J-shaped indicating a relatively higher mortality in the very early period of life, which drops to a low level after the hazards of early life and extends over a long period of life, and finally it rises sharply in old ages.

A typical ASDR plot looks like figure 1.1. it is seen that the mortality is relatively high at the age below 4 and at the ages above 60. The mortality at early age i.e. near birth has improved over the years in India and worldwide. This shows the improvement in the health and awareness among the population over time.

Figure 1.1



Data Source: SRS Statistical Report 2020. Office of Registrar General of India

Similarly, **The Age-Cause-Specific Death Rates** can be obtained by selecting deaths in specific age and cause group of the population as the numerator. Thus,

$$\text{Age} - \text{Cause SDR} = \frac{nD_x{}^c}{nP_x{}^c} \times 1000$$

The Age-Sex-Cause-Specific Death Rates can be calculated in the similar manner as Age Specific Death Rates. It should be noted that the sum of the cause-specific rates over all causes equals the crude death rate. Similarly, the sum of the age-cause-specific death rates equals the age-specific death rate at a given age.

9.3.4 Incidence Rate

This rate measures the annual incidence of a disease. The numerator of the rate is the number of times attacks of a certain disease are reported in a year, and the denominator is the population exposed to the disease during the same year. Note that the rate can exceed 1, as there could be repeated attacks of the disease to the same person.

$$\text{Incidence rate} = \frac{\text{C}}{\text{P}} * 1000$$

Where, C is the number of new cases of specialized disease during a given time period and P is the population at risk.

Example: Let us take a hypothetical population of 10,000 people, being observed over a one-year period to track the occurrence of a specific disease. During this year, 50 new cases of the specific disease were diagnosed.

Here:

- Number of New Cases = 50 (new cases of the disease during the year)
- Population at Risk = 10,000 (total population being observed)
- Time Period = 1 year (time period over which the cases were observed)

$$Incidence\ Rate = \frac{50}{10000 \times 1} \times 1000 = 5\ cases\ per\ 1{,}000\ population$$

The Incidence Rate in this example is 5 new cases per 1,000 people per year. This means that, on average, 5 out of every 1,000 people in this population were newly diagnosed with the disease over the course of one year.

9.3.5 Prevalence Rate

This rate is the proportion of the number of people reported as having a disease at a specific point in time. The numerator is the number of people with a disease and the denominator is the total population.

$$\text{Prevalence Rate} = \frac{\text{O} + \text{C}}{\text{P}} * 1000$$

Where, O is the Number of old cases of specified disease existing and C is the number of new cases at a given point of time and P is the estimated population at the same point of time

Example: Let us consider a small town with a population of 10,000 people. In this town, a study was conducted to find out how many people are living with diabetes.

Data:

- Total population of town = 10,000 people
- Number of people with diabetes (old or new) = 500 people

$$Prevalence\ Rate = \frac{500}{10{,}000} \times 1000$$

$$Prevalence\ Rate = 0.05 \times 1000 = 50\ per\ 1{,}000\ people$$

The prevalence rate of diabetes in the town is 50 per 1,000 people. This means that for every 1,000 people in town, 50 people are living with diabetes.

9.3.6 Infant Mortality Rate

The infant mortality rate (IMR), too, is an alternative to, and in a sense an improvement upon, the Age-Specific Death Rate for age 0 l.b.d. (last birth day) – in other words, upon the death rate for infants (i.e. children under 1 year of age). It is defined as-

$$\text{IMR} = \frac{\text{D}_0}{\text{B}} \times 1000$$

Where D_0 = number of deaths among children of age 01.b.d. (last birth day)

And B = number of live births.

This is also known as the conventional infant mortality rate (I_c).

Example: The Infant mortality rate for a region for any year say 2010 can be computed as follows, The total number of deaths among infants during year 2010 = 8,050

Total number of registered live births during the year 2010 = 90,000

The infant mortality rate for a region for the year 2010 is given by,

$\text{IMR} = \frac{D_0}{B} \times 1000 = \frac{8050}{90000} \times 1000 = 89.44 \text{ per } 1000 \text{ live births}$

The above IMR denotes that out of 1000 live births 89.44 infants died before attaining 1 year of life for the region in 2010.

Example: Suppose in a given year, a country recorded the following data:

Number of Infant Deaths (under 1 year of age): 500

Number of Live Births: 40,000

Now, calculate the Infant Mortality Rate (IMR) using the above formula:

$$IMR = \frac{500}{40{,}000} \times 1000$$

$$IMR = (0.0125) \times 1000$$

$$IMR = 12.5\ per\ 1{,}000\ live\ births$$

The Infant Mortality Rate (IMR) in this example is 12.5 per 1,000 live births. This means that for every 1,000 live births in that year, 12.5 infants died before reaching their first birthday.

9.3.7 Child Mortality Rate

It is defined as the total number of deaths of children aged 1 to 4 in the given year and geographical region and to the 1000 population of the same age in that year and same geographical regions.

$$\text{CMR} = \frac{\text{Number of deaths of children aged 1 to 4 year in the given Year and region}}{\text{Total population aged 1 to 4 in the given year and a given region}} \times 1000$$

Example: In a given year in a country, there were 150,000 live births and 1,200 deaths of children under the age of 5.

$$CMR = \frac{1{,}200}{150{,}000} \times 1000$$

$CMR = (0.008) \times 1000 = 8\ per\ 1{,}000\ live\ births$

This means that in this country, for every 1,000 live births, 8 children are expected to die before reaching the age of 5.

9.3.8 Neonatal Mortality Rate (NMR)

Number of infants dying within the first month (4 weeks)/(Up to 28 days) of life in a year and geographical region per thousand live births of same year and geographical region.

$$\text{NMR} = \frac{\text{Deaths of Infants upto 4 weeks}}{\text{Number of live births}} \times 1000$$

Example: Calculate the Neonatal Mortality Rate for the given data-

1. Number of live births in a year (in a specific region): 10,000

2. Number of neonatal deaths (infants who died between 0-28 days): 100

Then, NMR can easily be calculated as-

$$NMR = \frac{100}{10{,}000} \times 1{,}000$$

$$NMR = (0.01) \times 1{,}000 = 10\ per\ 1{,}000\ live\ births$$

The Neonatal Mortality Rate (NMR) in this example is 10 deaths per 1,000 live births, this means that for every 1,000 live births in the given year and region, 10 babies died within the first 28 days of life. This rate is an important indicator of the overall health conditions and quality of neonatal care in that region.

9.3.9 Post Neonatal Mortality Rate (PNMR)

The Number of infant deaths after 28 days to less than one year (between 4 weeks to 52 weeks) of age per thousand live births in a given year.

$$\text{PNMR} = \frac{\text{Number of deaths of children aged between 4 weeks and 1 year in the given year}}{\text{Number of live births during the same year}} \times 1000$$

Example: Suppose in a given year, a country recorded the following data:

The number of post neonatal deaths (28 days to under 1 year) = 150 and,

The total number of live births= 50,000

$$PNMR = \frac{150}{50{,}000} \times 1{,}000$$

$$PNMR = 0.003 \times 1{,}000 = 3\ deaths\ per\ 1{,}000$$

In this example, the Post Neonatal Mortality Rate is 3 deaths per 1,000 live births. This means that out of every 1,000 live births, 3 infants died between 28 days and under 1 year of age.

9.3.10 Foetal Death and Still Birth

It is used for deaths prior to the complete expulsion or extraction from its mother of a product of conception at any point of time of pregnancy. Still birth is defined as death of fetus after completing 28 weeks and till the time of birth.

9.3.11 Maternal Mortality Rate (MMR)

Maternal Mortality is one of the most important measures of mortality for the study of maternal and child health. Maternal health has proven to be very important for both mother and the child as it affects the next five years of the health of a child. This rate is sometimes represented as ratio as well.

Maternal Mortality Rate is defined by the formula-

$$1{,}000 \times \frac{D^P}{B},$$

Where,

D^P = total number of deaths from puerperal causes among the female population aged 15-49 in the given period in the given community. By puerperal we mean the causes related to pregnancy and child birth.

