Edition - I

STATISTICAL GROWTH MODELS

Dr. R. Panduranga Rao
Dr. Y.V. Seshagiri Rao
Dr. Y.S.N. Satyanarayana
Dr. Kalyan Kumar Palaparthi

www.sahasrapublications.com
info@sahasrapublications.com

Dr. R. Panduranga Rao awarded Ph.D. from Jawaharlal Nehru Technological University, Hyderabad. He is currently working as Associate professor in Vignan Institute of Technology and Science. (Autonomous) He had 23 years of Teaching Experience in various Engineering Institutions in and around Hyderabad. He had published more than 15 publications in reputed National and International Journals with Scopus and Quartile indexing along with 3 patents. His Area of specialization includerough sets relates applications in Artificial Intelligence, Data Mining and Machine Learning.

Dr. Y.V. Seshagiri Rao, has been associated with Vignan Institute of Technology and Science (Autonomous) as Associate professor of Mathematics in Basic Sciences and Humanities department. He was awarded with Ph.D. degree from Jawaharlal Nehru Technological University, Hyderabad. He has 22 years of teaching experience and published 22 research publications in various National, International, Scopus and SCI journals, participated in various workshops and has three patents. Also he worked as various positions in Academics as Head of the department. His area of interests are Lattice Theory, Fluid Mechanics.

Dr. Y.S.N. Satyanarayana awarded Ph.D. from Acharya Nagarjuna University, Guntur. He has more than two decades of teaching experience for both UG and PG course. At present working as an Associate professor in Malla Reddy Engineering College (MREC-Autonomous) in Humanities and Science department. Hyderabad. He published 10 research papers in peer reviewed international journals and got one patent, participated in national and International conferences worked as head of the department. His Area of interest probability & Statistics, Data Science. He is life member in ISTE and APSMS.

Dr. Kalyan Kumar Palaparthi awarded a Ph.D. from Jawaharlal Nehru Technological University, Hyderabad. He has been an Associate professor for almost 20 years, in and around Hyderabad, at various engineering Colleges. His papers published in more than 15 recognized National and International journals in Q3 and Q4 indexed Scopus Journals. He is currently working as an Associate Professor in KL (Deemed to be University), Hyderabad. His specialization includeNano-particle Fluidics. He is now working in the area of data science, artificial intelligence, and machine learning.

Sahasra Publications
Researchers Connect
www.sahasrapublications.com
Email: info@sahasrapublications.com

ISBN 978-81-961629-9-3
9 788196 162993

STATISTICAL GROWTH MODELS

Edition-I

Dr. R. Panduranga Rao
Associate Professor of Mathematics
Vignan Institute of Technology and Science
(Autonomous)

Dr. Y.V. Seshagiri Rao
Associate Professor of Mathematics
Department of Basic Sciences and Humanities
Vignan Institute of Technology and Science
(Autonomous)

Dr. Y.S.N. Satyanarayana
Associate Professor
Department of Humanities and Science
Malla Reddy Engineering College
(MREC-Autonomous), Hyderabad

Dr. Kalyan Kumar Palaparthi
Associate Professor
KL (Deemed to be University)
Hyderabad, Telangana

Published By

Sahasra Publications

203, Park view, Pushpak Co-op Housing Society

Prashanthi Nagar Industrial Estate

Kukatpally, Hyderabad, Telangana, India

Email: info@sahasrapublications.com

Author(s)

Dr. R. Panduranga Rao
Associate Professor of Mathematics
Vignan Institute of Technology and Science
(Autonomous)

Dr. Y.V. Seshagiri Rao
Associate Professor of Mathematics
Department of Basic Sciences and Humanities
Vignan Institute of Technology and Science
(Autonomous)

Dr. Y.S.N. Satyanarayana
Associate Professor
Department of Humanities and Science
Malla Reddy Engineering College (MREC-Autonomous). Hyderabad

Dr. Kalyan Kumar Palaparthi Associate Professor
KL (Deemed to be University) Hyderabad, Telangana

Book Title: Statistical Growth Models

Publication Year: 2023

ISBN: 978-81-961629-9-3

DOI: 10.58673/sp.2023.10.18.2

Price: Rs. 780/-

ABOUT AUTHORS

Dr. R. Panduranga Rao awarded Ph.D. from Jawaharlal Nehru Technological University, Hyderabad. He is currently working as Associate professor in Vignan Institute of Technology and Science. (Autonomous) He had 23 years of Teaching Experience in various Engineering Institutions in and around Hyderabad. He had published more than 15 publications in reputed National and International Journals with Scopus and Quartile indexing along with 3 patents. His Area of specialization includerough sets relates applications in Artificial Intelligence, Data Mining and Machine Learning.

Dr. Y.V. Seshagiri Rao, has been associated with Vignan Institute of Technology and Science (Autonomous) as Associate professor of Mathematics in Basic Sciences and Humanities department. He was awarded with Ph.D. degree from Jawaharlal Nehru Technological University, Hyderabad. He has 22 years of teaching experience and published 22 research publications in various National, International, Scopus and SCI journals, participated in various workshops and has three patents. Also he worked as various positions in Academics as Head of the department. His area of interests are Lattice Theory, Fluid Mechanics.

Dr. Y.S.N. Satyanarayana awarded Ph.D. from Acharya Nagarjuna University, Guntur. He has more than two decades of teaching experience for both UG and PG course. At present working as an Associate professor in Malla Reddy Engineering College (MREC-Autonomous) in Humanities and Science department. Hyderabad. He published 10 research papers in peer reviewed international journals and got one patent, participated in national and International conferences worked as head of the department. His Area of interest probability & Statistics, Data Science. He is life member in ISTE and APSMS.

Dr. Kalyan Kumar Palaparthi awarded a Ph.D. from Jawaharlal Nehru Technological University, Hyderabad. He has been an Associate professor for almost 20 years, in and around Hyderabad, at various engineering Colleges. His papers published in more than 15 recognized National and International journals in Q3 and Q4 indexed Scopus Journals. He is currently working as an Associate Professor in KL (Deemed to be University), Hyderabad. His specialization includeNano-particle Fluidics. He is now working in the area of data science, artificial intelligence, and machine learning.

CONTENTS

CHAPTER - 1 INTRODUCTION

1.1 GENERAL INTRODUCTION

Modeling of the growth is the heart of various fields of applied statistics such as Biometrics, Econometrics, Demo metrics, Business and Industrial Statistics, Time series and forecasting. Growth model methodology has been widely used in the modeling of various research problems in these fields. In recent years, in almost all applied fields of statistics, a great deal of research has been directed to either the mathematical or stochastic modeling of growth and establishing the functional relationships among different characteristics by fitting various linear or nonlinear growth models.

A large number of mathematical and statistical growth models have been developed in the literature and successfully applied to different situations in the real world relating to several research problems in the various fields of applied statistics. However, still, there are an equally a large number of situations, which have not yet been mathematically or statistically modeled, because of the situations may be complex or models formed are mathematically or statistically intractable.

1.2 CLASSIFICATION OF GROWTH MODELS

The various growth models can be broadly classified into two categories namely (i) Discrete Growth Models and (ii) Continuous Growth Models.

Suppose y_n be the value of y after n intervals of time passed. For discrete growth model, one may write a difference equation as

$$y_{n+1} = f\left(y_n, y_{n-1}, ..., t\right)$$

Mathematically, solving a difference equation means finding an explicit expression for y_n in terms of n and initial values such as y_0. A difference growth model connecting y_{n+1} and y_n and no other y values is called first-order difference growth model. If the growth model also involved y_{n-1} or y_{n+2} then it will be called Second order difference growth model. In other words, the order is difference between the highest and lowest subscripts appearing in the model. For instance, the first order linear difference growth model may be written as $y_{n+1} = \alpha y_n + \beta;\ \alpha, \beta$ are constants. A simple example of a nonlinear difference growth model may be written as

$$y_{n+1} = \alpha y_n \left(1 - y_n\right)$$

where α is constant.

Linear Difference Growth models involving more than one variable can be expressed using vectors and matrices.

In the case of continuous growth models, the techniques of differential calculus become very relevant. If f (t) is a function of time then the instantaneous rate of change of 'f' with respect to t is given by $\frac{d\ f(t)}{dt}$. When more than one variable changes at the same time, one may use partial derivatives.

Generally a continuous growth model is suitable when the time interval is small. The main advantage of continuous growth model is that a differential equation may be easier to manipulate and solve than a difference growth model. Some of the simple growth models are given by;

(a) **Discrete Growth Models**

(i) **Arithmetic Growth Model :** $y_{n+1} = y_n + c$

The arithmetic growth model, also known as linear growth, is a simple mathematical representation used to describe a situation where a quantity or value increases by a fixed amount in each time period. It is a linear relationship where the change is constant over time. This is in contrast to exponential growth, where the quantity increases by a fixed percentage in each time period.

(ii) **Geometric Growth Model :** $y_{n+1} = \alpha\ y_n$

The geometric growth model, also known as exponential growth, is a mathematical model that describes a situation in which a quantity or value increases by a fixed percentage or rate in each time period. It is a nonlinear relationship where the growth is proportional to the current value of the quantity. This is in contrast to arithmetic growth, where the quantity increases by a fixed amount in each time period.

(iii) **Linear First Order Growth Model :** $y_{n+1} = \alpha\ y_n + b$

The linear first-order growth model is a mathematical representation used to describe a situation where a quantity or value increases or decreases linearly over time with a rate of change that depends on a constant proportionality factor. Unlike the arithmetic growth model, which has a constant fixed increment, the linear first-order growth model involves a rate of change that is proportional to the current value of the quantity.

(b) **Continuous Growth Models**

(i**) Linear Growth Model:** $y = \alpha + \beta t$

The linear growth model, also known as linear growth or arithmetic growth, is a simple mathematical representation used to describe a situation where a quantity or value increases or decreases by a fixed amount in each time period. In this model, the change in the value of the quantity is constant over time, resulting in a straight-line relationship when graphed.

The linear growth model is a straightforward way to represent and understand linear trends in various fields, such as economics, finance, and physics. It provides a clear and intuitive way to make predictions about how a quantity will change over time when the rate of change is constant.

(ii) **Power Functional Growth Model :** $y = at^b$

The power functional growth model, also known as a power-law model or exponential growth model, describes a situation where a quantity or value increases or decreases exponentially with time. In this model, the rate of change of the quantity is proportional to its current value, resulting in exponential growth or decay.

(iii) **Polynomial Function Growth Model:** $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + ...$

A polynomial growth model describes a situation where a quantity or value increases or decreases according to a polynomial function of time. Polynomial functions are mathematical expressions involving variables raised to various powers, and they can represent a wide range of growth patterns, including linear, quadratic, cubic, and higher-degree growth or decay.

(iv) **Exponential Growth Model**: $y = \alpha e^{\beta t} \quad or \quad y = \alpha \beta^t$

The exponential growth model is a mathematical model that describes the process of growth in which a quantity increases rapidly over time. In this model, the rate of growth is directly proportional to the current size of the quantity. Exponential growth is characterized by a constant relative growth rate, which means that the quantity increases by a fixed percentage over a fixed time period.

The exponential growth model is often used to describe phenomena such as population growth, the spread of diseases, compound interest in finance, and the growth of microorganisms under ideal conditions. It's important to note that exponential growth is an idealized model and may not always accurately represent real-world situations. In reality, factors such as resource limitations, competition, and external constraints can limit exponential growth and lead to more complex growth patterns.

CHAPTER - 2 MATHEMATICAL ASPECTS OF GROWTH MODELS

2.1 INTRODUCTION

2.1.1 SIMPLE SITUATIONS REQUIRING MATHEMATICAL MODELLING

1. **Mathematical modeling** is a powerful tool for understanding and solving various real-world problems. Here are some simple situations that require mathematical modeling:
2. **Population Growth:** Modeling the growth of a population over time using equations such as the logistic growth model or exponential growth model.
3. **Compound Interest:** Calculating the future value of an investment based on interest rates, compounding frequency, and the initial principal amount.
4. **Projectile Motion:** Modeling the trajectory of an object thrown or launched into the air, considering factors like initial velocity, angle of projection, and gravity.
5. **Traffic Flow:** Analyzing traffic patterns on roads or highways to optimize traffic signals, lane configurations, and reduce congestion.
6. **Epidemiology:** Modeling the spread of diseases within a population using epidemiological models like the SIR (Susceptible-Infectious-Recovered) model.
7. Chemical Reactions: Predicting the outcome of chemical reactions by using chemical kinetics and reaction rate equations.
8. Finance: Developing models for stock price predictions, portfolio optimization, or option pricing using mathematical finance models like the Black-Scholes model.
9. Environmental Pollution: Simulating the dispersion of pollutants in air or water, taking into account factors like diffusion, wind speed, and chemical reactions.

10. Weather Forecasting: Building mathematical models to predict weather patterns and climate changes using data on temperature, pressure, humidity, and more.
11. Supply Chain Management: Optimizing supply chain logistics by modeling factors like demand, inventory levels, production capacity, and transportation costs.
12. Electrical Circuits: Analyzing and designing electrical circuits using Kirchhoff's laws and other mathematical tools.
13. Structural Engineering: Determining the stress, strain, and stability of structures under various loads, such as bridges, buildings, and dams.
14. Fluid Dynamics: Modeling the flow of fluids (e.g., air, water) in pipes, channels, or around objects to solve problems related to aerodynamics, hydrodynamics, or hydraulic systems.
15. **Optimization Problems**: Finding the best solution to a problem with constraints, such as linear programming for resource allocation or the traveling salesman problem for route optimization.
16. **Game Theory:** Analyzing strategic interactions and decision-making in situations involving multiple players, such as in economics, politics, or competition.
17. **Robotics and Motion Planning:** Developing algorithms for robotic movements and path planning, considering obstacles and constraints.
18. **Machine Learning and Data Analysis:** Using mathematical models for data analysis, classification, regression, and prediction in various domains.
19. **Geographic Information Systems (GIS):** Modeling geographical data for purposes like urban planning, resource management, and environmental analysis.
20. **Energy Consumption:** Modeling energy consumption patterns in buildings or industrial processes to optimize energy efficiency.

21. **Criminal Justice:** Modeling crime patterns and predicting crime hotspots to aid law enforcement and policy decisions.

These are just a few examples, and mathematical modeling can be applied to a wide range of fields to gain insights, make predictions, and solve practical problems. It involves using mathematics to represent and understand real-world phenomena, often requiring the development of equations, simulations, or computational models.

2.1.2 CLASSIFICATION OF MATHEMATICAL MODELS

Mathematical models can be classified into various categories based on their characteristics and applications. Here are some common classifications of mathematical models:

Deterministic vs. Stochastic Models:

Deterministic Models: These models are based on precise, predictable relationships and do not consider random factors. They provide a single, fixed solution for a given set of input parameters.

Stochastic Models: Stochastic models incorporate randomness and uncertainty into the modeling process. They often involve probability distributions and provide probabilistic
outcomes.

Continuous vs. Discrete Models:

Continuous Models: These models describe systems that change continuously over time or space, often using differential equations or functions.

Discrete Models: Discrete models represent systems with distinct, separate states or time steps, commonly used in fields like computer science and discrete mathematics.

Static vs. Dynamic Models:

Static Models: Static models represent systems at a single point in time or without considering changes over time. They are used for situations where time dynamics are not significant.

Dynamic Models: Dynamic models capture how systems change and evolve over time, making them suitable for modeling dynamic processes and systems.

Linear vs. Nonlinear Models:

Linear Models: Linear models assume that the relationships between variables are linear, adhering to the principles of superposition and proportionality.

Nonlinear Models: Nonlinear models allow for more complex relationships between variables, where changes in one variable may not produce proportional changes in another.

Descriptive vs. Predictive Models:

Descriptive Models: Descriptive models aim to understand and explain observed phenomena. They do not necessarily make predictions but provide insights into the underlying processes.

Predictive Models: Predictive models use historical data to forecast future events or outcomes. Machine learning algorithms, for example, are often used for predictive modeling.

Physical Models vs. Empirical Models:

Physical Models: Physical models are based on fundamental principles and laws of nature, such as Newton's laws of motion or the laws of thermodynamics.

Empirical Models: Empirical models are developed based on observed data and statistical relationships, often without a direct connection to underlying physical principles.

Deterministic Chaos Models:

Deterministic Chaos Models: These models exhibit chaotic behavior even though they are entirely deterministic. They are often used to describe systems that are highly sensitive to

initial conditions, such as the Lorenz system.

Agent-Based Models: These models simulate the interactions and behaviors of individual agents or entities within a system, often used in social sciences, economics, and ecology.

Spatial Models: These models consider the spatial distribution of variables and are used in fields like geography, urban planning, and ecology to study spatial relationships and patterns.

Optimization Models: These models aim to find the best solution among a set of feasible alternatives, often involving linear programming, integer programming, or nonlinear optimization.

Network Models: Network models describe relationships and interactions between interconnected entities, such as in transportation networks, social networks, or communication networks.

Hybrid Models: These models combine elements of multiple types of models to address complex, multidimensional problems. For example, a model might integrate deterministic and stochastic components.

The choice of model type depends on the specific problem, the available data, and the level of detail and accuracy required for the analysis or simulation. Researchers and practitioners select the appropriate model category based on the characteristics of the system being studied or modeled.

2.1.3 SOME CHARACTERISTICS OF MATHEMATICAL MODELS

Mathematical models are abstract representations of real-world phenomena, systems, or processes using mathematical language and principles. These models are used in various fields, including science, engineering, economics, and social sciences, to understand, predict, and analyze complex systems. Here are some key characteristics of mathematical models:

Abstraction: Mathematical models simplify real-world complexities by focusing on essential elements while disregarding less important details. This abstraction is necessary to make complex systems manageable and analyzable.

Mathematical Formulation: Models are expressed using mathematical equations, functions, and symbols to represent relationships and interactions within the system being studied. These equations can be differential equations, algebraic equations, or other mathematical expressions.

Assumptions: Models are based on assumptions about the behavior and characteristics of the

system. Assumptions may include simplifications, idealizations, or constraints that make the model tractable. Understanding the assumptions is crucial for interpreting model results.

Variables: Models include variables that represent the quantities or parameters of interest. These variables can be dependent (affected by other variables) or independent (driving the system's behavior).

Parameters: Models often include parameters, which are constants that characterize the system. Parameters can be estimated from data or based on expert knowledge.

Time Dependency: Many models are time-dependent, meaning they describe how the system changes over time. Time can be discrete (e.g., in discrete-event simulation) or continuous (e.g., in differential equations).

Predictive Capability: Mathematical models are used to make predictions about the behavior of the system under different conditions or scenarios. They can answer "what if" questions and help plan for various outcomes.

Validation and Verification: Models should be validated and verified to ensure that they accurately represent the real-world system. Validation involves comparing model predictions to observed data, while verification ensures the model's mathematical correctness.

Generalizability: A good mathematical model can be applied to a range of similar systems or situations. It should possess a degree of generality and be adaptable to different contexts.

Sensitivity Analysis: Sensitivity analysis assesses how changes in model inputs or parameters affect model outcomes. This analysis helps identify critical factors and uncertainties in the model.

Simulation: Some mathematical models are implemented as simulations, allowing researchers to study the behavior of the system through computational experiments. Simulations can provide insights into complex dynamic systems.

Interpretability: Models should be interpretable, meaning that their mathematical structure and parameters have meaningful connections to the real-world system. Interpretability aids in understanding the implications of model results.

Limitations: Models have limitations and may not fully capture all aspects of the real-world system. Recognizing and acknowledging these limitations is essential for responsible use of models.

Trade-offs: Models often involve trade-offs between complexity and simplicity. More

complex models may provide better representation but may be computationally intensive, while simpler models may lack accuracy.

Ethical Considerations: Mathematical models should be developed and used ethically, particularly when they have implications for human welfare, policy, or decision-making.

Iterative Process: Developing and refining mathematical models is often an iterative process. Models are continually improved as new data becomes available and as our understanding of the system deepens.

Communication: Models are a means of communicating insights and predictions to a broader audience, including policymakers, stakeholders, and the general public.

Mathematical models are valuable tools for understanding and solving complex problems, but their effectiveness depends on the quality of their formulation, assumptions, and validation processes. Properly designed and used, mathematical models can provide valuable insights and support informed decision-making.

1. ***Estimation of Parameters*****:** Every model contains some parameters and these have to be estimated. The models must itself suggest, experiments or observations and the method of calculation of these parameters. Without this explicit specification, the model is incomplete.
2. ***Modelling: Mathematics + Discipline*****:** For making a mathematical models of a situation, one must know both mathematics and the disciple in which the situation arises. Efforts to make a mathematical model without deeply understanding the discipline concerned may lead to infrutous models. Discipline insight must both precede and flow mathematical modelling.
3. ***Mathematical modelling and mathematical techniques*****:** Emphasis in applied mathematics has very often been on mathematical techniques, but the heart of applied mathematics is mathematical modelling.
4. ***Criteria for successful models:*** These include good agreement between predictions and observations, of drawing further valid conclusions, simplicity of the model and its precision.
5. ***Constraints of additively and normality:*** Models which are linear, additive and in which the probability distribution follows the normal law one relatively simpler, but relatively more realistic models have to be free from these constraints.
6. ***Validation by independent data*****:** Sometimes parameters are estimated with the help of same data and the same data are used to validate the model. This is illegitimate. Independent data should be used to validate the model.
7. ***Modelling in terms of modules****:* One may think of models for small modules and by combining them in different ways, one may get models for a large number of systems.
8. ***Complexity of models****:* This can be increased by subdividing variables, by taking more variables and by considering more details. Increase of complexity need not always lead to increase of insight as after a stage, diminishing returns begin to set in. the one of mathematical modelling consists in stopping before this stage.