B = total number of live births occurring in the given period in the community.

This rate may be looked upon as an alternative to, or a refined version of the corresponding cause-of-death rate.

First, here note is taken of the fact that only the part of the female population that goes through conception sometime during the period, and not whole population, is exposed to the risk of dying from puerperal causes (i.e. causes relating to child-birth). This population may be taken to be approximately the number of mothers giving birth to live-born children plus the number of those delivered of dead fetuses.

Example: Let us suppose that the number of maternal deaths in a particular population in one year = 25 and the number of live births in that same region in the same year = 10,000. Now the maternal mortality rate can be calculated as follows:

Using the formula: $MMR = \left(\frac{25}{10{,}000}\right) \times 1{,}000$

$MMR = (0.0025) \times 1{,}000$

$$MMR = 250\ maternal\ deaths\ per\ 1{,}000\ live\ births$$

In this example, the Maternal Mortality Rate is 250 maternal deaths per 1,000 live births. This means that for every 1,000 live births in this region, 250 women die due to causes related to pregnancy or childbirth.

9.4 Standardized Death Rates

Standardization is a technique which provides a summary measure of the rates (similar to the crude rates) while controlling for the compositional variation between the populations being compared. Thus, a comparison of the standardized rates gives a 'true' comparison of the phenomenon studied. We shall illustrate the calculations of the standardized rates with the help of the death rates.

The death rate calculated using the normal methods cannot be compared if the population in consideration differs in the age sex distribution. One way to deal with this is to consider a standard population and compute standardized death rates.

9.4.1 Method of Direct Standardization

In this method the distributions of the compositional variables (age, sex, marital status etc.) of the populations being compared are made identical and the standardized rates (similar to the crude rates) are calculated such that the difference between them is only due to the variation in the age-specific rates of their population. A Standard population is selected which is employed for deriving all the standardized rates in a set to be compared.

Data Needed

1.) For one compositional variable (say age) standardization, age distribution of the standard population, and

2.) Age-specific death rates in all populations to be compared.

Calculations

A. If M (i, x) represents the Age-Specific Death Rate at age (i) for population (x), and P(i, s) is the standard population at age (i).

The standardized death rate for population x is,

$$\mathrm{SDR} = \sum_{i} \frac{\mathrm{P(i,s)}}{\mathrm{P(s)}} . \mathrm{M(i,x)}$$

The numerator is the number of expected deaths in the standard population had the age-specific death rates of population (x) applied to the standard population, and the denominator is the total standard population. The rate is multiplied by 1,000 to express the rate as per 1,000 population.

(All the calculations are done with the rates per person. Finally, the standardized death rate is multiplied by the constant 1,000).

B. If the standardized death rate is required after controlling for the two characteristics of the population, say age and sex, the data needed will be the same as on the previous page but split by sex as well.

Thus, the standardized death rate for population x will be:

$$\frac{\sum_i [P(i,s,males).M(i,males) + P(i,s,females.M(i,females)]}{[P(s,males) + P(s,females)]}$$

C. If the death rates of males and females are to be compared, these are two different populations, and the method given under A is to be used. Thus,

The standardized death rate of males = $\dfrac{\sum_i P(i,s).M(i,males)}{P(s)}$

The standardized death rate of females = $\dfrac{\sum_i P(i,s).M(i,females)}{P(s)}$

Remarks:

- The selection of the standard population is, in theory, arbitrary. However, this population should be similar to the ones for whom the rates are being compared. The population of India at a most recent census date is appropriate for measuring state-differentials in mortality, or for comparing mortality trends over time for India. If two country's rates are to be compared either one country's age distribution or the average of the two country's distributions is appropriate.

9.4.2 Method of Indirect Standardization

When the distribution of the number of deaths by age is unavailable or not reliable then the direct standardization technique cannot be used because Age-Specific Death Rates for the populations cannot be calculated. But if the total number of deaths and the age distribution of populations whose rates are to be compared are available then we can make the comparison between two population by using indirect method of standardization of death rates.

Data Needed

1.) Observed number of deaths in all populations whose death rates are to be compared.

2.) Age distribution of all populations whose death rates are to be compared.

3.) Age-specific death rates for a population to be used as standard.

4.) Crude death rate in the standard population.

Calculations: If,

- P(i, x) represents the population x at age (i),
- M(i, s) is Age-Specific Death Rate at age i in standard population,
- M(s) is the crude death rate in the standard population,
- O(x) is the observed number of deaths in population x, and
- E(x) is the expected deaths in population x, then

The standardized death rate for population x = $\frac{O(x)}{E(x)}.M(s)$

The expected deaths in population x = E(x) = $\sum_i P(i,x).M(i,s)$

(All the calculations are done with the rates per person. Finally, the Standardized Death Rate is multiplied by the constant 1,000).

Rationale behind the formula for Indirectly Standardized Death Rate:

Following the symbols we have used so far, and adding a few more, P(i,A) and P(i,B) representing two populations A and B at age (i), P(A) and P(B) as the total populations A and B, P(S) as the total standard population, EM(A) and EM(B) as the expected death rates in population A and B, and O(S) as the crude death rate in the standard population, we have:

Expected death rate in population A = EM(A) = $\dfrac{\sum_i P(i,A).M(i,s)}{P(A)}$

Expected death rate in population B = EM(B) = $\dfrac{\sum_i P(i,s).M(i,s)}{P(s)}$

Observed death rate in the standard population = O(S) = $\dfrac{\sum_i P(i,B).M(i,s)}{P(B)}$

Note that the difference between EM(A) and EM(B) is only due to the fact that different age structures of the two populations are used. Both EM(A) and EM(B) are different from O(S), again due to the different age structure.

The ratio of O(S)/EM(A) and O(S)/EM(B) is the effect on the death rates of varying age structures of the two populations A and B.

The crude death rates of the two populations are adjusted for age effect, by multiplying the crude rates by the respective adjustment factor O(S)/EM(A)and O(S)/EM(B), to allow for the variation in their population age structures. The results are the standardized rates as can be seen from the following:

Standardized death rate for Population A is given as-

$$\mathrm{SDR_A} = \frac{\mathrm{O(A)}}{\mathrm{E(A)}} \times \mathrm{M(s)}$$

Similarly,

Standardized death rate for Population B is given as-

$$\mathrm{SDR_B} = \frac{\mathrm{O(B)}}{\mathrm{E(B)}} \times \mathrm{M(s)}$$

Advantages of Indirect Standardization:

- **Simplicity:** It is simpler to apply than direct standardization, especially when age-specific rates in the study population are unstable or not available.
- **Stability:** It provides more stable results in small populations or in populations with sparse data.
- **Comparative Analysis:** It allows comparison across different populations with varying age distributions by standardizing the rates.

Disadvantages of Indirect Standardization:

- **Dependence on Standard Population:** The results are dependent on the choice of the standard population. Different standard populations may yield different SMRs.
- **Less Detailed:** It does not provide age-specific standardized rates, which means it offers less detail compared to direct standardization.

Let us consider a few examples of Standardized Death Rates.

Example: Estimate CDR for the two populations and STDR for population B taking population A as the standard population using the data given in the following table:

Table 1.1

Age-Group (Years)	**A**		**B**	
	Population	**Deaths**	**Population**	**Deaths**

Under 10	10000	500	15000	450
10-20	15000	800	35000	790
20-40	50000	1450	60000	1260
40-60	40000	1250	20000	650
Above 60	10000	600	5000	250

Solution:

Table 1.2

Age-Group (Years)	**Population A**			**Population B**			
	P_x^a	D_x^a	m_x^a	P_x^b	D_x^b	m_x^b	$m_x^b P_x^a$
Under 10	10000	500	50	15000	450	30	300000
10-20	15000	800	53.3333	35000	790	22.57143	338571.4
20-40	50000	1450	29	60000	1260	21	1050000
40-60	40000	1250	31.25	20000	650	32.5	1300000
Above 60	10000	600	60	5000	250	50	500000
$\sum$	125000	4600		135000	3400		3488571

Crude death rate for population A and population B is calculated as-

$$CDR_A = \frac{\sum D_x^a}{\sum P_x^a} \times 1000 = \frac{4600}{125000} \times 1000 = 36.80\; per\; 1000\; population$$

$$CDR_B = \frac{\sum D_x^b}{\sum P_x^b} \times 1000 = \frac{3400}{135000} \times 1000 = 25.18\; per\; 1000\; population$$

Now the standardized death rate for population B taking A as standard population is,

$$STDR_B = \frac{\sum m_x^b \times P_x^a}{\sum {P_x}^a} = \frac{3488571}{125000} = 27.91\; per\; 1000\; population$$

It should be noted that the standardized death rate for population A taking A as the standard population is same the CDR of population A,

$$STDR_A = CDR_A = 36.80$$

Example: Estimate the standardized death rates for the two countries based on the given data:

Table 1.3

Age-Group (Years)	Death Rate per 1000		Standardised Population(in Lakhs)
	Country A	Country B	
0-4	10	3	150
5-14	2	1.5	180
15-24	1.5	1	170
25-34	2	2	200
35-44	3.4	4	100
45-54	6	6	80
55-64	14	3	90
65-74	50	12	40
75 and above	130	100	30

Solution:

Table 1.4

Age-Group (Years) (i)	Death Rate per 1000		Standardised Population (in Lakhs) (iv)	$m_x^a P_x^s (ii) \times (iv)$	$m_x^b P_x^s (iii) \times (iv)$
	Country A (ii)	Country B (iii)			
0-4	10	3	100	1000	300
5-14	2	1.5	180	360	270
15-24	1.5	1	170	255	170
25-34	2	2	200	400	400

35-44	3.4	4	100	340	400
45-54	6	6	80	480	480
55-64	14	3	90	1260	270
65-74	50	12	40	2000	480
75 and above	130	100	30	3900	3000
Total			990	9995	5770

Standardized death rate for country A is, $STDR_A = \frac{\sum m_x^a \times P_x^s}{\sum P^s} = \frac{9995}{990} = 10.09$

Standardized death rate for country B is, $STDR_B = \frac{\sum m_x^b \times P_x^s}{\sum P^s} = \frac{5770}{990} = 5.83$

Calculations are given in table 1.4

Example: Estimate the standardized death rates by direct and indirect methods for the given data:

Table 1.5

AGE	**Standard Population**		**Population A**	
	Population	**Specific Death Rate**	**Population**	**Specific Death Rate**
0-5	8000	50	15000	50
5-15	15000	15	20000	15
15-50	25000	20	25000	10
50 and above	5000	70	10000	60

Solution:

The calculations are shown in table 1.6-

Standardized death rate of country A by direct method is,

$$STDR_A = \frac{\sum m_x^a \times P_x^s}{\sum P^s} = \frac{1175000}{53000} = 22.17$$

And by indirect method is,

$$CDR_A = \frac{\sum m_x^a \times P_x^a}{\sum P^a} = \frac{1900000}{60000} = 31.67$$

Adjustment factor, $\hat{c} = \frac{\sum m_x^s \times P_x^s}{\sum P^s} \times \frac{\sum P_x^a}{\sum m_x^s \times {P_x}^a} = \frac{1475000}{53000} \times \frac{60000}{2250000} = 0.7421$

$$STDR_A = \hat{c} \times CDR_A = 0.7421 \times 31.67 = 23.50$$

Table 1.6

AGE	**Standard Population**			**Population A**				
	P_x^s	m_x^s	$P_x^s m_x^s$	P_x^a	m_x^a	$P_x^a m_x^a$	$m_x^a P_x^s$	$m_x^s P_x^a$
0-5	8000	50	400000	15000	50	750000	400000	750000
5-15	15000	15	225000	20000	15	300000	225000	300000
15-50	25000	20	500000	25000	10	250000	250000	500000
50 and above	5000	70	350000	10000	60	600000	300000	700000
Total	53000		1475000	60000		1900000	1175000	2250000

9.5 Self- Assessment Exercises

1. In a town with a population of 50,000, there were 450 deaths recorded in a year. Calculate the Crude Death Rate (CDR).

2. Two cities, City A and City B, have populations of 100,000 and 150,000, respectively. In a year, City A recorded 850 deaths, and City B recorded 1,200 deaths. Calculate the CDR for both cities and compare them.

3. In a country, the population and total deaths over three consecutive years are as follows:

- Year 1: Population = 1,000,000, Deaths = 9,000
- Year 2: Population = 1,050,000, Deaths = 10,000
- Year 3: Population = 1,100,000, Deaths = 11,500

Calculate the CDR for each year and analyse the trend over the three years.

4. A region has a recorded CDR of 8.5 per 1,000. If the number of deaths in the region was 5,100, estimate the population of the region.

5. In a town, the population increased from 200,000 to 250,000 over a decade. The number of deaths during this time was 1,800. Calculate the CDR at the start and end of the decade and discuss the impact of population growth on the CDR.

6. In a population of 10,000 people aged 40-49, there were 45 deaths recorded in a year. Calculate the Age-Specific Death Rate (ASDR) for this age group.

7. In a city, the following data was recorded:

- Population aged 20-29: 15,000, Deaths: 30
- Population aged 60-69: 8,000, Deaths: 120

Calculate the ASDR for both age groups and compare the rates.

8. The ASDR for the 50-59 age group in a country was as follows over three consecutive years:

- Year 1: Population = 500,000, Deaths = 2,500
- Year 2: Population = 520,000, Deaths = 2,600
- Year 3: Population = 540,000, Deaths = 2,700

Calculate the ASDR for each year and analyze the trend.

9. A region reports an ASDR of 12 per 1,000 for the age group 30-39. If the number of deaths in this age group was 240, estimate the population of this age group.

10. In two different regions, the following data was recorded for the age group 70-79:

- Region A: Population = 20,000, Deaths = 400
- Region B: Population = 15,000, Deaths = 200

Region A has recently improved its healthcare services significantly. Calculate the ASDR for both regions and discuss how healthcare improvements might affect the ASDR.

11. In a town, the population of males aged 50-59 is 10,000, with 50 deaths recorded, while the population of females in the same age group is 9,000, with 30 deaths recorded. Calculate the ASDR for both males and females and compare the results.

12. What do you understand by mortality? Explain different measures of mortality in detail.

9.6 Summary

This unit covers the basic concepts of mortality and its measures, their merits and demerits (if any), standardization methodology etc. It has also explained the trends and levels of mortality in India over the decades and the factors lying behind the decline of mortality over the years. Since, it is important practice to compare different populations on the basis of mortality, both direct and indirect standardization methods have been discussed in this unit.

9.7 References

- Pathak, K. B. and Ram, F. (1998) Techniques of Demographic Analysis, Himalaya Publishing House
- Kanitkar, T. and Bhinde, A. (2011) Principles of Population Studies, Himalaya Publishing House

9.8 Further Readings

- Bogue, D. (1969). Principles of Demography, New York: John Wiley and Sons.
- Yusuf, F., Martins, J. M., Swanson, D. A. (2014). Techniques of demographic analysis, Springer
- Srinivasan, K. (2015). Training Manuals on Demographic Techniques, UNFPA, Delhi.

UNIT - 10 LIFE TABLE

Structure

10.1 Introduction

Life table is statistical model which represents the mortality experience of a cohort. It gives a life history of a cohort of persons dismissed by deaths. Mortality is now commonly represented in the form of life table that also provides probability of death within one year at exact an age. These probabilities are based on tabulation of number of deaths in a specific population and estimates of the size of that population. It combines the mortality rates of a population at different ages into table and gives a more detailed investigation of mortality which is further useful for obtaining estimate of widowhood, chances of marriage and remarriage, population growth and for population projection and other demographic events. Life table columns are also used to indirectly estimate other demographic indicators such as Net Reproduction Rates and Age Specific Fertility Rates.