9. ***Robustness of models:*** Mathematical model is said to be robust if small changes in the parameters lead to small changes in the behaviour of the model. The decision is made by using sensitivity analysis for the models.

10. ***Relative precision of models:*** Different models differ in their precision and their agreement with observations.

11. ***Hierarchy of models:*** Mathematical modelling is not a one shot affair. Models are constantly improved to make them more realistic. Thus for every situation, we get a hierarchy of models, each more realistic than the preceding and each likely to be followed by a better one.

2.1.4 MATHEMATICAL MODELLING THROUGH GEOMETRY

Mathematical modeling through geometry involves using geometric principles, shapes, and transformations to describe and analyze real-world phenomena or solve practical problems. Geometry provides a powerful framework for modeling a wide range of situations, from simple spatial relationships to complex structures. Here are some examples of mathematical modeling through geometry:

Architectural Design: Architects use geometric modeling to design buildings, bridges, and other structures. They use concepts like symmetry, proportion, and geometric transformations to create aesthetically pleasing and functional designs.

City Planning: Urban planners use geometric models to lay out city streets, parks, and infrastructure. Concepts such as grid patterns, road intersections, and zoning regulations are based on geometric principles.

Robotics and Motion Planning: In robotics, geometric modeling is used to plan the movement and navigation of robots. Concepts like kinematics, inverse kinematics, and collision detection rely heavily on geometric computations.

Computer Graphics: Computer graphics and animation rely on geometric modeling to create and manipulate 2D and 3D objects. Geometric transformations, such as translation, rotation, and scaling, are fundamental operations in computer graphics.

Molecular Modeling: In chemistry and biology, geometric modeling is used to study the three-dimensional structures of molecules, proteins, and DNA. This helps in understanding chemical reactions and biological processes.

Cartography and Geographic Information Systems (GIS): Cartographers use geometric modeling to create maps that accurately represent geographical features and relationships. GIS technology relies on geometric data to analyze and visualize spatial information.

Mechanical Engineering: Engineers use geometric modeling to design and analyze mechanical parts and systems. Concepts like solid modeling, finite element analysis, and tolerance analysis are essential for product design and manufacturing.

Art and Sculpture: Artists often use geometric principles to create visually appealing and balanced compositions. Geometric shapes, such as circles, triangles, and rectangles, are used as the basis for many artworks.

Crystallography: Scientists use geometric models to understand the atomic arrangement in crystals. X-ray diffraction and geometric principles are employed to determine crystal structures.

Optics and Lens Design: Optical engineers use geometric modeling to design lenses and optical systems. Concepts like ray tracing and Snell's law are used to predict how light behaves in various optical setups.

3D Printing and Additive Manufacturing: Geometric modeling is essential in 3D printing and additive manufacturing processes, where digital models are converted into physical objects layer by layer.

Pattern Recognition: Geometric modeling is used in pattern recognition tasks, such as identifying shapes, objects, or patterns in images and data.

Astronomy and Celestial Mechanics: Astronomers use geometry to model the positions and movements of celestial bodies. Concepts like ellipses, orbits, and angular measurements are employed.

Navigation and GPS: Geometric principles are used in global positioning systems (GPS) to calculate the position of a receiver based on signals from multiple satellites.

Environmental Modeling: Geometric modeling is used to analyze and visualize environmental data, such as terrain elevation, land use, and the flow of water in watersheds.

In these and many other fields, geometric modeling serves as a valuable tool for representing, analyzing, and solving real-world problems. It helps researchers and professionals gain insights into the spatial and structural aspects of systems and phenomena.

2.1.5 MATHEMATICAL MODELLING THROUGH ALGEBRA

Mathematical modeling through algebra involves using algebraic equations and expressions to describe and analyze real-world situations, relationships, and problems. Algebra provides a versatile and powerful tool for modeling a wide range of phenomena. Here are some examples of mathematical modeling through algebra:

Population Growth: The logistic growth model and exponential growth model are classic algebraic models used to describe how populations of organisms grow over time. These models incorporate variables such as population size, growth rate, and carrying capacity.

Finance and Investment: Algebraic models are commonly used in finance to model various aspects of investments. For example, the compound interest formula is an algebraic model used to calculate the future value of an investment based on the principal amount, interest rate, and time.

Physics and Engineering: Algebraic equations describe the relationships between physical quantities in various engineering and physics applications. For example, $F = ma$ (Newton's second law) relates force (F), mass (m), and acceleration (a) through an algebraic equation.

Chemical Reactions: Chemical reactions are often represented using algebraic equations, such as stoichiometric equations, which show the proportions of reactants and products in a chemical reaction.

Economics: Algebraic models are used in economics to represent supply and demand relationships, production functions, and cost functions. Linear and nonlinear equations are common tools for modeling economic systems.

Electrical Circuits: Algebraic equations are used to analyze and design electrical circuits, such as Ohm's law ($V = IR$) for relating voltage (V), current (I), and resistance (R) in a circuit.

Statistics and Data Analysis: Algebraic models are employed in statistical analysis and regression analysis to describe relationships between variables and make predictions based on data.

Optimization Problems: Linear programming and quadratic programming involve algebraic modeling to optimize objective functions subject to constraints. These models are used in operations research and resource allocation.

Actuarial Science: In the insurance industry, algebraic models are used to calculate premiums, reserves, and probabilities of events, such as the probability of an insurance claim occurring.

Game Theory: Algebraic models are used to represent and analyze strategic interactions among players in various games. Payoff matrices and Nash equilibrium equations are examples of algebraic tools used in game theory.

Environmental Modeling: Algebraic models are employed to describe and predict environmental phenomena, such as pollution dispersion, species interactions, and ecosystem dynamics.

Epidemiology: Algebraic models, such as the basic reproduction number (R_0) in epidemiology, are used to understand the spread of diseases within populations.

Cryptocurrency and Blockchain: Algebraic models are used in cryptography and blockchain technology to describe algorithms and security protocols.

Network Analysis: Algebraic models represent relationships and connections in network analysis, including social networks, communication networks, and transportation networks.

Control Systems: Algebraic equations are used in control systems engineering to model and analyze the behavior of dynamic systems, such as those in robotics and industrial automation.

Algebraic models are valuable because they allow for the representation of relationships and behaviors of systems using equations that can be manipulated and solved analytically or computationally. They are essential tools in science, engineering, economics, and many other fields for understanding, predicting, and optimizing real-world phenomena.

This model was used about two thousand years ago. *A* and *B* are two points on the surface of the Earth with the same longitude and *d* miles apart. When the sun is vertically above *A* (i.e. it is a direction *OA*, where *O* is the centre of the earth) the Sun's rays make an angle of θ^0 with the vertical at *B* (i.e. with the line *OB*). If *a* miles is the radius of the earth,

$$\frac{d}{2\pi a} = \frac{\theta}{360} \qquad \text{(or)}$$

$$a = \frac{360}{2\pi\theta} \qquad (2.1.1)$$

2.1.6 MATHEMATICAL MODELLING THROUGH TRIGONOMETRY

Mathematical modeling through trigonometry involves using trigonometric functions and relationships to describe and analyze real-world phenomena and problems that involve angles, periodic behavior, and wave-like patterns. Trigonometry plays a crucial role in various fields where angular and cyclical aspects are significant. Here are some examples of mathematical modeling through trigonometry:

Wave Phenomena: Trigonometric functions like sine and cosine are used to model wave behavior in physics, such as the oscillation of a pendulum, the vibrations of a guitar string, or electromagnetic waves.

Mechanical Vibrations: Trigonometric functions are used to model the vibrations and oscillations of mechanical systems, including springs, pendulums, and mass-spring-damper systems.

Electrical Circuits: In alternating current (AC) circuits, trigonometric functions describe voltage and current waveforms. Phasor analysis uses trigonometry to represent complex AC quantities.

Sound Waves: Trigonometry is used to model sound waves, including their frequency, amplitude, and the Doppler effect, which describes the change in frequency of a sound source in motion.

Navigation and GPS: Trigonometric functions like the law of sines and law of cosines are used in navigation to calculate distances, angles, and positions on the Earth's surface.

Astronomy: Trigonometry is essential in astronomy for measuring distances between celestial objects, calculating their positions, and understanding the geometry of the night sky.

Mechanical Engineering: Trigonometry is used to analyze and design mechanical systems involving angles and rotational motion, such as gears, linkages, and cam mechanisms.

Optics and Light: Trigonometry is employed in optics to describe the behavior of light waves, including reflection, refraction, and the geometry of lenses and mirrors.

Periodic Phenomena: Trigonometric functions model periodic behavior in various fields, including the motion of planets, tides, and cyclical economic trends.

Seismic Waves: Trigonometry is used to analyze seismic waves generated by earthquakes, helping to determine their direction, magnitude, and epicenter.

Music and Sound Engineering: Trigonometry is used to analyze and synthesize musical tones, including frequency, amplitude, and harmonics.

Mechanical Waves: Trigonometric functions describe the propagation of mechanical waves, such as water waves, sound waves, and seismic waves.

Robotics and Kinematics: Trigonometry is used to model the movement and positions of robotic arms and other mechanical systems with rotating joints.

Geometry and Geodesy: Trigonometry plays a fundamental role in geometry, including calculating the measurements of angles, distances, and areas. It is also used in geodesy to measure the shape and size of the Earth.

Artificial Intelligence and Signal Processing: Trigonometric functions are used in Fourier analysis to analyze and process signals and data, including image and speech recognition.

Control Systems: Trigonometry is applied to model and analyze the behavior of dynamic systems in control engineering, such as systems with oscillatory responses.

Trigonometry provides a powerful set of tools for modeling and understanding cyclical and angular phenomena in various fields of science, engineering, and mathematics. It helps researchers and professionals describe and analyze complex systems with periodic or wave-like behavior.

From two points *A*, *B* the surface of the earth will be same longitude, one in the Northern hemisphere and the other in the Southern hemisphere, measure angles θ_1, θ_2 between verticals at *A* and *B* directions of the centre of the moon.

If *d* is the distance of the centre of the moon's disc from the centre of Earth,

$$\frac{d}{\sin\theta_1} = \frac{a}{\sin(\theta_1 - \varphi_1)}\ , \qquad \frac{d}{\sin\theta_2} = \frac{a}{\sin(\theta_2 - \varphi_2)} \tag{2.1.2}$$

Also

$$\varphi_1 + \varphi_2 = \alpha = \phi_1 + \phi_2 \tag{2.1.3}$$

where ϕ_1 is the northern latitude of *A* and ϕ_2 is the southern latitude of *B.*

Since ϕ_1, ϕ_2 are know, $\varphi_1 + \varphi_2$ is known. Eliminating φ_1, φ_2 from (2.`1.2) and (2.1.3), we get *d* in terms of a, ϕ_1, ϕ_2 which are all known.

2.1.7 MATHEMATICAL MODELLING THROUGH CALCULUS

Mathematical modeling through calculus involves using the principles of calculus, including differentiation and integration, to describe and analyze real-world phenomena, systems, and problems. Calculus provides powerful tools for understanding how quantities change with respect to one another, making it a valuable tool in many scientific and engineering applications. Here are some examples of mathematical modeling through calculus:

Motion and Dynamics: Calculus is used to model the motion of objects by describing their position, velocity, and acceleration as functions of time. This is essential in physics and engineering for understanding how objects move under various forces.

Rate of Change: Calculus is employed to model and analyze the rate at which quantities change. This is valuable in fields like economics to study production and consumption rates, or in biology to understand population growth rates.

Optimization Problems: Calculus is used to solve optimization problems where the goal is to find the maximum or minimum of a certain function, subject to constraints. Examples include finding the optimal design of structures or maximizing profits in business.

Population Dynamics: Calculus models population growth and decay, considering birth rates, death rates, immigration, and emigration. These models are used in demography, ecology, and epidemiology.

Chemical Kinetics: Calculus is used to describe the rates of chemical reactions, helping chemists understand how reactants transform into products over time.

Thermodynamics: Calculus is used to describe the behavior of thermodynamic systems, including processes such as heating, cooling, and phase transitions.

Economics: Calculus is used to model economic relationships, such as supply and demand curves, cost functions, and utility functions. It helps economists analyze market behavior and policy impacts.

Electricity and Magnetism: Calculus is essential in the study of electromagnetic fields and circuits, where it is used to describe electric and magnetic forces, voltage, current, and electromagnetic waves.

Fluid Mechanics: Calculus is employed to analyze the flow of fluids (liquids and gases) in pipes, channels, and around objects. It helps in understanding fluid dynamics and pressure distribution.

Probability and Statistics: Calculus is used in probability theory and statistics to calculate probabilities, find expected values, and derive statistical distributions.

Control Theory: Calculus plays a key role in control systems engineering, where it is used to model and analyze the behavior of dynamic systems and design controllers to achieve desired performance.

Environmental Modeling: Calculus is used to model environmental processes such as pollutant dispersion, groundwater flow, and ecosystem dynamics.

Finance and Portfolio Optimization: Calculus is employed in financial modeling to determine optimal investment strategies, risk assessment, and option pricing.

Machine Learning and Data Analysis: Calculus is foundational for many machine learning algorithms, including gradient descent used in training neural networks.

Medical Sciences: Calculus is used in medical imaging, pharmacokinetics (drug absorption and distribution), and modeling disease spread in epidemiology.

Astrophysics and Cosmology: Calculus helps model celestial phenomena, such as the motion of planets, the expansion of the universe, and the behavior of black holes.

Geophysics: Calculus is used to model seismic waves, study the Earth's interior, and analyze geological processes.

In each of these applications, calculus is a critical tool for developing mathematical models that describe the behavior of systems, analyze data, and make predictions. It provides the mathematical framework for understanding how variables change continuously and helps researchers and professionals solve complex problems in various fields.

Consider the problem of refraction of light from a point *A* in vacuum to a point *B* in a medium of refractive index μ. If light travels with velocity *V* in vacuum it travels with velocity *V/μ* in the second medium so that the time *T* of travel is given by

$$T=\sqrt{\frac{a^2+x^2}{V}}+\mu\sqrt{\frac{b^2+(c-x^2)}{V}} \tag{2.1.4}$$

So that

$$V\frac{dT}{dx}=\frac{x}{\sqrt{a^2+x^2}}-\mu\frac{c-x}{\sqrt{b^2+(c-x^2)}} \tag{2.1.5}$$

$$V\frac{d^2T}{dx^2}=\frac{a^2}{(a^2+x^2)^{3/2}}+\frac{\mu b^2}{(b^2+(c-x)^2)^{3/2}} \tag{2.1.6}$$

Thus *T* is minimum when

$$\frac{x}{\sqrt{a^2+x^2}}=\mu\frac{c-x}{\sqrt{b^2+(c-x^2)}} \quad or \quad \sin\alpha=\mu\sin\beta \tag{2.1.7}$$

2.1.8 LIMITATIONS OF MATHEMATICAL MODELLING

Mathematical modeling is a valuable tool for understanding and solving real-world problems, but it also has limitations. Here are some of the key limitations of mathematical modeling:

Simplification of Reality: Mathematical models are abstractions of real-world systems, and they often simplify complex phenomena to make them tractable. These simplifications may neglect important factors or interactions, leading to inaccuracies in predictions.

Assumptions: Models rely on assumptions about the behavior of the system being studied. If these assumptions are incorrect or oversimplified, the model's predictions may not be valid.

Data Availability and Quality: Models require data for parameter estimation and validation. If data is limited, incomplete, or inaccurate, it can lead to unreliable model outcomes.

Parameter Estimation: Many models involve parameters that need to be estimated from data. The accuracy of these parameter estimates can affect the model's performance, and inaccuracies can lead to incorrect results.

Complexity: Complex systems may require highly intricate and computationally intensive models, making them difficult to solve or simulate. This complexity can limit the practicality of modeling.

Nonlinearity: Real-world systems often exhibit nonlinear behavior, which can be challenging to model accurately. Linear approximations may not capture important dynamics.

Uncertainty: Models typically cannot account for all sources of uncertainty. Stochastic models attempt to address some of this uncertainty but may not fully capture the probabilistic nature of some phenomena.

Validation and Verification: It can be challenging to validate and verify models against real-world data, especially for systems with limited observational data or for predicting future events.

Human Factors: Models may not account for human behavior or decision-making accurately, especially in social sciences and economics, where human actions can be unpredictable.

Ethical Considerations: Models can be used for purposes that raise ethical concerns, such as profiling individuals or making decisions that impact people's lives.

Inherent Error: Models are approximations of reality and may contain inherent errors or inaccuracies due to simplifications and limitations in mathematical techniques.

Computational Resources: Some models, especially those involving high-dimensional systems or complex simulations, may require significant computational resources and time, limiting their practicality.

Domain Expertise: Developing and interpreting mathematical models often requires expertise in both mathematics and the specific domain being modeled. Misunderstandings or oversights can lead to incorrect conclusions.

Model Overfitting: In machine learning and statistics, overfitting can occur when a model is too complex and fits the training data too closely, leading to poor generalization to new data.

External Factors: Models may not account for external factors that can influence the system being studied, such as economic, political, or environmental changes.

Despite these limitations, mathematical modeling remains a valuable tool for gaining insights, making predictions, and solving problems in various fields. Recognizing these limitations and conducting sensitivity analyses can help improve the accuracy and reliability of mathematical models. It's also important to use modeling as a complement to empirical data and expert judgment, rather than as a sole source of decision-making.

2.7 MATHEMATICAL MODELLING SKILLS

Developing mathematical modeling skills is essential for effectively using mathematics to represent and analyze real-world phenomena. Whether you're a student, researcher, or professional in a field that involves mathematical modeling, here are some key skills and steps to develop and improve your modeling abilities:

Mathematical Knowledge: Build a strong foundation in mathematics, including calculus, linear algebra, differential equations, and probability/statistics. These are fundamental tools for creating and solving mathematical models.

Domain Knowledge: Understand the specific field or problem you're modeling. Domain knowledge is crucial for making informed assumptions, selecting relevant variables, and interpreting model results accurately.

Problem Formulation: Clearly define the problem you want to address with a mathematical model. Identify the key variables, parameters, and relationships that play a role in the system you're modeling.

Assumptions: Be explicit about the assumptions you're making when creating a model. Assumptions help simplify complex systems but should be realistic and justifiable.

Model Selection: Choose an appropriate modeling approach based on the nature of the problem. Decide whether a deterministic, stochastic, discrete, or continuous model is more suitable.

Equations and Formulation: Develop mathematical equations that describe how the variables in your model interact over time. These equations are often in the form of differential equations, difference equations, or algebraic equations.

Parameter Estimation: If your model involves parameters (constants), estimate or calibrate them using available data. Parameter estimation methods may include least squares fitting or maximum likelihood estimation.

Numerical Methods: Learn and apply numerical methods to solve mathematical models when analytical solutions are not possible. This includes techniques like finite difference methods, finite element methods, and numerical integration.

Simulation: Implement simulations of your model using software tools like MATLAB, Python, or specialized modeling software. Simulations allow you to explore how the system behaves under different conditions.

Validation and Verification: Validate your model by comparing its predictions with observed data. Ensure that the model's behavior aligns with real-world outcomes and that it is free from errors (verification).

Sensitivity Analysis: Conduct sensitivity analyses to assess how changes in model parameters or initial conditions affect the model's predictions. Identify which factors have the most significant impact on the results.

Model Complexity: Choose an appropriate level of model complexity. Avoid overcomplicating models if simpler ones can adequately represent the system.

Visualization: Use graphs, plots, and visualization techniques to help understand and communicate the results of your model. Visualization can make complex concepts more accessible.

Interpretation: Interpret the results of your model in the context of the problem. Explain the practical implications of your findings and whether they align with expectations.

Communication Skills: Develop effective communication skills to convey your modeling results to others, whether through written reports, presentations, or discussions with colleagues and stakeholders.

Continuous Learning: Stay up-to-date with advances in modeling techniques and software tools. Attend workshops, conferences, and online courses to expand your modeling knowledge.

Collaboration: Collaborate with experts in related fields who can provide valuable insights, data, or alternative perspectives on your modeling projects.

Peer Review: Seek feedback and peer review from colleagues or mentors to improve the quality and rigor of your modeling work.

Ethical Considerations: Consider the ethical implications of your modeling work, especially when it has real-world consequences. Ensure that your modeling is conducted with integrity and transparency.