John Graunt was precursor of modern life table, though more systematic approach was carried by Halley. A life table is a mathematical model for depicting mortality situation experienced by the population. Nathan Keyfitz has defined a life table as "a scheme for expressing the forms of mortality in terms of probabilities." According to Barclay, the life table is a life history of a hypothetical group, or cohort of people, as it is diminished gradually by deaths. The record begins at the birth of each member, and continues until all have died. Separate life tables are made for males and females.

Life table is categorized in two kinds based on the type of initial population namely: cohort life table and period life table. Consider a group of children born in a year. They will all be of exact age 0 in the year they were born; of exact age 1 in the next year and so on. Of course, not all will survive to age 1; some would die between exact ages 0 and 1. If we continue to follow these children throughout their lives until no body remains, we have complete data on their survival status at each exact age from 0 (when all were alive) to the end age of life (when none of the individuals of this group was left alive). These data can be placed on a Lexis diagram-along a diagonal-and the probabilities of dying from one exact age to another can be calculated. These probabilities are the sole basis of the construction of the life table. Such life is called the **cohort (or generation) life table** as we traced the mortality experience of a real cohort (children born in a year).

The other life table is a **cross-sectional (or period or current) life table** in which mortality experience of different generations in a year (or period) is traced. A current life table, therefore, may be viewed as a "snapshot" of current mortality. It is an excellent summary description of mortality in a year or a short period of time. Cross-sectional life tables are said to be tracing the mortality experience of a hypothetical cohort. It should be noted that the data for the cohort life table are not easy to assemble as one will have to follow the original cohort from its origin (birth) until each member of the cohort has died. Such data covers the whole life span of the cohort. The cross-sectional or current life tables, on the other hand, utilize death data for a year (or over a period) from vital registrations and the mid-year population estimates from census enumerations or the postcensal estimates; the data which are easy to assemble. In other words, the current Life Table is based upon the mortality experience of a community for a short period of time such as one year, three years of an intercensal period during which the mortality of a community has not changed substantially. The current life table does not depict the mortality experience of an actual cohort. It assumes a hypothetical cohort which is experiencing certain Age Specific Death Rates observed during a particular time.

10.2 Objectives

After reading this unit, the learner should able to understand about the

- Life table and its types
- Construction of life table
- Abridged life table
- Model life table
- Uses of life table

10.3 Types of Life Table

Complete and Abridged Life Table:

There are two usual ways of presenting a Life Table on the basis of types of age class, namely, complete and abridged life tables, according to the length of the age interval in which the basic data are presented. In a complete Life Table, information is given for every single year of age from birth until the last applicable age. In abridged Life Tables, information is given only for broader age intervals such as x to x + 5 years. The single abridged Life Table is preferred over the elaborated complete life table rather since the abridged is less laborious to prepare and sufficiently reliable for most purposes and often more convenient to use.

Conventionally the life tables are made by single years of age or by five years age group while splitting the age group 0-4 into 0 and 1-4 years. A life table calculated for single years of age is called a complete life table, whereas that based on the five years age grouping an abridged life table. If the death and population data are reliable, the complete life tables are more accurate than the abridged life tables; but there are ways of converting the complete life tables into life abridged tables and vice-versa.

10.3.1 Underlying Assumptions for a Life Table

1. A life table population is assumed to be closed to migration. A life table is supposed to trace the life history of a cohort that means the participation in the cohort cannot be changed due to any other reason such as migration.

2. Annual age specific death rates do not change over time.

3. Annual number of births remain constant over time; the annual number of births chosen is usually 100,000 and accordingly the synthetic population has 100,000 deaths annually over time; thus, the synthetic population is stationary in that it never changes in size.

10.4 Complete Life Table and its Construction

Under the assumptions stated above, let us consider the age of the person in yearly intervals. We begin with a cohort of babies who are all born in the same year. The following notations or relationships are used.

1.) Age (x)= x is the initial age of the group of individuals. In a complete life table x represents the exact age with respect to which the other column values are calculated.

2.) 1_x = The number of survivors at exact age x. l_0 (known as the radix of the life table) will be the size of the original cohort (survivors at an exact age 0). It is usually assumed to be 10,000 or 100,000.

3.) d_x= The number of deaths between two exact ages x and x + 1. It is also the number of deaths at central age x in the life table.

$$d_x = l_x - l_{x+1}$$

4.) q_x= Probability of dying between exact ages x and x + 1.

$$q_x = d_x / l_x$$

5.) a_x= Fraction of life lived by the person between the ages x to x+1. Or average number of years lived in the age interval x to x+1 by those dying during this age interval.

6.) P_x= Probability of survival between exact ages x and x + 1.

$$P_x = l_{x+1} / l_x$$

$$= \frac{l_x - d_x}{l_x}$$

$$= 1 - \frac{d_x}{l_x}$$

$$= 1 - q_x$$

7.) L_x= Number of person-years of exposure to the risk of dying between exact ages x and x+1. Also called as the number of person years lived between the exact ages (x, x+1). By the definition of L_x we can assume that L_x would be sum of two quantities: number of individuals who have survived whole age interval, those individuals will be l_{x+1} and the number of individuals deceased between the exact ages x and x+1. Persons aged (x +1) lived complete one year during the interval x to x+1, and those who died during the interval, d_x, on average they lived for half of the year (on the assumption of the uniform distribution of deaths during the year). Therefore,

$$L_x = l_{x+1}.(1) + d_x.(0.5)$$

$$= l_{x+1} + (l_x - l_{x+1}).(0.5)$$

$$L_x = \frac{l_x + l_{x+1}}{2}$$

$$Or\ L_x = l_x - \frac{d_x}{2}$$

At age 0 the distribution of deaths within the parallelogram is not uniform, only a few of the deaths at age 0 are expected to contribute to persons-years of exposure in the interval x, x + 1. L_0 is calculated in a different way.

$$L_0 = l_1.1 + d_0.f \quad = l_1.1 + (l_0 - l_1).f$$

$$= f.\ l_0 + (1-f).\ l_1$$

Where f is a separation factor.

For the last age, that is the age with indefinite upper age limit, the approximation of L_x is given as- $L_x = \frac{1}{m_{x+}}$

L_x function is also called the life table population.

8.) m_x = Life table death rate between exact ages x and x + 1. This is also known as central death rate values at exact age x.

$$m_x = d_x / L_x$$

9.) T_x= Total number of person-years lived beyond exact age x.

$$T_x = \sum_{x}^{u} L_x$$; u is the upper age at which no body survives.

$$T_x = L_x + T_{x+1}$$

Note that T_0 will be the total number of person-years lived by l_0 persons.

10.) e_x= Expectation of life at exact age x.

$e_x^o = T_x / 1_x$; the expectation of life at birth = T_0/l_0.

The expectation of life at an exact age x gives a summary measure of mortality. It shows that on average, how many years a person can look forward to live having lived to an exact age of x years.

11.) P_x = Survivorship ratio from central age x to x+1.

$P_x = L_{x+1}/L_x$

$P_o = L_0 / l_0$; it is the survivorship ratio from birth to central age 0 .

12.) M_x= Age Specific Death Rate

$= d_x/P_x$

There is another summary measure of mortality based on the life table. It is calculated as the ratio of total deaths to total population in the life table.

Thus,

Life table death rate = l_0/T_0; note that l_0 is total births which is equal to total deaths, and T_0 is total life table population. The rate is then multiplied by 1,000.

The pivotal column of a life table is the q_x column and it needs to be estimated from the observed mortality rates in a given population.

Suppose we have the value of q_x for every x from 0 upwards. We can then start with a suitable cohort – say, one or 100,000 (l_0) births.

Multiplying l_0 by q_0 we get $l_0q_0 = d_0$. Then $l_1 = l_0 - d_0$. Again, $d_1 = l_1 q_1$, $l_2 = l_1 - d_1$, and so on.

Having obtained the values in the l_x column, we can then fill in the other columns, viz. L_x, T_x (for which we start from the bottom of the table and get the values successively by using the relation $\mathbf{T_x = L_x + T_{x+1}}$ and by means of the relations stated above.

Relationship between probability of dying and life table death rate at age x.

We know that-

$$q_x = \frac{d_x}{l_x}$$

$$= \frac{d_x}{L_x} X \frac{L_x}{l_x}$$

Now, since $\frac{d_x}{L_x} = m_x,$

and, $\frac{l_x}{L_x} = \frac{L_x + \frac{d_x}{2}}{L_x} = 1 + \frac{m_x}{2}$

using these values in the expression for q_x –

we get $q_x = m_x \cdot \frac{2}{1+ m_x}$

$$\therefore \; q_x = \frac{2m_x}{1+ m_x}$$

Also $p_x = \frac{2- m_x}{2+ m_x}$

10.4.1 Basic Steps in Construction of a Complete Life Table

The construction of a complete life table can be considered in terms of three broad phases. Firstly, the basic data on deaths and population are checked for inconsistencies, biases and other errors and adjustments are made wherever possible. Secondly the mortality rates are computed and graduated (mathematically smoothed). Thirdly, the remaining functions of life table are computed.

In the computation and graduation of mortality rates, three different age segments are recognized as having their particular problems and are treated separately. The first is the youngest age segments which generally includes ages under five year when the data is sufficiently good and level of mortality is low, the age group (1-4) can be included in the second age segment and the first age segment can be limited to the first year of life only. The second age segment generally course the bulk of table from ages 5 to 9, the third segment relate to the rest of ages i.e. 90 and above.

10.4.1.1. Mortality at Age 0-4

The method used to compute mortality rates at ages 0-4 years is adopted according to the type & quality of data are available. As the quality of data improves there is a tendency to extent the age group 1-4. In the second age segment, owing the distinct nature of mortality during the first year of life, special procedures must be used to compute the mortality rate for first year of life.

10.4.1.2 Mortality at Ages 5 to 89

The main segment of the life table is that covering ages 5 to 89. for this range we usually have quite reliable data. However, it is necessary that various tests be performed to establish how satisfactory the data are and to make adjustment whenever necessary and possible. Once the data have been adjusted, the calculation of mortality rates for single year of age becomes a problem in graduation technique. The problem of computing the mortality rates is therefore, reduced to calculate these rates which are smoother than the observed rates. A large number of graduation methods have been used in the past and new ones are constantly being derived to make advantage of more in electronic computers.

10.4.1.3 Mortality at Ages 90 and above (85 & above)

Population and death data at older ages have always been considered to below accuracy. In practice, it has been found that regardless of total volume of data and their accuracy, there is some points at the older ages beyond which arbitrary methods must be applied. At these very old ages that data becomes either too scanty (very small) to the reliable or they are regarded as unreliable because of errors in age reporting. For practical purposes any reasonable method can be used since the effect of change in the mortality rates or the life table functions at younger ages would be relatively small. One method used is to assume that mortality rate increases at the older ages by about the same percentages as that found at the end of the main portion of the life table. Other method is to adopt a series of values based on mortality experience of the other population which are known to be more accurate.

Example: Let us consider a hypothetical cohort of size 100,000 for which the distribution of q(x) is given below. We have calculated the values of the remaining components from the formula given in section-2.3

Table 2.1 Construction of Life Table Components

Age	l(x)	d(x)	q(x)	L(x)	T(x)	m(x)	E(x)
0	100000	21898	**0.21898**	89051	366986	0.24590403	3.7
1	78102	27222	**0.34854**	64491	277935	0.42209923	3.6
2	50880	2605	**0.05119**	49578	213444	0.0525345	4.2
3	48276	860	**0.01782**	47846	163866	0.01797452	3.4
4	47415	726	**0.01531**	47052	116021	0.01542974	2.4
5	46689	1065	**0.02281**	46157	68969	0.02307367	1.5
6	45624	1193	**0.02615**	22812	22812	0.05229704	0.5

The calculations are explained as follows-

- d(0)= l(0) x q(0) = 100000 x 0.21898 = 21898
- l(1)= l(0)- d(0) = 100000 – 21898 = 78102
- similarly, l(2)= l(1) - d(1) =78102 – 27222 = 50880
- L(0) = [l(0) +l (1)]/2 = (100000+78102)/2 = 89051
- For L(6), we use approximation, L(6)= l(6)/2 = 22812
- T(0)= L(0)+L(1)+L(2)+L(3)+L(4)+L(5)+L(6)= 366986

- T(1)= L(1)+L(2)+L(3)+L(4)+L(5)+L(6)=277935 and so on
- m(0)= d(0)/L(0)= 21898/89051 = 0.24590
- e(0)= T(0)/l(0) = 366986/ 100000 = 3.7 years

As observed after completing the life table, we can see that the highest observed life expectancy is at age 2, among all the ages.

10.5 Abridged Life Table and Its Construction

In many cases, in which a life table is necessary, the degree of accuracy or detail needed is not so great as that provided by complete life table. In such cases an abridged life table serves the purpose. The assumptions for a complete life table and abridged life table is same.

The columns of an abridged life table are similar to those in the complete life table. We have the analogous functions l_x, ${}_nq_x$, ${}_nd_x$, ${}_nl_x$, T_x, and e_x where 'n' represents the length of age interval.

An Abridge Life Table is calculated in the same way as the complete Life Table but the formula allow for the intervals of age 0-1,0-4,5-9,10-14,…………..,85$^+$. The columns of the life table have the same name and meaning as complete life table. Formula for different life table columns is given below-

1.)

$$l_{x+n} = l_x - {}_nd_x$$

Or $l_{x+n} = \ l_x - nd_x$ For x=5,10,………

2.)

$$ {}_nq_x = \frac{{}_nd_x}{l_x} $$

$$ = \frac{l_{x+n}}{l_x} $$

$$ {}_np_x = 1 - {}_nq_x $$

3.)

$$ {}_nL_x = \frac{n}{2}\left[l_x + l_{x+n}\right] $$

Several approximations are used for L_{85+}

$$ {}_\infty L_{85} = \frac{n}{2}.l_{85} $$

$$ {}_\infty L_{85} = l_{85}.\log_{10}(l_{85}) $$

4.) Relationship of ${}_nq_x$ with ${}_nM_x$

$$ {}_nq_x = \frac{2n.{}_nM_x}{2 + n.{}_nM_x} $$

There are a number of methods for computation of abridged life tables.

The following table 2.2 shows an example of abridged life table of a hypothetical cohort of 100000 of West African Males, by United Nations

Table 2.2

Age	M(X)	Q(X)	l(X)	D(X)	L(X)	T(X)	E(X)	A(X)
0	0.30078	0.25033	100000.	25033.	83228.	2499992.	25.000	0.330
1	0.14122	0.41115	74967.	30822.	218250.	2416764.	32.238	1.352
5	0.01335	0.06460	44144.	2852.	213593.	2198514.	49.803	2.500
10	0.00450	0.02223	41293.	918.	204170.	1984921.	48.069	2.500
15	0.00381	0.01888	40375.	762.	200000.	1780752.	44.105	2.541
20	0.00568	0.02804	39613.	1111.	195396.	1580751.	39.905	2.599
25	0.00649	0.03195	38502.	1230.	189511.	1385355.	35.981	2.563
30	0.00819	0.04017	37272.	1497.	182761.	1195844.	32.084	2.597
35	0.01123	0.05467	35774.	1956.	174148.	1013082.	28.319	2.585
40	0.01375	0.06654	33819.	2250.	163626.	838934.	24.807	2.571
45	0.01808	0.08662	31568.	2734.	151217.	675308.	21.392	2.578
50	0.02393	0.11309	28834.	3261.	136241.	524091.	18.176	2.569
55	0.03196	0.14835	25573.	3794.	118684.	387850.	15.166	2.580
60	0.04834	0.21615	21779.	4707.	97389.	269166.	12.359	2.556
65	0.06766	0.28948	17072.	4942.	73045.	171777.	10.062	2.508
70	0.09894	0.39395	12130.	4779.	48296.	98732.	8.140	2.415
75	0.12093	0.45871	7351.	3372.	27885.	50436.	6.861	2.369
80	0.17645	******	3979.	3979.	22551.	22551.	5.667	5.667

Source: UN Model Life Tables for developing countries. Department of International Economic and Social Affairs Population Studies, No. 77

10.5.1 Computations of Components of Abridged Life Table

Over time various assumptions for the computation of life table functions are made by different researchers. The main problem lies in the conversion of central death rates values into probability of dying. For any specific population, the only information usually provided is the age specific death rates. there are various approximations of ${}_nM_x$ given by different demographer which are discussed below.