Persistence and Patience: Mathematical modeling can be challenging, and results may not always align with expectations. Maintain persistence and patience in refining and improving your models.

Remember that mathematical modeling is both an art and a science. It requires creativity, critical thinking, and problem-solving skills to build and refine models that accurately represent complex real-world systems. Practice and experience are key to becoming proficient in mathematical modeling.

2.7.2 SETTING UP MODELS

The purpose of modelling is to solve real practical problems, and this means that we must understand from the outset exactly what problem we are trying to solve. Successful mathematical modelling depends on getting things right from the start, and, as in most other scientific endeavours, we are more likely to succeed if we adopt a methodical approach. In most cases it is found useful to complete the following steps.

1. Clarify the problem.
2. List the factors.
3. List the assumptions.
4. Formulate a precise problem statement.

Clarifying the Problem

Essentially, a problem will be proposed, often by a non-mathematician, or indeed a non-scientist and a particular answer requested from the model outcome,the answer often to interpreted in non-mathematical terms.

A specific objective can then be identified. As mathematical modelers we must resist getting stuck into the mathematical formulation before this step has been achieved. Misunderstandings will be very costly if there is a failure to carry out a preliminary investigation first and the wrong problem is modelled and 'solved'.

Establishing the true nature of the problem involves asking a number of questions. No matter how simple the situation appears, it will be worthwhile to ask questions for the problem provider to establish what is needed from the subsequent model:

- Who is intending to use the model-how much sophistication and complexity are needed?
- Is there an underlying physical/scientific behavior to be taken into account?
- Is there data available or does it have to be collected or looked up?
- What underlying assumptions and simplifications about the problem can reasonably be made?
- When is the model 'solution' is required - what time limits have been set?
- In what form is a solution wanted-written report, short oral presentation?
- Are there conflicting outcomes at stake, perhaps concerning the cost of implementing two different courses of action.

This last remark is quite important, as the essence of mathematical modelling in the real world is often help answer conflicts between opposite viewpoints. For example, in the Post Office, when a large queue of customers forms and the cry goes up 'why don't they open more service counters?' , there is a conflict between the Post Office manager, who can't afford to employ an abundance of staff due to wage budget limits, and the customer, who reasonably wants a fast service and will go elsewhere if patience is tried too far.

As mathematical modellers we must be clear what we can achieve before we launch into the above questions. What do we need to establish, before starting on a model? We must appreciate both the scope and the limitations of modelling and that we may have some powerful computing tools at our disposal with which a mathematical solution can be found. What must be established in discussion with the problem provider is exactly what we can deliver. This will involve pinning down specific objectives and removing vague or irrelevant features from the context, at least until a first model has been completed and explained to the provider.

Our wider objectives are to produce a model, probably using algebra and other mathematical tools, which is sufficiently general that if can be reused for other similar situations. Do not forget, however, that problem providers are usually less interested in general mathematical models than in particular answers to their problem, and the capability of the model in dealing with questions of the 'what if' type.

Listing the Factors

Every problem involves a number of different 'factors' which may have a bearing on the solution. At the early stage of model-building we need a list of these factors. In mathematical modelling we tend to concentrate mainly on quantifiable factors, i.e. those which can be given numerical values (in terms of suitable units). Quantifiable factors can normally be classified as variables, parameters or constants, and each of these can be continuous, discrete or random. Brief explanations of these categories are as follows:

Continuous**:** takes all real values over an interval, e.g. time velocity, length, cost, area, etc.

Discrete: takes on certain isolated values only. Very often these will be whole numbers, e.g. the number of people, tickets, matches played etc., in which case. There are no units of measurement.

Random**:** unpredictable in advance, but governed by some underlying statistical model. For example, buses timed to arrive theoretically every 5 min but with actual random inter arrival times with a mean of 5 min. The model can either be based on data or assumed to be a particular theoretical form.

***Constants*:** Quantities whose values we can't change. These can be mathematical constants such, as ∏, or physical constants, such as acceleration due to gravity or the speed of light.

***Parameters*:** Quantities which are constant for a particular application of a model, but can have different values for another application of the same model. For example, fixed costs in a simple business model, the dimension of a room, price of a ticket, density of fluid, mean inter arrival time of a bus service.

***Input variables*:** Quantities which determine subsequent evolutions within the model, such as rainfall rate into a collecting butt, the number of people attending to disco, the number of months elapsed before you sell your can etc. Note that an input variable is expected to be known, or given or assumed, or can be considered to have any arbitrary value.

***Output variables**:* quantities which are consequences of given values of input variables and parameters and cannot be given arbitrary values. These represent the outcome from a model, such as the profit made on a business deal, the level in a result of user demand and evaporation, the time taken for a certain number of people to evacuate a room.

In order to make the mathematical work easier, all the factors need to have suitable algebraic symbols assigned to them and some which represent measurement will need units.

Listing Assumptions

While the factors provide the building blocks of the model, it is the assumptions which provide the glue with which to combine the factors together into a working model. How easy or difficult the model is to use and how successful it turns out to be depend very largely on what assumption we make. The most common and important types of assumptions that have to be made are:

(a) Assumption about whether or not to include certain factor.

(b) Assumption about the relative magnitudes of the effect of various factors.

(c) Assumption about the form of relationship between factors

Generally speaking, and especially when developing a new model for the first time, we try to choose *assumptions* which keep the model as simple as possible. Assumptions of type (a) and (b) help to keep the list of factors from being longer than strictly necessary. Assumptions of type (c) could be said to represent the heart of the model.

Problem Statement

We will have already discussed our problem under the heading ‘clarifying the problem’ above. In most problems we can identify the following ingredients:

(a) Something is known or given

(b) Something is to found, estimated or decided

(c) There is some condition to be satisfied or objective to achieve.

As a result of our considerations we should now be able to crystallise our thoughts into a precise problem statement, expressed in terms of the factors that we have listed under the heading ‘listing the factors’. This statement will have the general form:

Given {inputs, parameters, constants} find {outputs}. Such that {condition is satisfied or objective achieved}

This may look like an over-simplification, but in fact the vast majority of problems can be condensed into this precise form and it helps enormously in creating a model approximate to the problem. Note that the same problem can have different problem statements, depending on what we consider to be given and what we want the model to find.

2.7.3 DEVELOPING MODELS

Verbal statements are sometimes value and there may be a selection of possible equivalent mathematical statements. For example, the verbal statement as '*x* goes up *y* goes up' can be modelled mathematically in many ways. The simplest model is obtained by assuming that *y* is directly proportional to *x*. The equivalent mathematical statement is then $y \alpha x$, or as an equation, $y = \mathrm{k}\mathrm{x}$ where *k* is the constant of proportionality. By choosing this particular mathematical statement we are making a very clear assumption, which may well be criticized. A graph of *y* against *x* showing a straight line through the origin would be the ultimate justification.

The next simplest model is the linear form $y = a\mathrm{x} + b$ in which we are saying that *y* increases by 'a' units for every unit increase in x and that y=b when x=0. This also includes the case where 'y decreases as x increases'. In that case the parameter 'a' is negative (the gradient of the straight line graph) anther simple way of modelling the statement 'y decreases as x increases' is by inverse proportion is $\mathrm{y}\alpha^{1/x}$ or $\mathrm{y} = k / x$ this means y decreases more steeply with x than is the case in the linear model. One way of testing the validity of this assumption would be to check whether 'xy' remains nearly constant. Another way is to see if the plot of l_ny aginst l_nx is straight line of -1.

More general models can be created using the form $y=kx^a$, which gives convex curves for values of the parameter a>1 and concave curve for $a<1$. In biology, the relationship between the sizes of various parts of an organism can often be represented by such non-linear models. When there are several variable and y is assumed to be proportional to each of them is $y\alpha x_1$ and $y\alpha x_2$ and $y\alpha x_3$ for example, then the combine into the single model $y = kx_1x_2x_3$.

As part of the modelling process we need to

- Represent variables by mathematical symbols
- Make assumptions about now the variables are related

- Translate the assumptions into mathematical equation or inequalities.

When choosing symbols we usually use single letters and, as often as possible, the first letter of the name of the variable, such as t for time. It is also traditional to use Greek letters such as α,β,θ and Φ for angles.

When making assumptions, we choose the simplest versions which seem likely to reflect the behavior of the real variables. If we later realize that we have made an oversimplification, then a revised model will be necessary.

When making the translation the following simple forms may be useful to remember. More comlex forms can be often be broken down into combinations of these.

Verbal statement	**mathematical equivalent**
Sum/total	+
Difference between/change in	-
Less than	<
Greater than	>
At least	≥
Not more than	≤
Ratio	/
Y proportional to x	y=kx
Y inversely proportional to x	y=k/x
Y is x% of z	$y=\left(\frac{x}{100}\right)z$
Y is x% more than Z	$y=\left(\frac{1+x}{100}\right)z$
Rate of change of y with time t	$\frac{dy}{dt}$

Note that some case is needed with the words 'difference' and 'ratio'. The difference between A and B means A-B or it may mean the size of the difference between then is A-B where A>B, but B-A when A<B. these two can be combined in the single expression |A-B|. Similarly the "ratio of A to B" could be interpreted as A/B or B/A.

2.7.4 CHECKING MODELS

Here we discuss simple checks that can be made during the process of building the model. These checks help to determine whether the model is sensible, appropriate and adequate, without being unnecessarily complicated. Then will also be more checking to do at a last stage in the modelling process, in the validating and likely revising of the model.

Consistency

We can check whether a model is consistent with the assumptions and also whether it is logically (containing no contradictions) and dimensionally consistent. Consistency with the assumption is mainly a question of how the variables have behaved when one of them is changed. Logical consistency is usually fairly easy to check. Dimensional consistency involves checking that all physical quantities can be expressed in terms of the three fundamental dimensions mass(M), length(L) and Time(T).

Behavior

We need to examine in model's predictions' (a) qualitatively and (b) quantitatively. The qualitative examination usually involves investigating how the model predicts that one variable will change as a result of changes in other variables must of this can be carried out by remembering simple facts such as 'x' increases $\frac{1}{x}$ decreases. And 'exponentials increases or decreases faster than polynomials. A fraction like $\frac{a}{b}$ increases when 'a' increases and decreases when b increases (assuming both a and b are positive). it sometimes helps to do some algebraic simplification first. For example if a,b and c are positive parameters what happens to the expression $F = \frac{ax}{bx+c}$ as x increases ? On dividing thoroughly by x we can rewrite F *as* $\left(\frac{a}{b+\frac{c}{x}} \right)$.

now x increases $\frac{c}{x}$ decreases, so the denominator decreases and F itself increases.

Other general questions to ask about a model are what happens at extreme values of the variable (is very small or very large values) are there any special values when something interesting happens e.g. a variable reaches a local maximum or minimum value or becomes zero or infinite? Is there a square root term which can become negative? Are there numerical ranges of some variables restricted in some way because of their contextual meaning? The models quantitative behavior can be investigated so some extent by rough estimates but a more careful examination requires compression with data.

2.8 DISCRETE MODELS

One of the main points of modeling is to predict the future development of a system. A model of the economy, for example, can be used to predict future trends and son provide a basic for policy decisions. Any such model relies on assuming that the rate of change of the variable x is linked to or caused by some or all of

1. Present value of x
2. Previous value of x
3. Values of other variables
4. The rate of other variables.
5. The rate of change of other variables
6. The time ‘t’

The relationship that we want to model is the one that describes how x itself varies with time ‘t’ there are two very different ways of modelling such a relationship

(a) We could think of x(t) as a continuous function of continuous time ‘t’ in this case the graph of x against ‘t’ would show some continuous curve and our modelling objective would be no write an explicit formula for x(t) in terms of ‘t’

(b) We could think about values of x only at particular pints in time, for example at intervals of one hour or once a month in this case we use a symbol such as x_n to denote the value of x after ‘n’ intervals of time have passed. (The n is referred to as a subscript). A plot of x_n against n is now a set of separated points rather than a curve and we call this a discrete model.

For discrete models the essential ingredient is an equation of the form.

Next value= function of {present value and previous values and possibly time}

Or in terms of symbols,

$$X_{n+1} = f\left(x_n, x_{n-1}, \ldots\ldots\ldots t\right) \qquad (2.8.1)$$

This is usually called a difference equation note that before writing it down a considerable amount of thinking and choosing of modelling assumptions is normally necessary. Solving a difference equation means finding and explicit expression for x_n in terms of n and initial value such as x_0. Note, however, that this can be rather a difficult task and not strictly essential, because a model in the form of a difference equation can be used without knowing the mathematically solution of the equation. This is because if we know present the previous values of x we can always use difference equations to generate the net value followed by as many values as we like of course, the advantage of having a formula for x_n in terms of n is that we can substitute any value of n we like into the formula to get an immediate answer.

As a simple example of difference equation, suppose p_n represents an industry's production output in year 'n' and that production doubles every year so that

Net years production=2 x this years production

That is

$$p_{n+1} = 2p_n \tag{2.8.2}$$

If p_0= production in year 0 then $p_1 = 2p_0$, $2p_2 = 2p_1 = 2(2p_0) = 2^2 p_0$ and $p3 = 2p_2 = 2(2^2 p_0) = 2^3 p_0$.

The pattern is clean the p values are going up in a geometric progression and the general solution is $p_n = 2^n p_0$. A similar equation applies when even the growth rate is a constant percentage, for example if the growth rate is 25% per annum. Then the difference equation is $p_{n+1} = (1.25)p_n$ and the solution is $p_n = (1.25)^n p_0$ these are example of the difference equation $x_{n+1} = ax_n$ which corresponds to the assumption that a variable increases in a fixed ratio or (percentage) in each time step. The solution of this is $x_{n+1} = a^n x_0$ (as can be verified by changing in 'n' into n+1) the two statement $x_{n+1} = ax_n$ and $x_n = a^n x_0$ are intact equivalent. The first is the difference equation and the seconds its solution.

This kind to solution is very common with investment where interest accumulates at a constant r % per annum. If p_0 is the initial amount of money invested, the amount p_n after 'n' years satisfies the equation $p_{n+1} = \left(1 + \frac{r}{n}\right)p_n$ and the second is its solution

The next simplest type of difference equation is the first-order liner constant coefficient case, which we can write as

$$x_{n+1} = a.x_n + b \tag{2.8.3}$$

Starting with x_0 we get

$X_1 = aX_0 = b$ and $x_2 = ax_1 + b = a(ax_0 + b) + b = a^2 x_0 + (a+1)b$ its follows that

$x_3 = ax_2 + b = a\left(a^2x_0 + \left(a+1\right)b\right) + b$

$= a^2x_0 + \left(a^2 + a + 1\right)b$

Constituting in this way;

$$x_n = a^n x_0 + \left(a^{n-1} + \ldots\ldots\ldots\ldots\ldots + a^2 + a + 1\right)\ b = a^n x_0 + \frac{a^{n-2}b}{a-1} \qquad (2.8.4)$$

The last step comes from summing the geometric series and we have assumed that 'a' does not equal 1. If a does equal 1 then we can see that $x_2 = x_0 + 2b, x_3 = x_0 + 3b$ and so on, so in this case the solution is just $x_n = x_n + nb$ with some discrete modules of this kind we find that as 'n' increases the c values are approximating a "limit or equilibrium values", although they never quite reach it, in thus case x_{n+1} eventually becomes indistinguishable from then $x_{n+1} = x_n = L$, so from the difference equation $L = aL + b$ and therefore the equilibrium value is $L = \frac{b}{1-a}$ (provided we are not dealing with the case a=1) of course If we happen to start with $x_0 = L$ then we have $x_n = L$ forever.

A difference equation connecting x_{n+1}end x_n and no other x values is called a first-order difference equation. If the equation also involved x_{n-1}or x_{n-2}we would call it second order in other words, the order is the difference between the highest and lowest subscripts appearing in the equation. Not surprisingly, first-order difference equations are the easiest to deal with more significant than the order is the question of whether the difference equation is linear. An example of a linear difference equation (this one happens to be second order) is

$$x_{n+1} = 2\text{-}x_n + 3\text{-}\ x_{n-1} + n^2 + 7 \qquad (2.8.5)$$

Note that the presence of the n^2 does not make the equation non-linear. The important thing is that all the x-terms are only multiplied by constants. A simple(and important) example of a non-linear difference equation is

$$x_{n+1} = ax_n\left(1 - x_n\right) \tag{2.8.6}$$

where 'a' is a constant. The solution of non-linear equations reveal a much strange and more varied behavior than that of linear equations and in some cases show the kind of behavior described as 'chaos'. Difference equations are easiest to solve when there are homogeneous. This means that the equation can be satisfied by making the entire x' s equal to 0. For example the equation $x_{n+2} - 3\text{-}x_{n-1} + x_n = 0$ is homogeneous. While $x_{n+2} - 2xn = 3x_{n+1}$ is not. It also makes things easier if the coefficients are constant. For example $x_{n+2} - 3\text{-}x_{n-1} + x_n = 0$ has the constant coefficients, while $x_{n+2} - 3nx_{n-1} + x_n = 0$ does not.

2.8.1 MORE THAN ONE VARIABLE

Suppose that in a battle between two opposing faces each unit of army A is able to destroy units of Army B charging one time unit. Similarly each unit of Army B is able to destroy b units of Army A.

Let A_n denote the number of units of Army A surviving after n time steps, and similarly B_n for Army B. We therefore have two variables and their fates are obviously linked. How does this connection appear in a mathematical model? The answer is obtained by considering what happens in one time step.

The total number of units of Army A destroyed during that time step is bB_n because every one of the B_n units of Army B destroys b units of Army A. The number of surviving units of Army A at the beginning of the next time step is therefore.

$$A_{n+1} = A_n - bB_n \tag{2.8.7}$$

And similarly B,

$$B_{n+1} = B_n - aA_n \tag{2.8.8}$$

Here we have two explicit but coupled difference equations, and neither one can be solved on its own. However, given initial sizes A_0 and B_0 for the two armies, and also given the parameters *a* and *b*,we could compute A_1 and B_1 etc. directly from the above difference equations.

An alternative approach is to eliminate one of the variables by substitution. The first equation implies that

$$A_{n+2} = A_{n+1} - bB_{n+1}$$

$$= A_{n+1} - b\left(B_n - aA_n\right) \qquad \text{(substituting for } B_{n+1}\text{)}$$

$$= \mathrm{A_{n+1} + abA_n + \left(A_{n+1} - A_n\right)} \qquad \text{(substituting for } B_n\text{)}$$

This is

$$A_{n+2} - 2A_{n+1} + \left(1 - ab\right)A_n = 0 \tag{2.8.9}$$

a second order difference equation. From it can be generate a sequence of A_n values provided two starting values, e.g. A_0 and A_1 are available. Alternatively a mathematical solution for A_n in terms of *n* can be derived.

2.8.2 MATRIX MODELS

Linear difference equations involving more than one variable can be neatly expressed using vectors and matrices. The state of the battle in the previous example can be represented by the vector $x_n = \{A_n \ , B_n\}$ and pair of simultaneous difference equations can be written

$$\begin{bmatrix} A_{n+1} \\ B_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & -a \\ -b & 1 \end{bmatrix} \begin{bmatrix} A_n \\ B_n \end{bmatrix} \tag{2.8.10}$$

so the progress of the battle form one step to the next step can be written concisely as $x_{n+1} = Mx_n$,

Where M is the matrix

$$\begin{bmatrix} 1 & -a \\ -b & 1 \end{bmatrix} \tag{2.8.11}$$

and the solution can be written $x_n = M^n x_0$. This approach is especially useful for models representing transitions between states or compartments. When developing models for populations, for example, we often want to do more than just predict the total size of the population. At any time a human population will consist of a mixture of individuals of different age, sex, occupation etc. in order to make forward planning for the provision of resources such as schools and hospitals we need to make predictions about the future number of individuals in different categories within the population.

Let us take a simple and artificial example of a population of animals which become adult and capable of reproducing at the age of one year. Suppose we represent the population at time step *n* in terms of the numbers of animals in each of three categories.

B_n = number of babies and young animals up to one year old

A_n = number of young adult up to two years old

S_n = number of senior adults aged two or older

Then will be different annual birth and death rates for the three groups. Suppose we have the following information

Group	**Birth rate**	**Death rate**
B	0	0.1
A	0.3	0.2
S	0.1	0.3

This means, for example, that 90% of babies survive to become adults and that 10% of senior adults produce one offspring per year on average.

We can put this information into matrix form as

$$\begin{bmatrix} B \\ A \\ S \end{bmatrix}_{n+1} = \begin{bmatrix} 0 & 0.3 & 0.1 \\ 0.9 & 0 & 0 \\ 0 & 0.8 & 0.7 \end{bmatrix} \begin{bmatrix} B \\ A \\ S \end{bmatrix}_{n} \qquad (2.8.12)$$

and to find what happens to the population we only have to keep multiplying by the matrix(often called the transition matrix). The nature of the solution and whether we eventually reach a steady state depends on the largest eigenvalue, λ, of this matrix.

- If $\lambda > 1$ then the population grows without limit.
- If $\lambda = 1$ then the population converges to the eigenvector associated with λ.
- If $\lambda < 1$ then the population continually decreases.