10.5.1.1 Reed Merrell Method

Reed and Merrell (1939) method is one of the most frequently used short cut procedure for computing an abridged life table. In this method, the mortality rates are obtained from a set of standard conversion tables showing the mortality rates associated with various observed age specific death rates. (observed central death rates). The standard takes for ${}_3m_x$, ${}_5m_x$, ${}_{10}m_x$ have been prepare in the assumption that following relation holds

$${}_nq_x = 1 - e^{-n \cdot {}_nM_x - an^3 \cdot {}_nM_x{}^2}$$

Where, n is the size of the age interval $_nM_x$ is the age specific death rate, and a is a constant. Reed and Merrell found that a value of a equal to .008 would produce acceptable results.

The conversion of $_nM_x$ to $_nq_x$ by use of Reed Merrell tables is usually applied to five year or ten years data. But special age groups are employed at both ends of the life tables. At younger ages the most frequently used grouping is, ages under 1 and 1 to 4.

Once the mortality rates have been obtained the construction of abridged life table continues with the computation of every entry in the survival column and the death column by using the relationship

$$l_x + n = (1 - {}_na_x)\, l_x$$

$$_nd_x = l_x - l_{x+n}$$

In the Reed - Merrell method firstly T_x values are directly determined from l_x's for ages five and over or 10 and over by using the following equations.

i. If the age interval taken is 5- years.

$$T_X = -0.2833 l_x - 5 + 2.5 l_x + 0.20833 l_{x+5} + 5 \sum_{\alpha=1}^{\infty} l_X + 5\alpha$$

ii. If the age interval taken is 10- years.

$$T_X = 4.16667 l_x + 0.833 l_{x+10} + 10 \sum_{\alpha=1}^{\infty} l_X + 10\alpha$$

According to Reed and Merrell, for ages under 10, values of L_x values may be determined using below relations-

- $L_0 = 0.276\,(l_0) + 0.724\,(l_1) \approx L_0 = 0.3 l_0 + 0.7 l_1$
- $L_1 = 0.410\,(l_1) + 0.590\,(l_2)$
- $_4L_1 = 0.034\,(l_0) + 1.184\,(l_1) + 2.782\,(l_5)$
- $_3L_2 = -0.021\,(l_0) + 1.384\,(l_2) + 1.637\,(l_5)$

- ${}_5L_5 = -0.003(l_0) + 2.242\ (l_5) + 2.76\ (l_1)$

For others,

$$ {}_nL_x = \frac{n}{2}\left(l_x + l_{x+n}\right) $$

And for last age -

$$ {}_\infty L_x = \frac{l_x}{{}_\infty M_x} $$

${}_nL_x$ for ages 10 and above may be computed by differentiating the T_x values and e^x is computed in the same way as for a complete life table. $e^x = T_x/l_n$.

10.5.1.2 Greville's Method

A method suggested by Greville converts the Age Specific Death Rates into mortality rates

$$ {}_nq_x = \frac{{}_nM_x}{\frac{1}{n} + {}_nM_x\left[\frac{1}{2} + \frac{n}{12}\left({}_nM_x - \log_e c\right)\right]} $$

with the use of the formula

Where,

- ${}_nM_x$ represents the Age Specific Death Rate,
- n represents the length of the interval.
- ${}_nq_x$ represents the correspondence mortality rate and
- c is a constant, which comes from the assumption that ${}_nM_x$ follow an exponential curve.

Empirically the values of c have been found to be between 0.08 and 0.10 and hence $\log_e c$ could be assumed to be about 0.095. Thus, knowing the ${}_nM_x$ values and ${}_nq_x$ values can be obtained by using the above relationship.

In the Greville's method the central death rates in the life table and in the population is assumed to be the same. Hence ${}_nL_x$ values are calculated by the use of formula

$${}_nL_x = {}_nd_x / {}_nM_x$$

Also,

$${}_nL_x = \frac{n}{2}\left(l_x + l_{x+n}\right) + \frac{n}{24}\left({}_nd_{x+n} - {}_nd_{x-n}\right)$$

and for the last age interval it is computed by

$${}_\infty L_x = \frac{l_x}{{}_\infty M_x}$$

knowing the ${}_nL_x$ values T_x values may be obtained easily and the e_x values may be computed by using the relationship $e^x = T_x/l_x$

10.5.1.3 King's Method

The first step in the construction an abridged life table by King's method is to complete mortality rates at pivotal ages by using appropriate formula. Then we compute $P_x = 1 - q_x$. In order to proceed to the next life table function l_x, it is necessary to estimated ${}_5P_x$ (considering 6 intervals). Thus, to evaluate l_{x+5} we use the relationship $l_{x+5} = l_x \cdot {}_5P_x$

$$\text{Or, } \log l_{x+5} = \log l_x + \log ({}_5P_x)$$

Thus, it is necessary to estimate log 5px from the available values of px at the pivotal ages. For the first pivotal age log 5px is evaluated. Form the Newton forwarded formula as follows. Ignoring differences higher than the third, we have

$$\log p_{x+1} = \log p_x + .2\Delta \log p_x - .08\,\Delta^2 \log p_x + .048\,\Delta^3 \log p_x.$$

$$\log p_{x+2} = \log p_x + 4\Delta \log p_x - .12\Delta^2 \log p_x + .064\,\Delta^3 \log p_x$$

$$\log\ p_{x+3} = \log p_x + .6\Delta \log p_x - 12 \log\Delta^2 p_x + .56\log \Delta^3 p_x$$

$$\log p_{x+4} = \log p_x + .8\Delta \log p_x - .08\Delta^2\log_x + .032\Delta^3\log p_x$$

Hence, we get

$$\begin{aligned}\log\ {}_5p_x &= \log(p_x\ p_{x+1}\ p_{x+2}\ p_{x+3}\ p_{x+4}) \\ &= \sum_{i=0}^{4} \log\ p_{x+i} \\ &= 5\log p_x + 2\Delta \log p_x - .4\Delta_2 \log p_x + .2\Delta^3 \log p_x \\ &= 2.4\log p_x + 3.4\log p_{x+5} - \log p_{x+10} + .2\log p_{x+15}\end{aligned}$$

for the remaining pivotal ages, the Newton's forward formula based on log p_{x-5} is used. We get

$$\begin{aligned}\log p_x &= \log p_{x-5} + \Delta\log p_{x-5} \\ \log p_{x+1} &= \log p_{x-5} + 1.4\Delta\log p_{x-5} + .12\Delta^2\log p_{x-5} - .032\Delta^3\log p_{x-5} \\ \log p_{x+2} &= \log p_{x-5} + 1.4\Delta\log p_{x-5} + .28\Delta^2\log p_{x-5} - .56\Delta^3\log p_{x-5} \\ \log p_{x+3} &= \log p_{x-5} + 1.6\Delta\log p_{x-5} + .48\Delta^2\log p_{x-5} - .064\Delta^3\log p_{x-5} \\ \log p_{x+4} &= \log p_{x-5} + 1.8\Delta\log p_{x-5} + .72\Delta^2\log p_{x-5} - .048\Delta^3\log p_{x-5}\end{aligned}$$

Thus, we get

$$\begin{aligned}\log_5 p_x &= 5\log p_{x-5} + 7\Delta\log p_{x-5} + 1.6\Delta^2\log p_{x-5} - .2\Delta^3\log p_{x-5} \\ &= -.2\log p_{x-5} + 3.2\log p_x + 2.2\log p_{x+5} - .2\log p_{x+10}\end{aligned}$$

Having obtained the above values, we find the following sum

$$\lceil N'_{x,5} \rceil = \sum_{i=1}^{5} l_{x+i}$$

for each pivotal age x these sums are obtained in the same manner as desired above. Hence, the formula corresponding to first pivotal age is-

$$\left| N'_{x,5} \right| = 5l_x + 3\Delta l_x - .4\Delta^2 l_x + .2\Delta^3 l_x$$
$$= 1.4l + 4.4l_{x+5} - l_{x+10} + .2l_{x+15}$$

For the remaining pivotal ages, the formula becomes

$$\left| N'_{x,5} \right| = 5l_{x+5} + 8\Delta l_{x-5} + 2.6\Delta^2 l_{x-5} - .2\Delta^3 l_{x-5}$$
$$= -2l_x - 5 + 2.2l_x + 3.2\, l_{x+5} - .2l_{x+10}$$

In some cases, the formula may give negative value of $\left| N'_{x,5} \right|$ (specially at very old ages). In which case the value may be taken as zero. Then we compute,

$$N'_x = \sum \left| N'_{x,5} \right|$$
$$= \sum_{i=1}^{\infty} l_{x+i}$$

From this, the value of e_x, is evaluated by using the formula

$$e_x = .5 + \frac{N'_x}{l_x}$$

since

$$\frac{T_x}{l_x} = \frac{1}{l_x} \left[L_x + L_{x+1} + L_{x+2} . + ... \right]$$
$$= \frac{1}{l_x} \left[\frac{l_x + l_{x+1}}{2} + \frac{l_x + 1 + l_{x+2}}{2} + ... \right]$$
$$= \frac{1}{l_x} \left[\frac{l_x}{2} + l_{x+1} + l_{x+2} + ... \right]$$
$$= \frac{1}{l_x} \left[\frac{l_x}{2} + N'_x \right]$$
$$= .5 + \frac{N'_x}{l_x}$$

$$e_x = \frac{N'_x}{l_x} = e_x$$

Hence is known as curate expectation of life

i.e $e_x = e_x^0 - \frac{1}{2}$

10.6 Model Life Tables

It has been observed that the risk of death experienced by different ages & sexes segment of the population are often interrelated. It implies that if the death rates in some segments are relatively high then the normal expectations, then the death rate will be also higher in other segments of the population. It is easily understood by the fact that when health conditions are especially good or poor for one section of the society or segment of a population then the conditions tend to be good or poor for other segments as well. The result of this tendency would be used to estimate mortality of other sections using mortality for one section. It would be possible to construct a set of life table under mortality conditions ranging from the highest to the lowest mortality levels observed in the human population and a life table appropriate to a given population can then be chosen (with interpolation if necessary) from these set of life tables. Hence, a typical life table giving set of probabilities of dying at different ages corresponding to a given level of mortality (e_0^0) has been called Model Life Table.

The model life tables are useful to visualize the change in mortality in future. There are four sets of Model Life tables are in common use:

1.) Coale and Demeny's Regional Model Life Tables
2.) The United Nations Model Life Tables
3.) Lender man's system of Model Life Tables and
4.) The Logit System of Model Life tables.

10.6.1 Coale and Demeny Regional Model Life Tables

Ansley Coale and Paul Demeny produced Life Tables in 1966 known as Regional Model Life Tables consisting of 4 sets of model life tables, labelled as West, East, North and South. They analysed data from 192 life tables in total for males and females. Each set contains 24 tables calculated for males and females separately, at mortality levels designated as (1) to (24) with equal spacing of the values of the expectation of life at birth. For females, life expectancy at birth (e^0_0) was 20 years at level (1) and was 77.5 at level (24). Coale and Demeny identified for distinct mortality patterns corresponding to geographical areas of Europe. One pattern corresponded to Northern European countries, second to Southern European countries, third to Eastern and fourth to more heterogenous countries from Western Europe. The West model, based on 125 life tables from over 20 countries including Canada, United States, South Africa, Israel, Japan, Taiwan and a number of countries of West Europe, has a mortality pattern of an average world pattern and hence can be used as an overall one parameter average patter of mortality. The North set shows a relatively low old age mortality, and also low infant mortality as compared with rates at ages 1 to 4. Coale and Demeny (Coale & Demeny 1983) revised and have extended their mortality level to 25, with maximum value of life expectancy of females and males as 80 years and 76.65 years respectively. They are used by many demographers for the populations where accurate data is available for one or two censuses and the population can be assumed to be stable with negligible or no migration.

10.6.2 The United Nations Model Life Tables

The United Nations has published two sets of model life tables for the developing and under developed countries in 1982 respectively. The tables published in 1982 were prepared on the basis of 158 life tables collected from a wide selection of countries and representing different periods of time. The new set of UN Model Life tables have been extended up to 10 years of age for females and males.

United Nations (1982) has given some sets of age – sex patterns of mortality which are based on reliably documented developing country data and hence are expected to describe better the age pattern of mortality in the developing world. First, the different age patterns of mortality were stratified into clusters by graphical and statistical procedures, each cluster having a typical average age pattern of mortality. By fitting through the method of least squares, a second-degree parabola of the type $y = a + bx + cx^2$, a relationship developed between the mortality rates of the successive age groups.

The new set of UN Model Life tables given in 2011 have been extended up to 10 years of age for females and males. A new set of extended model life table by Gerland and Li imposed constraints to ensure accuracy at high levels of life expectancy at birth. The four sets of regional model life tables by Coale and Demeny and five UN model life table regional pattern are extended up to 100 years of life expectancy at birth, blended with the existing values to provide better estimates by age, sex and life expectancy at birth. A short sample of a complete life table under UN model life table system is given below. This life table has extended estimates for life table function till 130 years of age. The model life tables are used for making mortality assumptions for population estimation and projections as well as for smoothing out irregularities in the age-specific mortality rates to construct a life table for countries which have poor quality data. Model life tables can be generated for different region by using the software MortPack developed by the United Nations for finding various direct and indirect measures of demographic indicators.

Fig: 2.1 UN Model Complete Life Table for South East Asia Region for age 0-30 years

Age	m_x	q_x	l_x	d_x	L_x	T_x	s_x	e_x	a_x
0	0.35438	0.28803	100000	28803	81278	2000000	0.78797	20.00	0.350
1	0.19949	0.17945	71197	2777	64045	1918722	0.85348	26.95	0.440
2	0.12577	0.11768	58420	75	54661	1854678	0.90239	31.75	0.453
3	0.08338	0.07979	51546	113	49325	1800017	0.93304	34.92	0.460
4	0.05712	0.05543	47433	2629	46023	1750691	0.95365	36.91	0.464
5	0.03904	0.03825	44804	1714	43890	1704668	0.96761	38.05	0.467
6	0.02768	0.02728	43090	1175	42468	1660779	0.97652	38.54	0.471
7	0.02039	0.02017	41915	846	41471	1618311	0.98222	38.61	0.475
8	0.01584	0.01571	41069	645	40734	1576840	0.98583	38.39	0.480
9	0.01289	0.01280	40424	518	40156	1536106	0.98831	38.00	0.483
10	0.01078	0.01072	39907	428	39687	1495950	0.98998	37.49	0.486
11	0.00950	0.00946	39479	373	39289	1456263	0.99086	36.89	0.492
12	0.00900	0.00895	39106	350	38930	1416974	0.99091	36.23	0.499
13	0.00940	0.00936	38755	363	38576	1378043	0.99008	35.56	0.506
14	0.01065	0.01060	38393	407	38194	1339467	0.98855	34.89	0.511
15	0.01241	0.01234	37986	469	37756	1301274	0.98684	34.26	0.510
16	0.01404	0.01394	37517	523	3259	1263517	0.98541	33.68	0.508
17	0.01531	0.01519	36994	562	36716	1226258	0.98431	33.15	0.505