2.9 CONTINUOUS MODELS

2.9.1 LINEAR MODELS

Linear models are a class of mathematical models used to represent relationships between variables in a linear fashion. These models are widely used in various fields, including statistics, mathematics, science, engineering, economics, and social sciences, due to their simplicity and interpretability. Here are some key characteristics and types of linear models:

Characteristics of Linear Models:

Linearity: Linear models assume that the relationship between the variables is linear. This

means that when you plot the data on a graph, it forms a straight line (or a plane in higher dimensions).

Additivity: Linear models assume that the effects of individual variables on the response variable are additive. In other words, changes in one predictor variable have a constant effect on the response variable, regardless of the values of other predictors.

Parameters: Linear models involve parameters that need to be estimated from the data. These parameters include coefficients (slopes) for each predictor variable and an intercept (constant term).

Interpretability: Linear models are highly interpretable. The coefficients in the model represent the change in the response variable associated with a one-unit change in the predictor variable, holding all other variables constant.

Assumptions: Linear models make certain assumptions, such as the normality of residuals (errors), constant variance of residuals (homoscedasticity), and independence of observations. Violations of these assumptions can affect the validity of the model.

The following facts from coordinate geometry are useful

Gradient (slope) or straight line = $\frac{increase\,in\,y}{increase\,in\,x}$

$$= \frac{y_2 - y_1}{x_2 - x_1}$$

if $(x_1, y_1)\,and\,(x_2, y_2)$ are two points on the line

Equations of straight line:

1. With slope *m* and intercept *c* on the *y* axis $y=mx+c$
2. With intercepts *a* and *b* on the x-and y-axes respectively: $\frac{x}{a} + \frac{y}{b} = 1$
3. Passing through the point (x_0, y_0) with gradient m : $y - y_0 = m(x - x_0)$
4. Joining the points (x_1, y_1) and (x_2, y_2): $\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}$

Linear Models with Several Variables

If the value of variable *y* is thought to depend on the values of a no.of. other variables x_1, x_2........... the simplest way of expressing the dependence is through a linear model of the form $y = a + b_1x_1 + b_2x_2 + \ldots\ldots\ldots\ldots$ the condition for this kind of model to be valid is that *y* changes by equal amounts for equal changes in any one of the variables. This model is no more difficult to deal with than the single variable case, except when it comes to graphical interpretation.

Simultaneous Linear Models

We may have two or more dependent variables, all of which are modeled as linear functions of *x*. Questions of interest which then arise are: when are two variables equal (i.e. where do the lines cross) and when does one come above the other?

Suppose, for example, that we have a choice between hiring two machines. Machine A can be hired for Rs.25 a week while machine B can be hired for Rs.150 plus Rs.10 per week. Which is the cheaper machine if we want to use it for *x* weeks?

The cost of hiring A for *x* weeks is $y_A = 25x$ and for B the cost is $y_B = 150 + 10x$. The two are equal when $25x = 150 + 10x$, i.e. $x = 10$. Also when $x < 10$ we have $y_A < y_B$ with the opposite being true when $x > 10$. The conclusion from this is that if we need a machine for less than 10 weeks the cheaper choice is A, otherwise choose B.

Piecewise Linear Models

A model does not have to be represented by the same single formula for all values of the variables *x*. for example, suppose certain items cost Rs.10 each to by, but if you buy more than 100, the price of any extra items drops to Rs.9 per item. The model for the cost of buying *x* items is then

$$y = \begin{cases} 10x & 0 \le x \le 100 \\ 1000 + 9(x-100) = 100 + 9x & x \ge 100 \end{cases} \qquad (2.9.1)$$

The two different linear expressions agree at x=100, so there in no sudden jump (discontinuity) at the change over point.

In the previous example a piecewise linear model occurred naturally. We may sometimes choose to model a non-linear function approximately by a piecewise linear function. For example, suppose a car does about 40 miles per gallon of patrol at a speed of 30m.p.h decreasing to 20miles per gallon at a speed of 70 m.p.h. and also decreasing to zero as the speed decreases from 30 m.p.h to zero. If we don't know the detailed shape of the graph we could represent the mileage rate *R* as piecewise linear function of speed (V) using

$$R = \begin{cases} \dfrac{4V}{3} & 0 \le V \le 30 \\ 55 - \dfrac{V}{2} & 30 \le V \le 70 \end{cases} \qquad (2.9.2)$$

2.9.2 QUADRATIC MODELS

When a variable *y* does not change by equal amounts for equal changes in the *x* variable than a linear model is not suitable. A simple example of a non-linear model is the quadratic $y = ax^2 + bx + c$, whose is a parabola. Three separate pieces of information are needed to determine the three parameters *a, b* and *c* . The value of *a* determines whether the curve is concave upwards (if a>0) or concave downwards. There is a vertical axis of symmetry at $x = -\dfrac{b}{2a}$,which is also the *x* value at which the graph has a global maximum or minimum valve.The valve of the parameter *c* affects the vertical position of the curve relative to the coordinate axes.

2.9.3 OTHER NON-LINEAR MODELS

More flexibility is obtained by using higher degree polynomials with *n+1* pieces of information being needed to determine the *n+1* coefficients of a polynomial of degree *n*. There are also very simple non-linear models which are not polynomials, for example models expressing inverse relationship between the variables such as $y = \frac{k}{x}$ or $y = \frac{k}{x^2}$. Other examples are models based on rational functions of the form

$$\frac{a+bx}{c+dx}, \frac{a+bx+cx^2}{d+ex+fx^2}, ... \quad (2.9.3)$$

and models expressed in terms of standard mathematical functions such as square roots, exponentials and logarithmic and trigonometric functions. The choice of an appropriate form is based on a mixture of experience and experimentation the question of the best values to take for the parameters is the subject of fitting models to the data. It is useful to have insight into the effects of charging each parameter in turn. Note that the effect of replacing the variable x by x-c in a model is to move the curve horizontally *c* units to the right without altering its shape. The effect of adding *c* to the dependent variable *y* is to shift the curve vertically upwards without altering its shape. Replacing *x* by cx appears to alter the shape of the curve but can also be regarded as zooming in $(C > 1)$ or zooming out $(C < 1)$. A curve $y = f(x)$ can be reflected in a horizontal line $\mathrm{y = c}$ by taking $y = 2c - f(x)$, and for the reflection in a vertical line x=c take $y = f(x-c)$

2.9.4 MODELS TENDING TO A LIMIT

Very often physical variables increase gradually towards some upper limit or ceiling. An example is a living population whose size is limited by environmental factors examples of mathematical models with this kind of the behavior are

1. $a - be^{-ct} (\to a \ \ as \ \ t \to \infty)$
2. $\frac{at+b}{ct+d} \ \ when (ad > bc) \ (\to a/c \ \ as \ \ t \to \infty)$
3. $a + b\tan^{-1}(ct) \ \ (\to a + \frac{b\pi}{2} \ as \ \ t \to \infty)$
4. $\frac{1}{a+be^{-ct}} \ \ (\to \frac{1}{a} \ \ as \ \ t \to \infty)$

In these expressions *t* is time and *a, b, c and d* are positive constants.

The above examples are easily adopted to give curves that decrease gradually to a lower limit or floor, for example

1. $a + be^{-ct} (\to a \ \ as \ \ t \to \infty)$
2. $\dfrac{at+b}{ct+d} \ \ when(ad < bc) \ (\to a/c \ as \ \ t \to \infty)$
3. $a - b\tan^{-1}(ct) \ \ (\to a - \dfrac{b\pi}{2} \ as \ \ t \to \infty)$
4. $\dfrac{1}{a - be^{-ct}} \ \ (\to \dfrac{1}{a} \ \ as \ \ t \to \infty)$

2.9.5 TRANSFORMING VARIABLES

Starting from a variable with an infinite range, such as $[0,\infty] or [-\infty,\infty]$ we can construct variables with a finite range by using mathematical transformation. Suppose for example, that x goes from 0 to ∞. To get a range from 0 to 1 we could use for example $U = \dfrac{x}{1+x}$. We could make the range [a,b] by taking $U = \dfrac{ax+b}{1+x}$. Alternative ways of achieving the same thing are $U = a + (b-a)(1-e^{-x})$ and

$$U = a + 2(b-a)(\tan^{-1} x)/\pi \qquad (2.9.4)$$

Transforming variables will usually alter the shape of a curve and this can be useful in straightening curves. For example, if $y = kx^n$ then plot of $\ln y$ against $\ln x$ is a straight line and plotting logs often changes non-linear models into nearly straight lines.

2.10 MODELLING RATES OF CHANGE

One of the main points of modelling is try to predict what will happen, that is, we try to model how things change with time. Time itself can be modelled as continuous or discrete, and this is one of the modelling decisions which often have to be made. Sometimes there is a natural time interval at the end of which something happens, for example the breeding season of a living population. In cases like these a discrete model of time is probably appropriate, with a time step equal to the time between breading season. The variable whose rate of change we are modelling can also be continuous or discrete. For discrete variable whose valve after *n* time steps is x_n , the change or increment during the *n*th step is $\Delta x_n = \Delta X_{n+1} - X_n$

The progression of x_n with time can be shown in a no.of ways, the most obvious being a plot of x_n against n. It could also be useful to plot the incremental change Δx_n against *n*. Another possibility is to calculate the relative growth in each time interval from $R_n = \frac{\Delta X_n}{X_n}$ which gives the growth as a proportion of the value of the variable at the beginning of the interval. This third possibility could be relevant in modelling the growth of an animal. Foe example, the table gives the weight in pounds of one baby during the first few months of her life

Age(months)	n	0	1	2	3	4
Weight(lb)	X_n	8	10	13	16	20
	ΔX_n		2	3	3	4
	R_n		0.250	0.300	0.231	0.250

We see that the absolute rate of growth for this period (as shown by ΔX_n) was greatest in the fourth month, but the relative rate of growth (as shown by R_n) peaked in the second month.

Discrete models are obtained by making assumptions about the rate of growth which then translate into a difference equation satisfied by X_n. The simplest types are the geometric progression $\Delta X_{n+1} = aX_n$ and the linear form $\Delta X_{n+1} = aX_n + b$. It is useful to remember that a percentage change is equivalent to a multiplication, so if *X* changes by r% it becomes $\left(1+\frac{r}{100}\right)X$. If, for example, the price of an item goes up by 3%, 4% and 2% in three consecutive years its final price is (1.03)(104)(1.02)= (1.092) times the original price, so the net increase in price over the three year period is 9.2624%.

The rate of inflation is an introduction of the rate of increase in the average level of prices. This is measured by comparing the retail price index with the value of the index twelve months earlier. The inflation rate is then calculated as an annual percentage rate, say I%. The actual annual percentage change (say *X*%) in a particular cost or income can be misleading because it disregards the effect of inflation. The increase 'in real term' or 'allowing for inflation' is required. A very common mistake is to take differences *(X-I)* % which comes from forgetting the multiplicative nature of percentage changes as mentioned in the previous paragraph. The fact that this is wrong can be seen by noticing that if inflation is *I*% the value of Rs1 decreases to Rs $1/\left(1+\frac{I}{100}\right)$ in one year. An investment yielding an interest of *X*% p.a. gives for each Rs.1 an amount Rs.$\left(1+\frac{X}{100}\right)$ after one year, but these are deflated Rs worth $1/\left(1+\frac{I}{100}\right)$each. The effective rate of interest is therefore given by$\left(1+\frac{X}{100}\right)/\left(1+\frac{I}{100}\right)$.

When continuous model is chosen the techniques of differential calculus become very relevant. If $f(t)$ is a function of time then as time changes from t to $t+\Delta t$, $f(t)$changes from $f(t)$ to $f(t+\Delta t)$ and the change is represented by Δf. The quantities Δt and Δf are referred differentials.

The average rate of change of f is $\frac{change\, in\, f}{change\, in\, time}$, which is the ratio of the differentials or

$$\frac{f(t+\Delta t)-f(t)}{\Delta t} \tag{2.10.1}$$

Letting $\Delta t \to 0$ we get

$$\frac{\Delta f}{\Delta t} \to \frac{df}{dt} \tag{2.10.2}$$

Which is the instantaneous rate of change of f with respect to *t*. often referred to as the derivative and abbreviated to f^1. The differentials are related by

$$\Delta f \approx \frac{df}{dt}\Delta t \tag{2.10.3}$$

When more than one variable changes at the same time we use partial derivatives. In this example *h* is a function of θ and x .if θ Changes by $\Delta\theta$ and x changes by Δx the change in *h* is

$$\Delta h = \frac{\partial h}{\partial \theta}\Delta\theta + \frac{\partial h}{\partial x}\Delta x \tag{2.10.4}$$

In geometric terms the value of f^1 at any point gives us. The gradient of the tangent of the graph of f at that point.

If f^1 Is positive and increasing then the graph of *f* is concave up $\Rightarrow$ f is increasing at an increasing rate ($f^{11} > 0$).

If f^1 is positive but decreasing then the graph of *f* is convex up(or concave done) $\Rightarrow$ f is increasing at a decreasing rate($f^{11} < 0$)

The rate of change f^1 is the second derivative, f^{11} ,which tells us how quickly the rate of change is itself changing.

In economics, if C(x) is the cost producing *x* units, $C^1(x)$ is often called the marginal cost and is the approximate cost of producing the next unit. Another concept from economics is elasticity of demand. If Q(P) is the demand when the price is P, the elasticity is $E = -\frac{d(\ln Q)}{d(\ln P)} = -\left(\frac{P}{Q}\right)\left(\frac{dQ}{dP}\right)$. Its significance is that if demand is inelastic. (E<1) then raising the price will increase the revenue (=QP), while the reverse is true for elastic (E>1) demand.

Very often we have *y*= a function *x*, where *x*= a function of *t*, in which case we can find the rate of change of *y* from

$$\frac{dy}{dt} = \frac{dy}{dx}.\frac{dx}{dt} \qquad (2.10.5)$$

The fact that we have found a local minimum for the area and not a local maximum is obvious from the observation that we can make *A* is a large as we like by choosing *x* very small or very large. We do not need to calculate the second derivative etc. in pedantic fashion when we can use practical sense. This is the pragmatic modelling approach as opposed to the pure mathematical.

2.10.1 DISCRETE OR CONTINUOUS

In some problems a variable is very obviously discrete and no other way of representing it would be sensible. The harvesting of a crop, for example, usually happens once a year, not continuously though the year. In other cases the choice between a discrete model and a continuous one can be difficult. It is by no means true that the discrete choice is the easiest. It very much depends on the problem being investigated. Generally speaking a continuous model is suitable when the time interval between measurements is small. In going from a discrete model to a continuous one, the change in a unit time interval, which is $y_{n+1} - y_n$, is replace by y^1.

One advantage of a continuous model is that a differential equation may be easier to manipulate and solve than a difference equation; and(if we solve it exactly) we finish up with a solution which we can use for any value of *t*. paradoxically, if we use a numerical method to solve the continuous differential equation, we are in fact changing the problem back into discrete from! Note that some problems require both discrete and continuous elements together, for example, grass may grow continuously, butt is usually only cut on certain days.

The simplest models are

Discrete

Arithmetic $X_{n+1} = X_n + c$

Geometric $X_{n+1} = aX_n$

Linear first order $X_{n+1} = aX_n + b$

Continuous

Linear $y = a + bt$

Power law $y = at^b$

Polynomial $y = a + bt + ct^2 + dt^3 + ...$

Exponential $y = ae^{bt}$

Note the distinction between the power law and exponential models(the t is in the exponent in exponential model).Also note that $y = a^t$ and $y = e^{kt}$ are equivalent since $\ln y = t.\ln a$ In the first case and $\ln y = kt$ in the second, and they are the same if $k = \ln a$

CHAPTER - 3 STATISTICAL GROWTH MODELS

3.1 INTRODUCTION

Statistical aspects play a crucial role in the development, estimation, validation, and interpretation of growth models. Whether you are working with biological, economic, or population growth models, statistical techniques help you make inferences, quantify uncertainty, and assess the goodness of fit of your models. Here are some key statistical aspects associated with growth models:

Parameter Estimation: Growth models often have parameters that need to be estimated from observed data. Statistical methods like maximum likelihood estimation (MLE) or least squares fitting are used to find the values of these parameters that best fit the model to the data.

Model Selection: Statistical criteria, such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC), are used to compare different growth models and select the one that provides the best trade-off between goodness of fit and model complexity.

Model Validation: Statistical tests and diagnostics are applied to validate growth models. Residual analysis, hypothesis tests, and goodness-of-fit tests (e.g., chi-square test) can help assess whether the model adequately describes the data.

Confidence Intervals: Statistical techniques are used to calculate confidence intervals around parameter estimates. These intervals provide a measure of uncertainty about the true values of the parameters.

Heteroscedasticity: Growth models often assume homoscedasticity (constant variance of errors), but in practice, variances may change with time or other factors. Statistical tests for heteroscedasticity help identify violations of this assumption.

Autocorrelation: In time series growth models, autocorrelation (correlation of observations with past observations) can be a concern. Statistical tests, such as the Durbin-Watson test, help detect and correct for autocorrelation.

Prediction Intervals: In addition to point estimates of future values, statistical methods can provide prediction intervals, which give a range within which future observations are likely to fall. These intervals account for both parameter uncertainty and residual variability.

Model Assumptions: Growth models make assumptions about the distribution of errors (e.g., normally distributed errors in linear regression) and other underlying statistical properties. Checking these assumptions is crucial for model validity.

Outlier Detection: Statistical techniques can identify outliers or influential data points that may disproportionately affect model results. Robust regression methods can be used to mitigate the impact of outliers.

Nonlinearity: Growth models may involve nonlinear relationships. Statistical methods for nonlinear regression are used to estimate parameters in these cases.

Bootstrap Resampling: The bootstrap method can be used to estimate the sampling distribution of model parameters or to assess model uncertainty when the assumptions of traditional statistical tests are not met.

Bayesian Growth Models: Bayesian statistics offers a framework for estimating parameters and making predictions in growth models while explicitly incorporating prior information and quantifying uncertainty through posterior distributions.

Model Diagnostics: Various graphical and statistical tools, such as residual plots, Q-Q plots, and leverage plots, help diagnose potential issues with growth models, including heteroscedasticity, nonlinearity, and influential observations.

Cross-Validation: Cross-validation techniques, such as k-fold cross-validation, are used to assess the predictive performance of growth models and to prevent overfitting.

Time Series Analysis: Time series models, including autoregressive integrated moving average (ARIMA) models, are applied to growth data with temporal dependencies, helping to capture trends, seasonality, and autocorrelation.

Spatial Growth Models: For spatial growth models, spatial statistics, geostatistics, and spatial autocorrelation techniques are used to account for spatial dependencies and patterns.

Statistical aspects are essential for ensuring the robustness, validity, and generalizability of growth models. Effective use of statistical methods can help researchers and analysts draw meaningful conclusions from growth data and make accurate predictions about future growth trends.

Business is a dynamic affair and dynamism is related to time factor. There are many factors which change with the passage of time, i.e., as time passes, their values also changes. For example, the sales of a product, the population of a country, demand of commodities, prices etc., may increase with time. Arrangement of statistical data on a study variable in chronological order i.e., in accordance with occurrence of time is known as 'Time Series Data'.

3.2 MEASURES OF LINEAR AND COMPOUND GROWTH RATES

3.2.1 LINEAR GROWTH RATE

The linear growth rate, often referred to as the constant linear growth rate, is a measure of how a quantity increases or decreases over time in a linear fashion. It represents the rate at which a quantity changes by a fixed amount per unit of time. In a linear growth or decay model, the quantity increases or decreases at a constant rate over time.

The Linear Growth Rate (LGR) in a study variable (*Y*), for an absolute change in a time variable (*t*) is defined as the ratio of relative change in *Y* to the absolute change in t, multiplying by 100.

i.e.,

$$LGR = \left[\frac{relative\,change\,in\,Y}{absolute\,change\,in\,t}\right]100 \qquad (3.2.1)$$

Symbolically, for small changes in *Y* and *t*, LGR may be approximated by

$$\text{LGR}=\left[\frac{\Delta Y/Y}{\Delta t}\right]100=\left[\frac{(Y_2-Y_1)/Y_1}{t_2-t_1}\right]100 \qquad (3.2.2)$$

or

$$\text{LGR}=\left[\frac{(Y_2-Y_1)/Y_2}{t_2-t_1}\right]100 \tag{3.2.3}$$

or

$$\text{LGR}=\left[\frac{(Y_2-Y_1)/\bar{Y}}{t_2-t_1}\right]100 \tag{3.2.4}$$

Where $\bar{Y}=\frac{Y_1+Y_2}{2}$

Here, Y_1 and Y_2 are the values of Y for the time periods t_1 and t_2

3.2.2 COMPOUND GROWTH RATE

Compound growth rate, often referred to as the compound annual growth rate (CAGR), is a measure used to determine the annual growth rate of an investment, asset, or quantity when the value changes over multiple periods. CAGR takes into account the compounding effect, which means that each year's growth builds upon the previous year's growth. It provides a single, consistent growth rate over a specified time frame, allowing for easy comparison of different investments or the evaluation of an investment's performance.

Suppose Y_0 be the initial value of study variable Y and it will grow into the value Y_n after n years of annual compounding at the growth rate r per annum, then, by compound interest formula, we have,

$$Y_n = Y_0\left(1+\frac{r}{100}\right)^n \tag{3.2.5}$$

Given the values of Y_0, Y_n and n, the Compound Growth Rate (CGR) is given by

$$\text{CGR}= r = \left[\left(\frac{Y_n}{Y_0}\right)^{\frac{1}{n}} - 1\right]100 \tag{3.2.6}$$

3.3 STATISTICAL ESTIMATION OF LINEAR AND COMPOUND GROWTH RATES

Under statistical method of estimation, we first fit a linear or an exponential model to the given time series data by applying the Least Squares method. Later, the LGR or CGR may be computed by using the estimate of the parameter of the model. Since, this method uses all the observations in the data; it gives better approximation than the mathematical measure for the growth rate. By this method, the significance of the growth rate can also be tested.

We denote the coded time variable by *X* with the starting period as 1 and subsequent periods by 2,3,4…. n. Let the sample values of the study variable in sequence of time be $Y_1, Y_2, \ldots Y_n$. Now, the time series data may be represented as:

Time (t)	:	t_1	t_2 …………	t_n
Coded time (X)	:	1	2…………	n
Study variable	:	Y_1	Y_2	Y_n

(a) Linear Growth Rate (LGR)

Suppose there exists a linear relationship between a study variable (*Y*) and a time variable (*t*) as

$$Y_i = a + bt_i, \quad i = 1, 2, \ldots\ldots., n \tag{3.3.1}$$

By using g coded time variable *X* in the place of t, the linear model can be written as

$$Y_i = a + bX_i, \quad i = 1, 2, \ldots\ldots., n \tag{3.3.2}$$

or simply, $Y = a + bX$. By adding an error term ε, the statistical linear regression model is given by

$$Y = a + bX + \varepsilon \tag{3.3.3}$$

where, *Y*: dependent variable (Study variable)

X: Independent variable (Coded Time variable)

and *a*, *b* are the parameters of the linear model. The Least Squares Estimates (Best Estimates) of *a* and *b* are given by

$$\hat{b} = \left[\frac{\sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{n}}{\sum X^2 - \frac{\left(\sum X\right)^2}{n}} \right] \tag{3.3.4}$$

And

$$\hat{a} = \bar{Y} - \hat{b}\,\bar{X} \tag{3.3.5}$$

Here $\bar{X} = \frac{\left(\sum X\right)}{n}, \bar{Y} = \frac{\left(\sum Y\right)}{n}$ and *n* is the of observations. The estimated linear model is given by $\hat{Y} = \hat{a} + \hat{b}\,X$. This estimated model will be used for the prediction analysis.

An estimate of LGR is now given by

$$\text{LGR} = \left[\frac{\hat{b}}{\bar{Y}} \right] 100 \tag{3.3.6}$$

3.3.1 TEST OF SIGNIFICANCE OF LGR

To test for the significance of LGR, we use the following student's t-test statistic:

$$t = \frac{\hat{b}}{S.E(\hat{b})} \qquad (3.3.7)$$

or

$$t = \frac{\hat{b}\sqrt{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}}{\hat{\sigma}} \qquad (3.3.8)$$

Where

$$\hat{\sigma} = \sqrt{\frac{\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right] - \hat{b}\left[\sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{n}\right]}{n-2}} \qquad (3.3.9)$$

We compare the calculated value of $|t|$ with its critical value (table value) for (n-2) degrees of freedom at a desired level of significance and draw the inference accordingly. i.e., if the calculated valve of $|t|$ is greater than its critical value, then the LGR is said to be significant at the desired level of significance. Otherwise, LGR is said to be not significant.

Remark:

Here, the Standard Error (SE) of $\hat{b}$ is given by

$$S.E(\hat{b}) = \frac{\hat{\sigma}}{\sqrt{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}} \qquad (3.3.10)$$

(b) COMPOUND GROWTH RATE

Consider the formula for compound interest as

$$Y_n = Y_0\left(1+\frac{r}{100}\right)^n \tag{3.3.11}$$

If we replace Y_n by Y ; Y_0 by a, $\left(1+\frac{r}{100}\right)$ by *b* and *n* by a coded time variable X, then we will have an exponential functional relationship between *Y* and *X* as

$$Y = ab^X \tag{3.3.12}$$

Where Y: Dependent variable (study variable)

X: Independent variable (coded time variable)

and *a*,*b* are the parameters of the exponential model. Since, directly fitting of exponential model involves several difficulties, we generally transform the model into a log-linear form and then it may be fit to the data.

Taking logarithms on both sides of the model, we get

$$\log Y = \log a + X \log b \tag{3.3.13}$$

or

$Y = A + BX$,which is a linear model.

Here, $Y = \log y, \quad A = \log a, \quad B = \log b$

The least squares estimates of A and B are given by

$$\hat{B} = \left[\frac{\sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{n}}{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}\right], \tag{3.3.14}$$

$$\hat{A} = \bar{Y} - \hat{B}\bar{X} \qquad (3.3.15)$$

Here, *n* is the number of observations.

The estimated linear model is given by $\hat{Y} = \hat{A} + \hat{B}X$. An estimated of original parameter b is given by $\hat{b} = Anti\log(\hat{B})$

Thus,

$$\hat{b} = \left(1 + \frac{\hat{r}}{100}\right)b \qquad (3.3.16)$$

Hence, an estimate of CGR is now given by

$$\text{CGR} = \hat{r} = (\hat{b} - 1)100 \qquad (3.3.17)$$

3.3.2 TEST OF SIGNIFICANCE OF CGR

To test for the significance of CGR, we use the following student's t-test statistic

$$t = \frac{\hat{B}}{S.E(\hat{B})} \qquad (3.3.18)$$

or

$$t = \frac{\hat{B}\sqrt{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}}{\hat{\sigma}} \qquad (3.3.19)$$

Where

$$\hat{\sigma} = \sqrt{\frac{\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right] - \hat{B}\left[\sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{n}\right]}{n-2}} \tag{3.3.20}$$

We compare the calculate value of $|t|$ with its critical value for (n-2) degrees of freedom at an appropriate level of significance and draw the inference accordingly.

Remark:

1. Suppose $\hat{b}_1$ and $\hat{b}_2$ be the estimates of parameters of linear models of two time series data respectively. The Linear Growth Rates of the two time series data can be compared by using the following student's t-test statistic as

$$t = \frac{\hat{b}_1 - \hat{b}_2}{\sqrt{\left[S.E(\hat{b}_1)\right]^2 + \left[S.E(\hat{b}_2)\right]^2}} \tag{3.3.21}$$

where, S.E. denotes the standard error of estimate.

We compare $|t|$ value with the critical value of t-test statistic for $(n_1 + n_2 - 4)$ degrees of freedom at a desired level of significance and draw the inference accordingly. Here, n_1 and n_2 a number of observations in two time series data respectively.

2. Suppose we wish to investigate whether there is any change in the growth of production between wartime and peace time periods (or any two different time periods) .Such a change is referred to as a 'structural change in the growth'. It can be tested by using a test given in 'Econometrics' known as Chow test for structural change.

3.4 THE GOMPERTZ CURVE AS A GROWTH CURVE

The Gompertz curve, named after Benjamin Gompertz, is a mathematical function that is often used as a growth curve to describe the growth of biological populations, particularly those that exhibit sigmoidal or S-shaped growth patterns. It has applications in fields such as biology, epidemiology, demography, and economics. The Gompertz curve is characterized by its ability to model both exponential growth in the early stages and decelerating growth in the later stages. Here's a description of the Gompertz curve as a growth curve

In 1825 Benjamin Gompertz published a paper in the philosophical transaction of the Royal Society "On the Nature of the Function Expressive of the Law of Human Mortality", in which showed that "if the average exhaustions of a man's power to avoid death were such that at the end of equal infinitely small intervals of time, he lost equal portions of his remaining power of oppose destruction" then the number of survivors at any age x would be given by the equation.

$$L_x = kg^{c^x} \qquad (3.4.1)$$

(It is clear that Gompertz means equal proportions, not equal absolute amounts, of the "power to oppose destruction.")

The Gompertz come was for long of interest only to actuaries. More recently, however, it has been used by various authors as a growth curve, both for biological and economic phenomena. It is the purpose of the present note to consider some of the mathematical properties of this curve, and to indicate to some extent its usefulness and its limitations as a growth curve.

For actual purposes the curve is generally written in the form (3.4.1), but for our purpose it is more convenient to write it

$$y = k.e^{-e^{a-bx}} \qquad (3.4.2)$$

in which k and b are essentially positive quantities.

From (3.4.2)) it is clear that as x becomes negatively infinite y will approach zero, and as x becomes positively infinite y will approach k. Differentiating (3.4.2) we have.

$$\frac{dy}{dx} = kbe^{a-bx}e^{-e^{a-bx}} = bye^{a-bx} \quad (3.4.3)$$

and it is apparent that the slope is always positive for finite values of x, and approaches zero for infinite values of x. Differentiating again we have

$$\frac{d^2y}{dx^2} = b^2ye^{a-bx}(e^{a-bx}-1) \quad (3.4.4)$$

From (3.4.4) we see that there will be a point of inflection when $x = \frac{a}{b}$

The ordinate at the point of inflection is $y = \frac{k}{e}$

Or approximately, when 37% of the final growth has be reached.

Equations

Gompertz $\quad y = e^{-e^{-x}}$

Logistic $\quad y = \frac{1}{1+e^{-x}}$

When we describe, therefore, to fit growth data, which show a point of inflection in the early part of the growth cycle, when approximately 35 to 40 percent of the total growth has been realized, we may use the Gompertz curve with the expectation that the approximation to the data will be good. There seems, however, no particular reason to expect that the Gompertz curve will show any wider range of fitting power than any other three-constant S-shaped curve. For example the logistic

$$y = \frac{k}{1+e^{a-bx}} \quad (3.4.5)$$

possesses the same number of constants as the Gompertz curve, but has the point of inflection mid-way between the asymptotes. The degree of "skewness" in the Gompertz curve is just as fixed as in the logistic, and it is clear that to introduce a variable degree of skewness into a growth curve will require at least four constants.

To illustrate the mathematical properties of the Gompertz curve and the logistic, the following table has been prepared

Property	Gompertz	Logistic
Equation	$y = ke^{-e^{a-bx}}$	$y = \frac{k}{1+e^{a-bx}}$
Number of constants	3	3
Asymptotes	$y = 0$ $y = k$	$y = 0$ $y = k$
Inflection	$x = \frac{a}{b}$ $y = \frac{k}{e}$	$x = \frac{a}{b}$ $y = \frac{k}{2}$
Straight line form do equation	$\log\log\frac{k}{y} = a - bx$	$\log\frac{k-y}{y} = a - bx$
Summary	Asymmetrical	Symmetrical about inflation
Growth rate	$\frac{dy}{dx} = bye^{a-bx} = by\log\frac{k}{y}$	$\frac{dy}{dx} = \frac{b}{k}y(k-y)$
Maximum growth rate	$\frac{bk}{e}$	$\frac{bk}{4}$
Relative growth rate as function of time	$\frac{1}{y}\frac{dy}{dx} = be^{a-bx}$	$\frac{1}{y}\frac{dy}{dx} = \frac{b}{1+e^{-a+bx}}$
Relative growth rate as function of size	$\frac{1}{y}\frac{dy}{dx} = b(\log k - \log y)$	$\frac{1}{y}\frac{dy}{dx} = \frac{b}{k}(k-y)$

The parallelism between the Gompertz curve and the Logistic may be carried further. It has been found useful, for example, to add a constant term to the Logistic, giving it a lower asymptote different from zero.

$$y = d + \frac{k}{1+e^{a-bx}} \tag{3.4.6}$$

Clearly this procedure is equally applicable to the Gompertz curve, giving

$$y = d + ke^{-e^{a-bx}} \tag{3.4.7}$$

It is also clear that in general the sum, or the average, of several Gompertz curves will not be a Gompertz curve, just as several logistics do not; in theory, give a logistic when added or when averaged. But just as it has been found in practice that the sum of a number of logistics does in fact often approximate closely a logistic as has shown by reed and pearl(1927). It will be true that Gompertz curves will often add to give something very close to a Gompertz curve. And the general theory of averaging growth curves work by Merrell (1931) and applied by her to the logistic can be applied without modifying to the Gompertz curve.

It may be further pointed out that the Gompertz curve may be generalized in the manner in which peral and reed (1923) have generalized the logistic. Pearl and Read tset

$$\frac{dy}{dx} = \frac{b}{k} y(k-y) f(x) \tag{3.4.8}$$

And on integration obtain.

$$y = \frac{k}{1+e^{F(x)}} \tag{3.4.9}$$

And assuming that F(x) can be expressed as a Taylor's series, they reach

$$y = \frac{k}{1+e^{a_0x+a_1x+a_2x^2+a_3x^3+\dots\dots\dots\dots}} \tag{3.4.10}$$

In a similar fashion, the differential equation of the Gompertz curve may be written

$$\frac{dy}{dx} = by(\log k - \log y) \tag{3.4.11}$$

and if we add an arbitrary function of time on the right hand side of this equation

$$\frac{dy}{dx} = by(\log k - \log y) f(x) \tag{3.4.12}$$

We have on integration

$$y = ke^{-e^{F(x)}} \tag{3.4.13}$$

and if F(x) can be expressed in a Taylor's series. We have

$$y = ke^{-e^{a_0x+a_1x+a_2x^2+a_3x^3+.................}} \qquad (3.4.14)$$

It will be noted that if we wish to use only a finite number of terms in the power series, we must keep and odd power of *x* for our highest term, if our curve is to run from *y*=0 to *y*=*k*. This restriction applies to both the generalized logistic and the generalized Gompertz curve.

We may rationalize the derivation of the Gompertz curve along the lines indicated by Ludwig (1929). Ludwig postulates that the relative growth rate $\frac{1}{y}\frac{dy}{dx}$ must decrease monotonically with continued growth.

If now we write

$$\frac{1}{y}\frac{dy}{dt} = m - ny \qquad (3.4.15)$$

We have the differential equation of the logistic. In this case the relative growth rate decreases as a linear function of growth already reached.If we set

$$\frac{1}{y}\frac{dy}{dx} = pe^{-qx} \qquad (3.4.16)$$

Characteristics:

Sigmoidal Shape: The Gompertz curve has an S-shaped or sigmoidal shape, which means that initially, growth is close to exponential, then it decelerates, and eventually levels off as the population approaches an asymptote.

Inflection Point: The inflection point is the point on the curve where the rate of growth is at its maximum. This is also the point where the curve changes from accelerating growth to decelerating growth.

Asymptote: The asymptote is a horizontal line that the curve approaches but never reaches. It represents the maximum attainable population or quantity.

Parameters: The curve is defined by the parameters

These parameters determine the initial population, growth rate, and timing of the inflection point, respectively.

Applications:

The Gompertz curve is used in various applications:

Population Biology: It can describe the growth of populations of organisms, such as bacteria, animal populations, or tumor growth.

Epidemiology: The Gompertz curve is used to model the spread of diseases within a population. It helps predict when an epidemic is likely to peak and when it will slow down.

Demography: In demography, the Gompertz curve can model the mortality rate of a population as it ages. The curve shows an increasing mortality rate with age.

Economics: It can be applied to modeling the adoption of new technologies or products in a market, where initial adoption is rapid, but growth slows as saturation is reached.

Quality Control: In manufacturing and quality control, the Gompertz curve can be used to model the failure rates of products over time.

Finance: The Gompertz curve can describe the growth of investment portfolios or financial assets over time.

Overall, the Gompertz curve is a versatile tool for modeling growth and saturation processes in a variety of fields. Its sigmoidal shape and parameterization make it suitable for describing a wide range of growth and decline phenomena.

3.9.3 METHOD SELECTION

We have described two graphical methods and three analytical methods for estimating Weibull parameters β and η . Now, the question is which method do we use? The answer depends on whether one needs a quick or an accurate estimation. In what follows, the methods are ranked according to their accuracy or speed. The order of the methods based on their speed (computing time) are

1. Any graphical method.
2. Least Squares Method.
3. Maximum Likelihood Estimator.
4. Method of Moments.

The order of the methods based on their accuracy are

1. Applying the three analytical methods (MLE,MOM and LSM) and selecting the best one which gives the minimum mean squared error.
2. Method of Moments.
3. Maximum Likelihood Estimator.
4. Least Squares Method.
5. Any graphical method.

3.10 FITTING OF SOME SPECIAL TYPES OF GROWTH CURVES

The various nonlinear growth curves such as the Modified exponential, Gompertz and Logistic curves cannot be fitted by the principle of least squares became the models for these curves involve the number of parameters more than the number of variables, these growth curves can be fitted to the time series data by using some special methods such as Method Of Three Selected Points, Method Of Partial Sums, Yule's Method, Hotelling's Method, Successive Approximation Method, Maximum likelihood Method and Linearization by Taylor Series expansion Method.

3.10.1 FITTING OF MODEIFIED EXPONENTIAL GROWTH CURVE

Consider the growth model for the modified Exponential Growth Curve as

$$Y_t = a + bc^t \tag{3.10.1}$$

Where Y_t is the value of the study variable at the time period t and *a, b, c* are unknown parameters

(A) Method Fof Three Selected Points

One may take three ordinates $Y_1, Y_2 \text{ and } Y_3$ to three equidistant values of $t = t_1, t_2 \text{ and } t_3$ respectively such that $t_2 - t_1 = t_3 - t_2$. (3.10.2)

Substituting values of $t_1, t_2 \text{ and } t_3$ in (3.10.1) one may get

$$Y_1 = a + bc^{t_1} \tag{3.10.3}$$

$$Y_2 = a + bc^{t_2} \tag{3.10.4}$$

$$Y_3 = a + bc^{t_3} \qquad (3.10.5)$$

$$Y_2 - Y_1 = bc^{t_1}(c^{t_2-t_1} - 1) \qquad (3.10.6)$$

And

$$Y_3 - Y_2 = bc^{t_2}(c^{t_3-t_2} - 1) \qquad (3.10.7)$$

$$\Rightarrow \frac{Y_3 - Y_2}{Y_2 - Y_1} = c^{t_2-t_1} \qquad (3.10.8)$$

$$c = \left[\frac{Y_3 - Y_2}{Y_2 - Y_1}\right]^{\frac{1}{(t_2-t_1)}} \qquad (3.10.9)$$

Substituting for *c* in (3.10.6) we get

$$Y_2 - Y_1 = b\left[\frac{Y_3 - Y_2}{Y_2 - Y_1}\right]^{\frac{t_1}{(t_2-t_1)}}\left[\frac{Y_3 - Y_2}{Y_2 - Y_1} - 1\right]$$

$$\Rightarrow b = \frac{(Y_2 - Y_1)^2}{Y_3 - 2Y_2 + Y_1}\left[\frac{Y_2 - Y_1}{Y_3 - Y_2}\right]^{\frac{t_1}{(t_2-t_1)}} \qquad (3.10.10)$$

Substituting *b* and *c* in (3.10.3)

$$a = Y_1 - bc^{t_1} = Y_1 - \left[\frac{(Y_2 - Y_1)^2}{Y_3 - 2Y_2 + Y_1}\right] = \frac{Y_1Y_3 - Y_2^2}{Y_3 - 2Y_2 + Y_1}$$

Substituting for *a*, *b* and *c* from (3.10.3), (3.10.4) and (3.10.5) get the equation of the Modified Exponential Curve fitted to the given time series data

Y_1, Y_2 and Y_3 being ordinates of the free hand curve corresponding to the three selected points $t = t_1, t_2$ and t_3

(B) Method Of Partial Sums

The given time series data are split up into three equal parts each containing, (say) *n* consecutive values of Y_t corresponding to $t=1,2,.....,n$; $t=n+1,n+2,.....,2n$; $t=2n+1,2n+2,.....,3n$. Let S_1, S_2 and S_3 represent the partial sums of the three parts respectively so that

$$S_1=\sum_{t=1}^{n} Y_t,\ S_2=\sum_{t=n+1}^{2n} Y_t \ and\ S_1=\sum_{t=2n+1}^{3n} Y_t \quad (3.10.11)$$

Substituting for Y_t (3.10.1), one may get

$$S_1=\sum_{t=1}^{n}\left(a+bc^t\right)=na+b(c+c^2+......+c^n)=na+bc\left(\frac{c^n-1}{c-1}\right) \quad (3.10.12)$$

Similarly we shall get

$$S_2=na+bc^{n+1}\left(\frac{c^n-1}{c-1}\right) \quad (3.10.13)$$

And

$$S_3=na+bc^{2n+1}\left(\frac{c^n-1}{c-1}\right) \quad (3.10.14)$$

Substituting (3.10.12) from (3.10.13), and (3.10.13) from (3.10.14) one may get

$$S_2-S_1=bc\frac{\left(c^n-1\right)^2}{c-1} \quad (3.10.15)$$

And

$$S_3-S_2=bc^{n+1}\frac{\left(c^n-1\right)^2}{c-1} \quad (3.10.16)$$

Dividing (3.10.16) by (3.10.15), we have

$$\frac{S_3-S_2}{S_2-S_1}=c^n \Rightarrow c=\left[\frac{S_3-S_2}{S_2-S_1}\right]^{1/n} \quad (3.10.17)$$

Substituting for c^n in (3.10.15), we get

$$S_2 - S_1 = \frac{bc}{c-1}\left[\frac{S_3 - S_2}{S_2 - S_1} - 1\right]^2$$

$$\Rightarrow b = \frac{(c-1)(S_2 - S_1)^3}{c(S_3 - 2S_2 + S_1)} \qquad (3.10.18)$$

Finally substituting the values of *b* and *c* in (3.10.12)

$$a = \frac{1}{n}\left[S_1 - \frac{bc}{c-1}\left(c^n - 1\right)\right]$$

$$= \frac{1}{n}\left[S_1 - \frac{(S_2 - S_1)^3}{(S_3 - 2S_2 + S_1)^2}\left(c^n - 1\right)\right] \quad \left[\text{From}(3.10.18)\right]$$

$$= \frac{1}{n}\left[S_1 - \frac{(S_2 - S_1)^3}{(S_3 - 2S_2 + S_1)^2}\left\{\frac{S_3 - S_2}{S_2 - S_1} - 1\right\}\right] \quad \left[\text{From}(3.10.17)\right]$$

$$= \frac{1}{n}\left[S_1 - \frac{(S_2 - S_1)^2}{S_3 - 2S_2 + S_1}\right]$$

$$= \frac{1}{n}\left[\frac{S_1 S_3 - S_2^{\ 2}}{S_3 - 2S_2 + S_1}\right] \qquad (3.10.19)$$

3.10.2 FITTING OF GOMPERTZ CURVE

Gompertz curve is given by the equation

$$y_t = ab^{c^t} \qquad (3.10.20)$$

where y_t is Time-Series value at time *t* and a, b, c are its parameters

$\log y_t = \log a + \log b.c^t$

$$Y_t = A + Bc^t \qquad (3.10.21)$$

where $Y_t = \log y_t, A = \log a \ and \ B = \log b$

Now (3.1.21) is a modified exponential curve and the constants *A*, *B* and *C* can be estimated by the method of three selected points or by the method of partial sums.