18	0.01628	0.01615	36432	588	36140	1189542	0.98343	32.65	0.503
19	0.01712	0.01697	35844	608	35541	1153402	0.98266	32.18	0.502
20	0.01784	0.01768	35235	623	34925	1117861	0.98204	31.73	0.501
21	0.01838	0.01821	34612	630	34297	1082937	0.98163	31.29	0.500
22	0.01866	0.01849	33982	628	33667	1048639	0.98149	30.86	0.499
23	0.01867	0.01850	33354	617	33044	1014972	0.98158	30.43	0.498
24	0.01849	0.01832	32737	600	32435	981928	0.98183	29.99	0.497
25	0.01820	0.01803	32137	579	31846	949492	0.98207	29.54	0.497
26	0.01801	0.01785	31558	563	31275	917647	0.98218	29.08	0.498
27	0.01799	0.01783	30994	553	30718	886372	0.98208	28.60	0.499
28	0.01820	0.01804	30442	549	30167	855654	0.98177	28.11	0.500
29	0.01862	0.01845	29893	552	29617	825487	0.98129	27.61	0.501
30	0.01916	0.01898	29341	557	29063	795869	0.98081	27.12	0.500

10.6.3 Lederman's Model Life Table

Ledermann (1969) published 7 sets of model life tables based on 154 life tables that had greater geographical coverage. This basic method is again a regression model of the type the seven tables are based on seven different measures of mortality such as e_0^3, q_0, ${}_5q_0$, ${}_{15}q_0$, ${}_{20}q_{30}$, ${}_{20}q_{35}$, and m_{50+}.

$$\log {}_nq_x = a_{x_0} + a_{x_1} \log {}_nq_j$$

and

$$\log {}_nq_x = b_{x_0} + b_{x_1} \log {}_nq_i + b_{x_2} \log {}_nq_j$$

Thus, one has a choice of using the most easily available or most reliable piece of information about mortality as the basis of regression estimates of the rest of the life table.

10.6.4 The Logit System of Model Life Tables

An empirical study of the relationship between the logits of 1_x values of different life tables showed this relationship to be linear. This prompted Brass to relate the logit of observed $1_{(x)}$ to the logit of same $1_{(x)}^s$ of a standard life table by the equation

Where logit $l_{(x)} = \alpha + \beta \, . \, \text{logit } l_x^s$

or $y_{(x)} = \alpha + \beta \, . \, y^s_{(x)}$

where logit $l_{(x)} = \frac{1}{2} \log e \frac{l_{(x)}}{1 - l_{(x)}}$

and $-\text{logit } l_{(x)} = \text{logit } [1 - l_{(x)}]$

and $y_{(x)}$ = logit $l_{(x)}$ of population, $y_{(x)}^s$ is the logit of l_x^s the function of standard population. If α and β are known, the table of the study population can be derived with the help of standard life tables.

The parameters α and β may be estimated from the reliable estimates of childhood survival, p(2) and an overall estimate of adult mortality in the form of p(30) or p(50) given the accepted value of p(2). The values of α and β can then be estimated using

$$\text{logit } p(2) = \alpha + \beta \text{ logit } p^s(2)$$

$$\text{logit } p(50) = \alpha + \beta \text{ logit } p^s(50)$$

Apparently the logit system of model life tables are better than the earlier existing models, but its use in the developing countries is quite limited because of unreliable estimates of p(2) and p(5).

10.7 Uses of Life Table

Although the primary purpose of a life table is to present a clear picture of mortality of a population, it may be put to other important uses also. These are:

1.) Life tables are used by life insurance companies in determining rates of premium for policies of persons of different ages. In fact, here the probabilities of survival of a person of given age are computed up to different ages and accordingly the premium rate is computed.

2.) Government or a firm also utilizes life tables for the determination of rates of retirement benefit for its employees.

3.) Life tables are also used for population projection for future dates.

4.) Life tables are used in computation of net reproduction rate and probability of widowhood, orphan hood, etc.

5.) In demography, life table are used for measuring mortality are used in the study of fertility, reproductivity, migration and population structure. It is widely used in the estimation and projection of population size, structure and change at a future date.

6.) Expectancy of persons marrying or remaining singles

7.) Expectancy of persons admitted to a mental hospital

8.) Expectancy table for school going population with drop-out rates

9.) Expectancy of working force life table.

10.8 Assorted Examples

Example 10.1 In a population, if the number of males at the age 55 was 10,53 and the probability of dying between ages 55 and 56 was 0.01 then after one year how many males will reach an age 56?

Here

$$\begin{aligned} l_{55} &= 1056 \\ q_{55} &= 0.01 \\ d_{55} &= l_{55} \times q_{55} = 1056 \times .01 \\ &= 10.56 \cong 11 \end{aligned}$$

Therefore, the number of males at age 56 will be

$$\begin{aligned} l_{56} &= l_{55} - d_{55} = 1056 - 11 \\ &= 1046 \end{aligned}$$

Example 10.2 In a sample survey of a locality number of males between ages 45 and 46 were 30,450 and 30,320 respectively. Calculate q_{45}

Here

$$l_{46} = 30320 \textit{ and } l_{45} = 30450$$

$$\therefore d_{45} = 30450 - 30320$$

$$= 130$$

$$\therefore m_{45} \cong \frac{dx^*}{\frac{l_x + l_x + 1}{2}} = \frac{130}{\frac{30450 + 30320}{2}}$$

$$= \frac{130}{30385} = .0046$$

$$\therefore q_{45} = \frac{2 \times .0046}{2 + .0046} = \frac{.0092}{2.0046}$$

$$\cong .00458 \ \textit{Ans.}$$

(* Here D_x is approximated by d_x)

10.9 Self-Assessment Exercises

Q1. Fill in the blanks:

(i) In a complete Life Table, the information iss given for __________, however, in an Abridged Life Table, these functions are computed for ________________ .

(ii) A Life Table is based on the assumption that the cohort is closed against ______________________

(iii) The cohort originates with some standard number of births that is called___________________ of the life table.

(iv) The column q_x represents _____________ between ages x to x + 1.

(v) T_x represents total number of years lived beyond age ___________ and is symbolically written as T_x= ______________ + L_x.

(vi) e^o_x represents expectations of life at ______________and is written as e^o_x _____________/ l_x.

Q2. A part of a complete life table is given below with missing entries in the rest of the columns. Calculate the missing values and complete the life table. Also find the probability-

(a) that a child of age 1 will live at least 7 years more,

(b) that out of two children ageing 4 and 5, at least one will die within 9 years.

Age x	l_x	d_x	q_x	L_x	m_x	T_x	e_x^o
0	100,000		0.435				
1			0.541				
2			0.680				
3			0.733				
4			0.724				
5			0.801				
6			1.058				
7			1.002				
8			0.921				
9			1.013				

Q3. Give a description of complete life table and discuss the relationships among various columns of a life table.

Q4. Write uses of life tables.

Q5. If the values of q_x columns are known, how other columns can be computed.

Q6. Explain the construction of Abridged life table using Reed and Merrell and Greville's method of approximations for life table functions.

10.10 Summary

Life table is one of the most important tools of demography to estimate mortality and various other vital demographic indicators. Over the time, different life tables have been constructed with increasing level of precision. We have various life tables for different regions, sex and mortality level. In this unit, we have thoroughly explained the concept of life table and how to construct complete as well as abridged life table. We have discussed different model life tables developed in the history, their relevance and their applications.

10.11 References

- Yusuf, F., Martins, J. M., Swanson, D. A. (2014) Techniques of Demographic Analysis, Springer
- Srinivasan, K. (2015). Training Manuals on Demographic Techniques, UNFPA, Delhi.
- Keyfiitz, N. (1968). *Introduction To the Mathematics Of Population*, Reading, Mass: Addison-Wesley.
- United Nation Model Life Tables For Developing Countries: Https://Www.Un.Org/Development/Desa/Pd/Sites/Www.Un.Org.Development.Desa.Pd/Files/Files/Documents/2020/Jan/Un_1982_Model_Life_Tables_For_Developing_Countries.Pdf
- Life Tables for The United States Social Security Area 1900-2100
- SRS Based Abridged Life Tables 2016-20, Census Of India, Office of Registrar General of India

10.12 Further Readings

- Yusuf, F., Martins, J. M., Swanson, D. A. (2014) Techniques Of Demographic Analysis, Springer

- Pathak, K. B. and Ram, F. (1998) Techniques of Demographic Analysis, Himalaya Publishing House