3.10.3 FITTING OF LOGISTIC GROWTH CURVE

Let us consider the Logistic growth curve

$$Y_t = \frac{k}{1+e^{a+bt}}; b<0, k=\max(Y_t) \tag{3.10.22}$$

Where t, Y_t are variables and *a*, *b*, and *k* are the unknown parameters. One may discuss below various methods of fitting of Logistic curve.

(A) Method Three Selected Points

Given time-series data is first plotted on a graph paper and a trend line is first drawn by freehand method. Three ordinates Y_1, Y_2, Y_3 taken from the trend line corresponding to selected equidistant points of time, say $t=t_1, t=t_2, t=t_3$ respectively such that $t_2-t_1=t_3-t_2$.

Substituting the values $t=t_1, t_2 \; and \; t_3$ in (3.10.22) we get

$$Y_1 = \frac{k}{1+e^{a+bt_1}} \Rightarrow a+bt_1 = \log\left(\frac{k}{Y_1}-1\right)$$

$$Y_2 = \frac{k}{1+e^{a+bt_2}} \Rightarrow a+bt_2 = \log\left(\frac{k}{Y_2}-1\right) \tag{3.10.23}$$

$$Y_3 = \frac{k}{1+e^{a+bt_3}} \Rightarrow a+bt_3 = \log\left(\frac{k}{Y_3}-1\right)$$

$$b(t_2-t_1) = \log\left(\frac{(k/Y_2)-1}{(k/Y_1)-1}\right)$$

$$b(t_3-t_2) = \log\left(\frac{(k/Y_3)-1}{(k/Y_2)-1}\right)$$

(3.10.24)

Since the points are equidistant i.e. $t_2-t_1=t_3-t_2$ we get

$$\log\left(\frac{(k/Y_2)-1}{(k/Y_1)-1}\right)=\log\left(\frac{(k/Y_3)-1}{(k/Y_2)-1}\right)$$

$$\Rightarrow\left(\frac{k}{Y_3}-1\right)\left(\frac{k}{Y_1}-1\right)=\left(\frac{k}{Y_2}-1\right)^2$$

$$\Rightarrow Y_2^{\,2}(k-Y_3)(k-Y_1)=Y_1Y_3(k-Y_2)^2$$

$$\Rightarrow Y_2^{\,2}\left[k^2-k(Y_1+Y_3)+Y_1Y_3\right]=Y_1Y_3(k^2+Y_2^{\,2}-2kY_2)$$

$$\Rightarrow k^2\left(Y_2^{\,2}-Y_1Y_3\right)=k\left[Y_2^{\,2}(Y_1+Y_3)-2Y_1Y_2Y_3\right]$$

Since $k\neq 0$, one may get

$$k=\frac{Y_2^{\,2}(Y_1+Y_3)-2Y_1Y_2Y}{Y_2^{\,2}-Y_1Y_3} \qquad (3.10.25)$$

From (3.10.24) and (3.10.23), one may get

$$b=\frac{1}{t_2-t_1}\log\left[\frac{(k-Y_2)Y_1}{(k-Y_1)Y_2}\right] \qquad (3.10.26)$$

$$a=\log\left(\frac{k-Y_1}{Y_1}\right)-bt_1 \qquad (3.10.27)$$

(B) Yule's Method

Suppose that the value of *k* is approximately known or obtained by other methods. Then the Logistic Growth Curve (3.10.22) contains two parameters *a* and *b*, and two variables t and Y_t.

Hence the principle of least squares can be used to estimate *a* and *b*

From (3.10.22)

$$a+bt=\log\left(\frac{k}{Y}-1\right) \qquad (3.10.28)$$

Or $$U=a+bt$$

Where $U=\log\left(\frac{k}{Y}-1\right)$ (3.10.28) represents a linear trend and according to the principle of least square, the normal equations for estimating *a* and *b* are

$$\sum U = na + b\sum t$$

$$\sum Ut = a\sum t + b\sum t^2$$

(C) Hotelling Method

A very elegant and ingenious method for fitting a Logistic Growth curve is given by Hotelling. One may have

$$Y = Y_t = \frac{k}{1+e^{a+bt}};\ b<0$$

The rate of growth is given by

$$\frac{dY}{dt} = \frac{-k}{\left(1+e^{a+bt}\right)^2}.b.e^{a+bt}$$

$$= -b\left(\frac{k}{1+e^{a+bt}}\right)\left(\frac{k}{1+e^{a+bt}}\right).e^{a+bt}$$

$$= -bY.\frac{Y}{k}\left(\frac{k}{U}-1\right)$$

$$= -bY.\left(1-\frac{Y}{k}\right)$$

$$\Rightarrow \frac{1}{Y}\frac{dY}{dt} = -b.\left(1-\frac{Y}{k}\right)$$

If the interval is not too large, then as an approximation to $\frac{1}{Y}\frac{dY}{dt}$ One may take $\frac{1}{Y}\frac{\Delta Y_t}{\Delta t}$ thus one may get

$$\frac{1}{Y_t}\frac{\Delta Y_t}{\Delta t} = -b + \frac{b}{k}Y_t \qquad (3.10.29)$$

Or

$$U = A + BY \qquad (3.10.30)$$

Where $U = \frac{1}{Y_t}\frac{\Delta Y_t}{\Delta t},\ A = -b\ \ and\ = \frac{b}{k}$

A and *B* and consequently *b* and *k* can be estimated from (3.10.30) by the principle of least squares

The constant *a* is then obtained by assuming that the curve $Y_t = \frac{k}{1+e^{a+bt}}; b<0, k=\max(Y_t)$ passes through the mean value of *Y* And mean of *t*.

(D) Method Of Successive Approximation

If some approximation values of the parameters *k, a, b* are known, then a first correction for each of these can be obtained by the principle of least squares. Let us suppose the first correction for *k, a* and *b* are $\mu, \lambda \text{ and } \delta$ respectively so that

$$Y_t = \frac{k+\mu}{1+\exp\left[a+\lambda+(b+\delta)t\right]}$$

$$= \frac{k+\mu}{1+\exp(a+bt)\exp(\lambda+\delta t)}$$

$$= \frac{k+\mu}{1+e^{(a+bt)}(1+\lambda+\delta t)}$$

Higher powers of $\lambda \text{ and } \delta$ being ignored since $\lambda \text{ and } \delta$ are very small

$$Y_t = \frac{k+\mu}{\left(1+e^{a+bt}\right)}\left[1+\frac{\left(\lambda+\delta t\right)e^{a+bt}}{1+e^{a+bt}}\right]^{-1}$$

$$= \frac{k+\mu}{\left(1+e^{a+bt}\right)}\left[1-\frac{\left(\lambda+\delta t\right)e^{a+bt}}{1+e^{a+bt}}\right],$$

Higher powers being neglected.

$$\therefore Y_t = \frac{k}{1+e^{a+bt}} + \frac{\mu}{1+e^{a+bt}} - \frac{k\lambda e^{a+bt}}{\left(1+e^{a+bt}\right)^2} - \frac{kt\delta e^{a+bt}}{\left(1+e^{a+bt}\right)^2} \qquad (3.10.31)$$

Terms involving $\lambda\delta$ and $\mu\delta$. being ignored (3.10.31) may be rewritten as, say,

$$Y_t = A_t + \mu B_t + \lambda C_t + \delta D_t$$

Where A_t, B_t, C_t and D_t are known. Since *k, a* and *b* are known. μ, λ and δ can be obtained by the principle of least squares.

CHAPTER - 4 POPULATION GROWTH MODELS

4.1 POPULATION GROWTH RATE

Population growth is the change in a population over time and can be quantified as the change in the no. of individuals of any species in a population using "per unit time" of for measurement.

(A) Determination of Population Growth

Population growth is determined by four factors Births (B), Deaths (D), immigration (I) and emigrants (E) using a formula expressed as

$$\Delta P = B - D + I - E$$

In the other words, the population growth of a period can be calculated in two parts, natural growth of population (B-D) and mechanical growth of population (I- E) in which Mechanical growth of population is monthly affected by a social factors e.g. the advanced Economics are growing faster while the backward economies are growing slowly even with negative growth.

(B) Population Growth Rate

In demographic and ecology, population growth rate (PGR) is the fractional rate at which the no. of individuals in a population increases. Especially PGR ordinarily refers to the change in population over a unit time period often expressed as a percentage of the no. of individuals in the population at the beginning of that period. This can be written as the formula

$$Growth\ rate = \frac{Population\ at\ end\ of\ period - Population\ at\ beginning\ of\ period}{Population\ at\ beginning\ of\ period}$$

(In the limit of a sufficiently small time period)

The above formula can be expanded to

Growth rate = Crude birth rate – Crude death rate + net immigration rate.

Or

$$\frac{\Delta P}{P} = \left(\frac{B}{P}\right) - \left(\frac{D}{P}\right) + \left(\frac{I}{P}\right) - \left(\frac{E}{P}\right)$$

Where P is the total population,

B is the no. of Births

D is the no. of Deaths

I is the no. of immigrants

E is the no. of Emigrants

This formula allows for the identification of the source of population growth whether due to natural increase or an increase in the net immigration rate. Natural increase is an increase in the native born population. Stemming from either a higher birth rate or , a lower Death rate or a combination of the two. Net immigration rate is the difference between the no. of immigrants and the no. of Emigrants.

The most common way to express population growth is as a ratio, not a rate. The change in population over a unit time period is expressed as a percentage of the population at the beginning of the time period. That is

$$\textit{Growth ratio} = \textit{growth rate} \times 100\%$$

A positive growth rate indicates that the population is increasing, while a negative growth ratio indicates the population decreasing. A growth ratio of zero indicates that there were the same number of people at the two times is net difference between births, deaths and migration is zero.

(C) Excessive Growth and Decline

Population exceeding the carrying capacity of an area or environment is called overpopulation. It may be caused by growth in population or by reduction in capacity. Spikes in human population can cause problems such as pollution traffic congestion, these might be resolved or worsened by technological and economic changes.

Conversely, such areas may be considered “under populated” if the population is not large enough to maintain an economic system.

4.1.1 INFERENCES FROM A DETERMINISTIC POPULATION DYNAMICS MODEL FOR BOWHEAD WHALES

Definition of the Model

The population Dynamics model (PDM) that we consider was developed for Bowhead whales by Brejwick, Eberhandt and Braham (1984) and is a special case of the well known one sex age structured Leslie matrix population projection model (Lewis 1942, Leslie 1945, 1948). A fairly general form of this model is as follows. Let n_{xt} be the no. of females aged x next birthday on January 1 of calendar year *t*, where *t =0* is the initial year, then the model is specified by the equations.

$$n_{t,t+1} = \sum_{x=1}^{\infty} f_{xt}(n_{xt} - c_{xt}) \qquad (4.1.1)$$

$$n_{x+1,t+1} = S_{xt}(n_{xt} - c_{xt}), \;\; x = 1, 2, ... \qquad (4.1.2)$$

where f_{xt} is the average number of female calves that survive to age 1 born in year *t* to a female aged x, S_{xt} is the natural survival rate of females aged *x* in year *t* and c_{xt} is the no. of females aged 'x' killed by hunting in year 't'.

This can be written in the matrix form as

$$N_{t+1} = A_t(N_t - C_t) \qquad (4.1.3)$$

Where $N_t = (n_{1t}, n_{2t}, ...)^T$, $C_t = (c_{1t}, c_{2t}, ...)^T$ and $A_t = (A_{txy} : x, y = 1, 2, ...)$ is a doubly infinite square matrix defined by

$$A_{txy} = \begin{cases} f_{yt} & if\, x = 1 \\ s_{yt} & if\; x = y+1 \\ 0 & otherwise \end{cases} \qquad (4.1.4)$$

As it stands the model has an infinite number of parameters, and Brejwick et.al.(1984) proposed the following restrictions for the bowhead whale case.

Mortality: It is assumed that an immature survival rate, s_0 prevails from age 1 to age' a' and that a mature (adult) survival rate applies from age *a+1* onwards. To approximate senescence, it is assumed that all individuals aged *w* at time *t* die before time *t+1*.

Mortality is assumed to be constant over time and in particular to involve no density dependence so that S_{xt} does not depend on t. Thus we have

$$S_{xt} = \begin{cases} s_0 & if\, x = 1, \ldots.., a \\ s & if\ x = a+1, \ldots..., w-1 \\ 0 & x = w \end{cases} \tag{4.1.5}$$

Calf mortality is accounted for in the specification of fertility.

Fertility : Fertility is assumed to be constant with respect to age between at sexual mortality, m and age *w-1* . It is assumed to be density dependent with a functional form corresponding to a modified logistic growth curve. First parturition is assumed to occur one year after age at sexual maturity this yields

$$f_{xt} = \begin{cases} 0 & x = w \\ f_t = f_0 + \left(f_{\max} - f_0 \right) \left[1 - \left(\dfrac{p_t}{p_0} \right)^2 \right] & x = m+1, \ldots\ldots.., w-1 \end{cases} \tag{4.1.6}$$

where P_t is the female population size at the beginning of year *t*, $f_{\max}$ is the maximum fertility, attained when the stock is near extinction, and z is the density dependence Parameter. Assuming that the population was in equilibrium at $t = 0$, before the start of Commercial whaling, yields a value for f_0 by solving the matrix equation $A_0 N_0 = N_0$, namely

$$f_0 = (1-S) / \left[\left(S_0 / S \right)^{\alpha} \left(S^m - S^{w-1} \right) \right] \tag{4.1.7}$$

one estimate of f_t note that f_{xt} is the product of the no. of female calves per mature female with the calf survival rate. We have no information which would enable us to separate fertility from calf survival. By including calving rate and Calf survival in a single term we acknowledge that density dependence may occur through changes in reproductive rate changes in first year survival or both.

Hunting Mortality: The length of whales killed in the past three decades suggests that the recent subsistence harvest has predominately selected immature whales. Historically, the Commercial catch was probably biased towards larger, mature animals.

With these assumptions, the original model of Brejwick et al(1984) first divides the catch into two shares : for immature whales and one for mature whales. It then distributes each share among its corresponding age classes in proportion to the relative abundance of each class at the beginning of the year. With this distribution of the annual catch, it is possible to obtain negative age class counts without population extinction. We have modified the original model so that the number removed from each age. Class is never more than the current class size and what remains of the mature and immature catch shares is distributed proportionally as before but among empty age classes.

The model requires values of the eight input parameters $s_0, s, a, w, m, f_{\max}, z \; and \; p_0$ as well as hunting mortality by year. Given these, it outputs a full age distribution of the female population for each year. It is assumed that the sex ratio is 1:1, so doubling this gives the total population.

4.1.2 YIELD QUANTITIES AND THEIR RELATIONSHIP TO THE MODEL

Several quantities used by the IWC (International Whaling Commission) for making policy decisions are related to the inputs and outputs of this PDM (Allen 1976, Cooke 1987, Butterworth and Best 1990). One is maximum sustainable yield (MSY). Once an unexploited stock of size P_0 begins to be exploited, it can sustain indefinitely any level of catch less than MSY. The MSY level (MSYL) is the lowest population level at which MSY is attained, expressed as a proportion of P_0. Under the assumption of density dependence in reproductive rate and/ or Calf survival the population increases at a higher rate when it has been reduced below P_0 than when it is at or near its carrying capacity and therefore prevented by environmental limitations from increasing. Thus MSYL is less than one; it has often been assumed to 0.6 by the SC (Scientific Committee).

For protected species like the bowhead; replacement Yield (RY) is a key management concept. RY is the catch from the recruited stock which, if taken, would leave the recruited population at the same level at the beginning of the next season (IWC, 1988). For Bowheads we assume that the recruited stock consists of all whales aged at least one year. The maximum sustainable yield rate (MSYR) is defined as RY at MYSL, expressed as a proportion of the population at MSYL. We define MSYL and MSYR in terms of the total population aged 1 or above as proposed by Butterworth and Punt (1992).

Relationship between the model parameters and MSYL and MSYR are included by the Characteristic equation of the Leslie matrix and by the density dependence equation in (4.1.6).

The Characteristic equation of the Leslie matrix, A_t is

$$\lambda^{m+1} - s\lambda^{m} - s_0^{\alpha} s^{m-\alpha}\left[1-\left(s/\lambda\right)^{w-m-1}\right]=0 \qquad (4.1.8)$$

where λ is the Eigen value or population multiplier So that $N_{t+1} = \lambda N_t$ t by (4.1.3) (Brejwick et al 1984).

If the time *t* is such that p_t / p_0 = MSYL and $s_0, s, a, w, m, and\ MSYR$ are known, Then can find the fertility rate at MSYL, f_{MYSL} by selecting λ = 1+MSYR and solving for f_t in equation (4.1.8). Similarly, f_0 the relatively rate of a Stable Unexploited stock can be found by setting λ = 1 in equation (4.1.8). The only remaining unknown quantities in equation (4.1.6) an $f_{\max}$ and Z as given $f_{\max}$, we can solve equation (4.1.6) to the corresponding value of Z. Note that, by definition $f_{\max} \geq \max(f_0, f_{MSYL})$

4.2 POPULATION AND ECONOMIC GROWTH

Analysis of economic growth may treat population growth as an independent variable, as endogenous variable, or as an instrument to be altered according to economic and social criteria. All these approaches except one will be reviewed. The subject of population and optimal economic growth more appropriately precedes the discussion of optimal population control.

The descriptive neoclassical growth models is so well known , that ist need be dealt with only briefly. Many of its properties are dependent on the fact that the proportional rate of growth of population is assumed constant. As was seen in provide the evidence for there being some simple set of economic determinants of population growth does not seem strong. fertility in particular would appear to be only weakly associated with economic variable, so far as mortality is concerned , whilst Adelman's study did not show a relationship with income per capita, it is also true that for underdeveloped economics there is evidence that relations inexpensive public health programs are a much more significant influence . Nevertheless, the idea tht the minimum requirements of subsistence must limit population endogenous in a variety of ways and the result models some items lead to surprising conclusions.

The models presented are described in some detail, and a brief appraisal of their results is given.

4.2.1 THE DESCRIPTIVE NEOCLASSICAL GROWTH MODEL

Suppose population grows at the constant percentage rate, *n*, if it has been growing at this rate for sometime the ration between those of workforce age and dependents may be used interchangeably without confusion. Output is produced by two inputs capital (*K*) and Labor (*N*) under conditions of constant returns to scale, so that output per head may be related to capital per head (*K*) by

$$y = f(k),\ f \in c^2$$

$$f^1 > 0,\ for\ all\ k \geq 0 \qquad (4.2.1)$$

Further assumption (which ensures unique results) is that f Is strictly concave.

Assume that capital depreciates exponentially at the rate λ. if all saving is invested and if saving is a constant proportion of income(S) . it follows that

$$\frac{dk}{dt} = I - \lambda k = sf(k)N - \lambda k \qquad (4.2.2)$$

Where λ is the rate of depreciation of capital? The proportional rate of change in the capital-labour ratio is, by definition

$$\frac{dk}{dt}.\frac{1}{k}=\frac{dk}{dt}.\frac{1}{k}-\frac{dk}{dt}.\frac{1}{N} \tag{4.2.3}$$

So that using (4.2.2) and the assumed constancy of the population growth rate(n)

$$\frac{dk}{dt}=sf(k)-(n+\lambda)k \tag{4.2.4}$$

It is not difficult to make father assumptions to ensure the existence of a positive capital labour ratio(k) to which the system would converge it may be deduced that in long-run equilibrium, capital and output both grow at the same rate as labour (n) but that raising the saving ration raises the equilibrium level of output per head.

Technical progress of Harrods-neutral type may easily be added to the model. If it occurs at a constant proportional rate, output per head will rise “forever” at a constant rare. In essence, the efficiency of the labour force continually improves. Other assumptions about technical progress considerably alter the conclusions, except for special cases.

Leaving technical progress aside, what part does population growth play in this mode? Being an exogenous variable it is likely to be of considerable importance. Indeed the equilibrium rate of growth of capital and of output are the same as the population growth rate. There is nothing to prevent population growing forever at the given rate. Of course, the assumption of constant returns to scale to capital and labour is essential reason for this.

If a choice is considered amongst various constants but non-negative population growth rate if is easily seen that the highest output per head is achieved with the stationary population ($n=0$) even higher levels of output per head could be achieved by Makin the growth rate negative, but the implications of the continuity assumption when applied to the population variable are not easy to accept. Population could decline forever at a constant proportional rate only if people (or their labour) were perfectly divisible the fact that scale does not matter in the system leads to the result that whatever the level of population when that variable is stationary the equilibrium level of output per head would be the same. The absolute sixe of population plays no essential part in determining the results of this model.

The originators of this approach were, of course, aware of this sort of implication of their models, and it is true that for some purpose it needed not be a serious limitation. In the sprint of the 1950's and 1960's restrictions applicable to a continued growth of population may have seemed relevant only to the distinct future. Further, by assuming that population growth was independent of economic considerations it may be that they were recognizing the greater independence with respect to child-bearing and the advances in medical care which have become available. Yet there were under developed counties where there seemed to be conflict between population growth and economic welfare, and casual observation of the real world would suggest that the time must come when the same applied to developed countries. For these situations the neoclassical growth model was not relevant.

4.2.2 ENDOGENOUS POPULATION GROWTH

A model which is close to tha5 Ricardo has been studied by Niehans(1963). Using a cob-douglass production function,

$$Y = K^{\alpha} N^{\beta}; \quad \alpha, \beta \in (0,1),\ \alpha + \beta < 1; \qquad (4.2.5)$$

he assumes that noth laour and capital growth rate relate to their respective marginal products so that

$$\frac{dN}{dt}\frac{1}{N} = p\left(\frac{\partial Y}{\partial N} - w_m\right); \ w_m, p > 0, \ \text{constant}; \qquad (4.2.6)$$

$$\frac{dK}{dt}\frac{1}{K} = s\left(\frac{\partial Y}{\partial K} - r_m\right); \ r_m, p > 0, \ \text{constant}; \qquad (4.2.7)$$

w_m and r_m are the marginal products at which growth of the respective factor becomes zero. this Niehans calls" two-class model" although really it must involve a third class holding rights to a fixed factor when it assumed that there are decreasing returns to scale from (4.2.5)

$$\frac{\partial Y}{\partial N} = \beta \frac{Y}{N} \qquad (4.2.8)$$

$$\frac{\partial Y}{\partial K} = \alpha \frac{Y}{K} \qquad (4.2.9)$$

To analyze this model it is useful to plot the curves for zer population and capital growth given by equating each of (4.2.6) and (4.2.7) to zero.

Population and capital follow the arrows shown in the diagram towards the stationary state $\left(\bar{N}, \bar{K}\right)$ at which point the return to labour and capital will be at minimum levels.

The constant and increasing returns to scale case do not necessarily lead to such a situation as sustained growth may be possible.

$$\frac{\dot{K}}{K} = s.\frac{Y}{K}. \qquad (4.2.10)$$

Further from (4.2.1)

$$\frac{\dot{Y}}{Y} = \alpha.\frac{\dot{K}}{K} + \beta\frac{\dot{N}}{N} = \alpha s\frac{Y}{K} + \beta\frac{\dot{N}}{N}. \qquad (4.2.11)$$

Now swan assumes an extreme form of the Malthusian hypothesis namely that population grows at such a rate that output per head remains constant this means that

$$\frac{\dot{N}}{N} = \frac{\dot{Y}}{Y} \qquad (4.2.12)$$

Combining (4.2.11) and (4.2.12)

$$\frac{\dot{Y}}{Y} = \frac{\dot{N}}{N} = \left(\frac{\alpha}{1-\beta}\right) s\frac{Y}{K} \qquad (4.2.13)$$

He then plots (4.2.10), (4.2.12) and (4.2.13) against the output-capital ratio to obtain the construction in (4.2.3).

Capital gains faster than labour for any positive *Y/K*. So that the growth rates of population and Output fall along the 'Recondian Line'. It does not however, follow that the system approaches a stationary state. To see this note that a feature of the system to the output per head is constant or some level. Say $\bar{y}$. From (4.2.5) it follows that

$$K = \bar{y}^{\frac{1}{\alpha}} N^{\frac{1-\beta}{\alpha}} \qquad (4.2.10^1)$$

Which, because $\alpha + \beta < 1$

Now observe from (4.2.10) that capital continues to gram unless K and N zero. As both are initially positive and neither declines capital continues to grow down the Recondian Line or up the iso-productivity curve and can therefore exceed any bound. Population also grows along the parabola and also would exceed and bound. There is no stationary state for the system as it stands.

However, the model may be easily adapted so that a stationary state must exist. It capital is supposed to determine exponentially (4.2.10) becomes.

$$\frac{\dot{K}}{K} = s\frac{Y}{K} - \lambda, \qquad (4.2.14)$$

And it is clear that rare so growth of capital is zero if

$$\frac{Y}{K} = \frac{\lambda}{s}, \qquad (4.2.15)$$

Or using 4.3.5

$$K = \left(\frac{\lambda}{s}\right)^{\alpha-1} N^{\frac{\beta}{1-\alpha}} \qquad (4.2.16)$$

The output capital ratio converges to $\frac{\lambda}{s}$. As N and K increases along $\bar{y}$ locus the output-capital ratio is falling and the system approaches the stationary state characterized by $\left(\bar{N}, \bar{K}\right)$

Without exponential depreciation it was possible for the capita-labour ratio to grow without limit at a rate sufficient to offset the adverse effects of increasing scale the rise in the capital labour ratio means a rise in the capital-output ratio, so that a situation can be approached in which capital in sufficiently large in comparison with output that the saving out of output (sY) is just sufficient to offset the depreciation capital (λK).

Both Enke[1963] and Wiehans [1963] have produced models which are similar to that Swan by contrast with his two-class model, Niehans one-class model does not relate a factor's growth to its earning. Istead both capital and Labour growth depend on output per head. Whist the seems reasonable for labout in classless economy; I cannot see why classless citizens should accumulate. Capital in way independent of its rate of return, unless output head is very low.

By contrast with Swan's model Niehans and Enke allow for the negative investment, and hence it would seem for depreciation of capital. Pespite this assumption Niehans comes to conclusion that theme one situation in which 'Capital and population will never cease to grow, domination returns not withstanding (1963 p.362).

As in his earlier model Niehans again uses a Cobb-Douglas production function with diminishing returns to scale to capital and labour. The behavior of population and capital respectively are assumed to be given by

$$\dot{N} = p\left(\frac{Y}{N} - m_L\right)N; \;\; p, m_L \text{ constant,} \qquad (4.2.17)$$

$$\dot{K} = s\left(\frac{Y}{N} - m_K\right)N; \;\; s, m_K \text{ constant,} \qquad (4.2.18)$$

Niehans Calls (4.2.10) nothing but an old fashioned Keynesian savings function [1963 P.358], but it is clearly more than that is has implications for capital depreciation as well as for saving. When output per head is below some level m_K, Saving and new investment is insufficient to offset depreciation, and the capital stock declines without enamoring what rationable them may be for this function, its would be noted that if $m_K < m_L$ depreciation does not set a limit to the scale of the economy. The point which is relevant for the present discussion is that whilst exponential depreciation must result in bounds to growth these sorts of economies, it is possible to design and perhaps to justify depreciation assumption which does not have this property.

Thus leaving depreciation aside if diminishing returns no scale is no enough to embody the essentials a bounded environment; what alternative approaches will continue this idea? The answer would seem to include the shape of the iso-output/head curves. It was seen that the swan model grows without bounds because both K and N approach infinity as $t \to \infty$ In general, the slope of an iso-productivity curve is given by

$$\frac{dK}{dN}[(Y/N) = const.] = \frac{-(F_N - (F/N))}{(F_K)} \tag{4.2.19}$$

Where F is a twice differentiable production function with arguments N and K , and F_N and F_K are the marginal product of labour and capital, respectively.

If the slope becomes infinite for some finite $\left(\bar{N}, \bar{K}\right)$, that is,

$$F_K\left(\bar{N}, \bar{K}\right) = 0, with F_N\left(\bar{N}, \bar{K}\right) \neq F\left(\bar{N}, \bar{K}\right)/\bar{N} \;\; for\, some\left(\bar{N}, \bar{K}\right), \tag{4.2.20}$$

the size of the population is bounded above condition (4.2.10) Simply means that output can't be further increased by applying more capital, so that additional labour must reduce output per head surely it is reasonable to assume that with given technology in a finite environment there must be some scale of operation and input combinations at which further capital is no longer productive. If this were not so, the ridiculous extreme could be reached at which the mass of all capital equipment exceeded that of the universe, but an extra unit of capital could still produce further output indeed, taking exrenalities into negative account a negative managerial production of capital is not unreasonable

Using Swan's population assumption the economy, moves upward along the iso-population curves with population bounded above by $\bar{N}$. Capital is not bounded in this process, but if investment behavior were made to depend in some way on the managerial production of capital it is reasonable to assume that accumulation would stop at or before the point at which the marginal product of capital was zero.

Suppose a production function linear and homogenous in the three factors labour capital and land, with a fixed of land. It is required to investigate the following problem. It is possible for population and the capital stock to grow forever in such a regime without the marginal productivity of capital becoming zero the growth rate of the output per head (*y*) will be given by

$$\dot{y}/y = \theta_K(\dot{K}/K) - (\theta_L + \theta_K)(\dot{N}/N) \tag{4.2.21}$$

Where θ_K and θ_L Are the production classified of capital and labour respectively, and will be assumed non-negative. From 21 If output per head is constant.

$$\dot{K}/K = \left(1 + \frac{\theta_L}{\theta_K}\right)\frac{\dot{N}}{N} > \frac{\dot{N}}{N}, \tag{4.2.22}$$

So that if population grows with a constant *y*, both capital and capital-labour ratio(*k*) must also rise more over the land labour ratio (*l*) Must fall, what then happens to the marginal product of capital (f_k) writing the production function in the form.

$$Y = Nf(k, l) \tag{4.2.23}$$

It follows that

$$\frac{df_k}{dk} = f_{kk} + f_{kl}\frac{dl}{dk} \tag{4.2.24}$$

It has been deduced that $\frac{dl}{dk} < 0$, and if diminishing returns to a single factor is assumed, there $f_{kk} < 0$ However, no assumption has yet been made on the sign of f_{kl} so that it is possible that the marginal product of capital could remain constant or even rise. If it were assumed that f_{kl} was positive this would e sufficient to ensure that the marginal product of capital fell during the process.

A reasonable model which has apparently not been considerable would seem to be one which related capital growth to its return and population growth to income per head. Thus, using earlier notation,

$$\dot{K} = I - \lambda K$$

$$I = \begin{cases} \mu(F_k - r)K, & F_k \ge r;\ \mu > 0, const. \\ 0, & F_k \le r, \end{cases} \tag{4.2.25}$$

$$\dot{N} = \xi((Y/N) - m_L)N, \quad \xi > 0, const. \tag{4.2.26}$$

It is not difficult, using a decreasing returns to scale Cobb-Douglas production function, to show that this model approaches a unique stationary state from any initial position with *N* and *K* Both positive.

Output is assumed to be produced under conditions of constant returns to scale to capital and labour. Capital accumulation per head of population is explained by

$$\frac{dK}{N} = \begin{cases} b(y - x); & y > y^1\ , b > 0 \\ -C\ ; & y \le y^1. \end{cases} \tag{4.2.27}$$

Above $(Y/N)^1 = y^1$ saving per head is increasing function of income per head.

The percentage rate of growth of population is also related to output per head. up to y^1 the death rate is assumed to be a decreasing function of output per head whilst above that level the rate of growth of population is assumed constant.

The working of the system can be illustrated readily if a Cobb-Douglas production function is assumed. It may be shown that for such a function

$$\frac{Y}{K} = \left(\frac{Y}{K}\right)^{\frac{\alpha-1}{\alpha}} \tag{4.2.28}$$

and

$$\frac{N}{K} = \left(\frac{Y}{N}\right)^{\frac{-1}{\alpha}} \tag{4.2.29}$$

Using these results the rate of growth of capital as a function of output per head is given by

$$\frac{dK}{K}=\begin{cases} b\left(\frac{Y}{N}\right)^{\frac{-1}{\alpha}}\left\{\left(\frac{Y}{N}\right)^{\alpha-1}-X\right\}; & y>y^1; \\ -C\left(\frac{Y}{N}\right)^{\frac{-1}{\alpha}}; & y\le y^1 \end{cases} \tag{4.2.30}$$

substituting in 4.2.11 the rate of growth of income for $y > y^1$ is

$$\frac{dY}{Y}=\alpha b\left(\frac{Y}{N}\right)^{\frac{-1}{\alpha}}\left[\left(\frac{Y}{N}\right)^{\alpha-1}-X\right]+(1-\alpha)\frac{dN}{N} \tag{4.2.31}$$

and for $y \le y^1$

$$\frac{dY}{Y}=-\alpha C\left(\frac{Y}{N}\right)^{\frac{-1}{\alpha}}+(1-\alpha)\frac{dN}{N} \tag{4.2.32}$$

Thus as output per head rises from a low level the rate of growth of output rises, but approaches $(1-\alpha)\left(\frac{dN}{N}\right)^{*}$ as $y\to\infty$.

4.3 POPULATION GROWTH AND ASSET PRICES

There are many authors examine the relationship between a population’s age distribution and asset prices. Mankiw and Weil [1989] argue that the maturing of the baby boomers during the 1970’s accelerated the rate of household formation, which in turn, increased the demand for housing and its price. Bakshiand Chen [1994] incorporate Mankiw and Weil’s findings into a “life-cycle investment hypothesis” which argues that individuals change their allocations of wealth as they age so an aging population alters aggregate demand for assets and thus their prices. Yoo [1994], motivated by a similar intuition, finds that the real return to U.S. T-bills is negatively correlated with the size of the age group that has the highest increment to its wealth.

Some of the authors show that an individual's demand for an asset varies with age. They then argue that because aggregate demand is merely the sum of individual demands, changes in a population's age distribution affects the aggregate demand for that asset and thus affect the price of the asset.

Although all of them share a common intuition, they are mostly empirical. This section examines the theoretical underpinnings of the relationship between a population's age distribution and asset prices. It presents a general equilibrium model that aggregates individual's optimizing behavior to derive equilibrium asset prices. The theoretical relationships between the two variables are shown by simulations based on the model.

The model suggests four conclusions about the relationship between asset prices and a population's age distribution. First, changes in a population's age distribution affect asset prices, as noted by the empirical literature. Second, although fluctuations in the population growth rate like the US post-war baby boom affect asset prices, such changes generate low frequency movements in asset prices. Third, the treatment of expectations matter; a small response of individuals to changes in asset prices has large implications for the path of asset prices. Finally, incorporating a supply of assets by interpreting an asset as a claim on physical capital diminishes the magnitude of the relationship but does not change the sign or timing of the relationship between a population's age distribution and asset prices.

4.3.1 A MODEL OF AGE DISTRIBUTION AND ASSET PRICES

The intuition that explains the potential relationship between a population's age distribution and asset prices is rather simple. First, it assumes that an individual's demand for an asset varies with age. This premise underlies all three empirical papers cited and it is consistent with life-cycle behavior if an individual saves in assets, rather than saves storable consumption goods. Given the assumption, variations in a population's age distribution will alter the aggregate demand for that asset by changing the distribution of asset holders. So, a young population has many savers which generate a high total demand for assets, but an old population has many dissevers so total demand for assets is low. It is this variation in aggregate demand for an asset that produces a relationship between a population's age distribution and asset prices.

A Simple Model

Let D_t be the aggregate quantity of consumption goods invested in an asset in period t. Also, let S_t be the number of shares of that asset outstanding in the economy. Then in equilibrium, the price of the asset P_t adjusts so that,

$$D_t = P_t S_t \tag{4.3.1}$$

If an agent lives for T_d periods and saves $a_{t,s}$ when *s* years old in period *t*,and $N_{t,s}$ is the number of individuals s years old, then the aggregate demand for the asset is

$$D_t = \sum_{s=1}^{T_d} N_{t,s} a_{t,s} \tag{4.3.2}$$

Combining the two equations shows how changes in a population's age distribution affect asset prices in a manner suggested by the empirical literature,

$$P_t = \frac{\sum_{s=1}^{T_d} N_{t,s} a_{t,s}}{S_t} \tag{4.3.3}$$

An additional individual aged *s* affects asset prices by an amount proportional to $a_{t,s}$. So variations in a population's age distribution affect asset prices as long as age is a related to an individual's demand for the asset.

A pricing equation very similar to equation (4.3.3) can also be derived by combining an individual's budget constraint and an aggregate resource constraint for an endowment economy. The individual's budget constraint is,

$$a_{t,s} = \frac{P_t}{P_{t-1}} a_{t-1,s-1} + e_s - c_{t,s} \tag{4.3.4}$$

Where e_s the endowment of consumption is good received by an individual aged *s* and

$c_{t,s}$ is the consumption of that agent in period *t*. If the endowment goods are non-storable, then in equilibrium, total consumption equals total endowment each period,

$$\sum_{s=1}^{T_d} N_{t,s}c_{t,s} = \sum_{s=1}^{T_d} N_{t,s}e_s \,. \tag{4.3.5}$$

Substituting (4.3.4) into (4.3.5) yields

$$P_t = \frac{\sum_{s=1}^{T_d} N_{t,s}a_{t,s}}{\sum_{s=1}^{T_d} N_{t-1,s}a_{t-1,s}} P_{t-1} \,. \tag{4.3.6}$$

The two pricing equations (4.3.3) and (4.3.6) are equal if

$$S_t = \frac{\sum_{s=1}^{T_d} N_{t-1,s}a_{t-1,s}}{P_{t-1}} \,.$$

Although the derivations of the asset pricing equations are fairly simple, both equations capture the intuition and the methodology of the empirical literature. The empirical findings start with a cross- sectional estimate of an individual’s demand for an asset and aggregate using the population’s age distribution to determine how demographic factors affect asset prices.

4.3.2 MODIFYING THE MODEL’S ASSUMPTIONS

The basic model presented in the previous section and make two strong assumptions when using equation (4.3.3) to show the relationship between the age distribution and asset prices. This section examines the impact of relaxing the assumptions about expectations and supply of assets.

Expectations

Estimating and simulating (4.3.6) in two steps assumes that an individual’s demand for an asset does not respond to its price, a strong assumption given that saving is implicitly a forward-looking behavior. This section presents a simulation that replaces the static expectations assumption with a perfect foresight one. Equation (4.3.6) still captures the relationship between asset prices and age distributions but future asset prices now affect $a_{t,s}$.

The simulation strategy is similar to that of the previous section. The population growth rates generate the $N_{t,s}$'s and supplies the $a_{t,s}$'s for the asset pricing equation but unlike the previous simulation, individuals now consider the future path of asset prices when making their saving decisions. Incorporating forward-looking behavior generates a perfect foresight path for the price of an asset.

Computationally, I use an iterative technique to simulate a perfect foresight path for asset prices. Agents in each iteration use the path of prices calculated by the previous iteration as the future path of asset prices. The iterations continue until each period's price changes by a negligible amount. The system is closed by assuming that asset prices return to their steady state growth rate within a finite horizon. Auerbach and Kotlikoff use this technique to simulate their model.

Changing the assumption about individual's expectations has a noticeable effect on the response of asset prices to a baby boom. The model under either assumption about expectations shows that demographic shocks like a baby boom affects asset prices albeit with different timing and magnitudes.

Assumptions about the responsiveness of $a_{t,s}$ to changes in asset prices are responsible for the differences between the responses of asset prices to a baby boom.

Supply of Assets

The model specified above shows how the demand for an asset changes with demographic changes but it assumes that the supply of the asset is fixed. This section addresses that problem by incorporating a production function into the economy so that assets now represent claims to physical capital.

An individual still maximizes her utility function but wages replace the age-dependent endowments in her budget constraint. So,

$$a_{t,s} = \frac{P_t}{P_{t-1}} a_{t-1,s-1} + e_s w_t - c_{t,s}, \tag{4.3.7}$$

Where e_s now represents an age-dependent labor productivity parameter and w_t is the wage paid per effective unit of labor.

A Cobb-Douglas production function represents the productive capacity of the economy

$$Y_t = K_t^{\alpha} L_t^{1-\alpha} \tag{4.3.8}$$

Where Y_t is the net output of the economy,

$$K_t = \sum_{s=1}^{55} N_{t-1,s} a_{t-1,s} , \tag{4.3.9}$$

And

$$L_t = \sum_{s=1}^{45} N_{t,s} e_s \tag{4.3.10}$$

The factors of production receive their marginal product, so that

$$\frac{P_t}{P_{t-1}} = \alpha k_t^{\alpha-1} \tag{4.3.11}$$

And

$$w_t = (1-\alpha) k_t^{\alpha} , \tag{4.3.1}$$

where k_t is the capital-labor ratio in t

The above equations replace (4.3.5) and (4.3.6) in determining the equilibrium asset prices of the economy. Otherwise the simulation strategy is similar to that of the endowment economy. As before, the simulation uses an iterative technique to determine the perfect foresight path of asset prices.

4.4 DENSITY-INDEPENDENT POPULATION GROWTH

Density-independent growth models offer an extremely simple perspective on changes in population size by assuming away many potential complications. For example, two sets of counteracting processes affect population size; birth and immigration increase populations while death and emigration decrease them. To simplify, assume that (a) immigration and emigration balance, leaving birth and death as the only determinants of population density. Let's also assume that (b) all individuals are identical (especially with respect to their probabilities of dying or producing offspring), (c) the population consists entirely of parthenogenetic females, so that we can ignore complications associated with mating, and (d) environmental resources are infinite, so that the only factors affecting population size are the organisms' intrinsic birth and death rates. These assumptions allow a simplistic model of population growth, and it is instructive to present the model in two formats for different kinds of life histories.

4.4.1 EXPONENTIAL GROWTH WITH CONTINUOUS BREEDING

First we will consider an organism like Homo sapiens or the bacteria in a culture flask, with continuous breeding and overlapping generations. All ages will be present simultaneously, and population size will change steadily in small increments with the birth and death of individuals at any time. This continuous population growth is best described by a differential equation, with instantaneous rates defined over infinitely small time intervals.

If: N = population size

b = instantaneous birth rate per female

d = instantaneous death rate per female

then population growth is given as:

$$\frac{dN}{dt} = (b-d)N \qquad (4.4.1)$$

If we collect the per capita birth and death rates in a single parameter $r = b - d$ called the intrinsic rate of increase or exponential growth rate, then:

$$\frac{dN}{dt} = rN \qquad (4.4.2)$$

This expression states that population growth is proportional to N and the instantaneous growth rate r. When $r = 0$, birth and death rates balance, individuals just manage to replace themselves, and population size remains constant. When $r < 0$, the population shrinks toward extinction, and when $r > 0$, it grows.

We integrate the differential form of this continuous growth model to project future population sizes:

$$N(t) = N(0)e^{rt} \tag{4.4.3}$$

Although r is an instantaneous rate, its numerical value is only defined over a finite interval. If this rate remains constant, then we can predict future population size,

$N(t)$ from a knowledge of the constant growth rate (r), the present population size, $N(0)$, and the time over which growth occurs (t).

4.4.2 GEOMETRIC GROWTH WITH DISCRETE GENERATIONS

Now we consider a density-independent growth model that is more appropriate for many plants, insects, mammals, and other organisms that reproduce seasonally. Individuals in such a population comprise a series of cohorts whose members are at the same developmental stage. Assume that an interval begins with the appearance of newborns, and that if individuals survive long enough, they produce another cohort of offspring at the beginning of the next interval. Parents may all die before the offspring are born (like annual plants), or they may survive to reproduce again so that generations are partially overlapping (like many mammals). In either case youngsters appear in nearly synchronous groups separated by intervals without recruitment.

This discrete population growth is best described by a finite difference equation.

If: N_t = population size at time t

b = births per female per interval

p = probability of surviving the interval, then:

$$N_{t+1} = pN_t + pbN_t = (p + pb)N_t \tag{4.4.4}$$

Redefining the collective term with birth and death rates as a single parameter $\lambda = (p + pb)$, which gives the number of survivors plus their progeny,

$$N_t = \lambda N_{t-1} = \lambda(\lambda N_{t-2}) = \lambda^t N_0 \tag{4.4.5}$$

λ is the geometric growth factor , or per capita change in population size over a discrete interval, t . If $\lambda = 1$, then individuals just manage to replace themselves and population size remains constant. If $\lambda < 1$, the population shrinks toward extinction, and if $\lambda > 1$, it grows larger. As long as remains constant, we can predict future population sizes from the growth rate (λ), the present population size (N_0), and the interval over which growth occurs (t), using the equation

$$N_t = \lambda^t N_0 \tag{4.4.6}$$

4.5 DENSITY-DEPENDENT POPULATION GROWTH

This section simulates density-dependent population growth, assuming a linear negative feedback of population size on per capita growth. It requires specification of a starting population size $N(0)$, a maximum sustainable population size or environmental carrying capacity K , a per capita intrinsic growth rate r , and (optionally) a feedback lag τ . The program includes continuous, lagged continuous and discrete simulations.

Density-dependent models assume that population size affects per capita growth. While the feedback of density on growth can take many forms, the logistic model imposes a negative linear feedback. Note that if *K* is the environmental carrying capacity (quantified in terms of individuals, *N*), then *K* – *N* gives a measure of the unused carrying capacity, and *(K-N)/K* gives the fraction of carrying capacity still remaining. Thus

$$\frac{dN}{dt} = rN\left(\frac{K-N}{K}\right) \tag{4.5.1}$$

If N is near zero, the carrying capacity is largely unused, and $\frac{dN}{Ndt}$ is near r .If N=K , the environment is totally used or occupied, and $\frac{dN}{Ndt} = 0$ In this continuous, differential equation model, r is an instantaneous rate, but its numerical value is defined over a finite time period.

To project a time trajectory of logistic population growth, we need to integrate the differential equation from time (0) to time (t).

$$N(t) = \frac{K}{1+\left(\frac{K-N(0)}{K(0)}\right)e^{-(rt)}} \tag{4.5.2}$$

A plot of *N(t)* with respect to time gives a sigmoid (S-shaped) trajectory, where growth is nearly exponential when N is near zero, and slows to equilibrium at *N*=*K* . When initial population size exceeds the carrying capacity, numbers fall in an asymptotic approach toward *K.*

Sometimes the feedback of density on per capita growth rate is not instantaneous. For example, the effect of malnutrition on population growth might not be strongly evident before malnourished juveniles reach reproductive age. We can simulate this process by assuming that growth rates are affected by population size in some previous time period. Thus

$$\frac{dN}{dt} = rN(t)\left(1-\frac{N(t-\tau)}{K}\right) \tag{4.5.3}$$

where τ is a time lag. There is no definite integral for this equation, so we project time trajectories by summing instantaneous changes in population size via numerical integration. Because the lag is delayed by an amount τ , a growing population may reach and overshoot the carrying capacity before the negative feedback term causes the population to stop growing or decline. The resulting oscillation may damp to a stable equilibrium or continue indefinitely as a limit cycle.

A population with discrete generations or cohorts cannot adjust instantaneously to changes in density-dependent feedback, because births occur only once in each generation or cohort interval. There is an implicit lag associated with the period of discrete population growth increments. With the lagged logistic model of the previous section, lag time, t, could vary in length; but with a discrete logistic model it is constant, fixed by the interval of discrete time steps. As a result, *r* and *K* alone determine the dynamics. When *r* is small, the population may not grow fast enough to overshoot carrying capacity within the lag time of a single cohort interval; but as *r* increases, sustained oscillations are more likely. Several approaches have been used to formulate difference equations analogous to the continuous logistic equation, but they yield similar results.

The version implemented is:

$$N_{t+1} = N_t e^{r\left(1-\frac{N_t}{K}\right)} \tag{4.5.4}$$

With a small population-growth rate, r , this discrete model gives a sigmoid approach to equilibrium, just like continuous and lagged logistic models. With increasing r values, discrete logistic dynamics show damped oscillation; then 2-point cycles of constant period and amplitude; and then cycles that include 4 points, 8 points, 16 points, etc., before repeating. Finally, very large r values cause population size to fluctuate in a way that is extremely sensitive to initial conditions, and never settles into a precisely repeating cycle, a regime that mathematicians term “chaotic.”

4.6 AGE-STRUCTURED POPULATION GROWTH

Youngsters and oldsters give birth and die at different rates. To keep track of these differences and their effect on population growth, biologists divide an organism's life into a series of discrete intervals, each representing a cohort of individuals that are about the same age and have similar expectations of survival and fertility. By listing S_x , the number of surviving individuals in each cohort by age, x , we can specify the composition of an age-structured population. We can also tabulate age-specific changes of fertility and survival in a life table, or $l_x m_x$ schedule. The first component, l_x , is the average probability of survival from birth to age x . The second component, m_x, is the average number of female offspring that a female can expect to acquire when she reaches age x . With these life history parameters, we can then project population growth by cohort or with a weighted average of fertility and survival rates over all ages.

This section provides three different visual representations of a life history, allowing students to see a life-table or $l_x m_x$ schedule, a life-cycle graph of the age classes and provide the data to initiate a demographic projection in any of the three formats. There are output graphs showing changes in population size, population composition, the expectation of future progeny, and a tabular output illustrating the computations that project population composition, based on the $l_x m_x$ schedule, the initial S_x values, and assumptions about the timing of reproduction and population censuses.

There are several ways of specifying the age-specific fertilities and probabilities of survival for a demographic projection. Fertilities can be tabulated as the average number of offspring accruing to a female when she reaches age x (this is the m_x of an $l_x m_x$ schedule), or the number of progeny of an age x female that are expected to be alive after the next projection interval (this is the f_x from the first row of a Leslie Matrix). Survival probabilities can be specified over x projection steps from age 0 to age x (this is l_x) or a single projection step from ages x -1 to x , or x to x +1 (this is the p_x transition probabilities, or a population state vector listing S_x, the number of survivors in each age class, with the Leslie projection matrix. Students can compare views, and just below the diagonal of a Leslie Matrix). The different styles of visual representation require converting and manipulating these survival and fertility parameters, and the exact details depend on assumptions about the timing of population censuses and reproduction of the organisms in question. When reproduction occurs seasonally at discrete intervals and the population census comes immediately after reproduction, then

$$p_x = \frac{l_x}{l_{x-1}} \quad and \quad f_x = p_x m_x \tag{4.6.1}$$

If censuses are made immediately before reproduction, then newborn individuals must survive a full projection interval before they are tabulated in their first census, so

$$p_x = \frac{l_{x+1}}{l_x} \quad and \quad f_x = l_1 m_x \tag{4.6.2}$$

Finally, when reproduction is continuous rather than pulsed in discrete seasons, the length of projection time steps is arbitrary, and probabilities of survival and reproduction are averages of the values at the beginning and end of each interval.

$$p_x = \frac{l_x + l_{x+1}}{l_{x-1} + l_x} \quad and \quad f_x = \left(\frac{1+l_1}{2}\right)\left(\frac{m_x + p_x m_{x+1}}{2}\right) \tag{4.6.3}$$

Projecting a Constant $l_x m_x$ Schedule

Consider a hypothetical population with discrete reproduction, comprised at an initial post reproductive census only of S_0 newborn individuals. At the next census, $l_1 S_0$ of those newborns will still be alive; they will have just passed age 1, each producing m_1 newborn progeny of their own, so that the new population is comprised of two cohorts. If l_x and m_x values remain constant, this process can be projected indefinitely.

The first three columns in this table give x, l_x and m_x. The shaded fourth column gives a hypothetical initial population, consisting in this case of 4 newborn individuals. The projection of this population to subsequent time steps 1-5 is made for each succeeding time step by first tabulating the number of 1- and 2-step-old adults, and then adding the progeny they can expect on reaching each age x . The projection shows that if l_x and m_x remain unchanged, the ratio of successive population sizes $\lambda = \frac{N_{t+1}}{N_t}$, often converges on a constant value, and the proportional representation of each age class then reaches a constant Stable Age Distribution. Thus, a population with a constant age-specific schedule of survival and reproduction may be started with any arbitrary composition, but will usually settle down to a constant growth rate, λ , and a stable age distribution.

It is also possible to project the constant growth of an age structured population with some simple weighted average rate estimates. The Net Reproductive Rate R_0, gives the number of female progeny expected to accrue during the entire lifetime of an individual female. It is calculated as

$$R_0 = \sum l_x m_x \qquad (4.6.4)$$

which is the sum of offspring produced in each age interval, weighted by the mother's probability of surviving to that age. The mean generation length or cohort generation time, T_c, is estimated as

$$T_c \approx \frac{\sum x l_x m_x}{\sum l_x m_x} = \frac{\sum x l_x m_x}{R_0} \qquad (4.6.5)$$

the weighted average of a female's ages when each of her progeny are born. From these two averages, we can approximate population growth, λ or *r*, as

$$\lambda \approx \frac{R_0}{T_c} \qquad or\ r \approx \frac{\ln R_0}{T_c} \qquad (4.6.6)$$

This approximation is fairly accurate for semelparous life histories (where organisms only breed once, like the Onchorhynchus salmon of the North Pacific) or populations that are not growing significantly. For iteroparous (multiple-brooding) life histories in growing populations, we determine *r* with any desired precision by successive approximation using the Lotka-Euler equation,

$$\sum e^{-rx} l_x m_x = 1 \tag{4.6.7}$$

Matrix Projection

This is represents population composition as a vector whose elements are S_x values, the numbers of individuals in each age class. To project a subsequent composition, this vector is multiplied by a transformation matrix (the "Leslie Matrix"), which has age-specific fertility values, f_x, in its first row, and probabilities of surviving from one age to the next, p_x, below the diagonal. The product of this matrix multiplication is a new vector, specifying the S_x values of the population at the next succeeding census. Students who need to review the basics of matrix multiplication. The Lotka-Euler equation is the characteristic equation of this projection matrix, and λ is its dominant eigenvalue.

$$\begin{pmatrix} S_1(t+1) \\ S_2(t+1) \\ S_3(t+1) \\ \vdots \\ S_n(t+1) \end{pmatrix} = \begin{pmatrix} f_1 & f_2 & f_3 & \cdots & f_n \\ p_1 & 0 & 0 & \cdots & 0 \\ 0 & p_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & 0 & 0 & p_{n-1} & 0 \end{pmatrix} \begin{pmatrix} S_1(t) \\ S_2(t) \\ S_3(t) \\ \vdots \\ S_n(t) \end{pmatrix} \tag{4.6.8}$$

Computational Notes

Reproductive Value, V_x, is a function of age. This is the expected number of future female progeny for a female of age x , relative to the expected future output of a newborn female, R_0

$$V_x = \left(\frac{e^{rx}}{l_x}\right)\left(\sum_{y=x}^{\infty} e^{-ry} l_y m_y\right) \tag{4.6.9}$$

When students elect to initiate a demographic simulation by specifying elements of the Leslie Matrix and population state vector, it is necessary for our program to specify l_x in terms of p_x, and m_x in terms of f_x and p_x. For continuously reproducing organisms, I am not aware that these relations have been published previously. The solution is

$$l_x = \prod_{i=0}^{i=x-1} p_i - l_{x-1} \quad \textit{and} \;\; m_x = \frac{4 f_{x-1}}{p_0 p_{x-1}} - \frac{m_{x-1}}{p_{x-1}} \tag{4.6.10}$$

Any m_x can then be found recursively, working backward from the first $m_x = 0$.

4.7 LOTKA-VOLTERRA COMPETITION

Density-dependent growth models like the logistic equation simulate an intraspecific competitive process; resources become limiting as the population increases, and the per capita growth rate declines. In this section, an additional term is added to the logistic to represent interspecific density-dependent effects, and a pair of the resulting expressions comprise the "Lotka-Volterra competition equations," which provide a simple and historically important vehicle for thinking about competitive interactions.

In the Lotka-Volterra equations, densities of both species are subtracted from the carrying capacity to give a density-dependent feedback term, and the number of interspecific competitors is weighted by a term called the competition coefficient which varies with the species' similarity in resource requirements. Thus

$$\frac{dN_1}{dt} = r_1 N_1 \left(\frac{K_1 - (N_1 + \alpha N_2)}{K_1} \right) \tag{4.7.1}$$

$$\frac{dN_2}{dt} = r_2 N_2 \left(\frac{K_2 - (N_2 + \beta N_1)}{K_2} \right) \tag{4.7.2}$$

where N_1 represents the density of species 1, K_1 is the environmental carrying capacity of species 1, and r_1 is its intrinsic rate of increase, and α is the competition coefficient, a proportionality constant defining the amount of K_1 used by every individual of species 2. In the second expression, β is an analogous coefficient weighting the effect of each species 1 individual on K_2.

Although we have no closed form solution for these equations, we can still gain interesting insights about their dynamics near the equilibrium where $\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0$.

Trivial equilibria occur when r or $N = 0$; a more interesting case occurs when $N_1 = K_1 - \alpha N_2$ and $N_2 = K_2 - \beta N_1$.

These are the equations for straight lines in N_2 vs N_1 coordinate space, the "Zero-Net-Growth Isoclines" which specify density ratios where $\frac{dN_1}{dt} = 0$ and $\frac{dN_2}{dt} = 0$, respectively.

CHAPTER - 5 POPULATION PROJECT IN TECHNIQUES

5.1 INTRODUCTION

Population projection techniques are methods used to estimate the future size and composition of a population based on historical data, current demographics, and assumptions about future trends. These projections are valuable for a variety of purposes, including urban planning, resource allocation, and policy development. Several techniques are commonly employed for population projections:

Arithmetic Growth: This basic method assumes a constant annual increase in population. It is calculated by adding a fixed number of individuals each year to the current population. While simple, it is generally not suitable for long-term projections, as it does not consider changes in growth rates.

Geometric Growth: Geometric growth is similar to exponential growth, where the population increases by a fixed percentage each year. This model is more realistic than arithmetic growth and can be used for short- to medium-term projections.

Exponential Growth: Exponential growth models assume that population growth occurs at a constant rate over time. It uses the formula N(t) = N0 * e^(rt), where N(t) is the projected population at time t, N0 is the current population, e is the base of the natural logarithm, r is the annual growth rate (expressed as a decimal), and t is the number of years into the future.

Logistic Growth: Logistic growth accounts for resource limitations and assumes that population growth levels off as it approaches a carrying capacity (K). This model is suitable for situations where populations cannot continue growing indefinitely.

Cohort Component Method: This method breaks down the population into cohorts or age groups and projects each cohort separately based on birth rates, death rates, and migration rates specific to each group. It provides detailed age-specific projections.

Rate-Based Projections: These projections use vital rates (birth rates, death rates, and migration rates) to estimate future population sizes. They can incorporate various demographic factors and are often used in official population projections.

Survival Analysis: Survival analysis techniques, such as Kaplan-Meier estimates and Cox proportional hazards models, are used to project survival rates for specific populations, such as cancer patients or participants in clinical trials.

Time Series Analysis: Time series methods, like autoregressive integrated moving average (ARIMA) models, use historical data to project future population trends, accounting for seasonality and temporal dependencies.

Demographic Transition Models: These models describe the historical shifts in birth rates, death rates, and population growth that occur as societies move from high fertility and high mortality to low fertility and low mortality. They can be used to project future trends based on a country's stage of development.

Agent-Based Modeling: Agent-based models simulate the behavior of individual agents (e.g., people) within a population and can project population dynamics based on agent interactions and rules governing their behavior.

Probabilistic Projections: Instead of producing a single projection, probabilistic methods provide a range of possible future scenarios, each with its own likelihood. Monte Carlo simulations and Bayesian methods are used for probabilistic projections.

Scenario Analysis: Scenario analysis involves developing multiple projection scenarios based on different assumptions about factors like birth rates, mortality, and migration. These scenarios can represent different policy or development options.

Projection Software: Specialized software packages, such as the United Nations' World Population Prospects or the U.S. Census Bureau's Population Projections, are often used to perform complex population projections, incorporating various demographic and statistical methods

.Effective population projection techniques depend on the availability of reliable demographic data, the understanding of current demographic trends, and the consideration of various future scenarios. Projections should also be periodically updated to account for changing circumstances and assumptions.

5.2 FREJKA’S COMPONENT METHOD OF POPULATION PROJECTION

Frejka's component method of population projection is a demographic modeling technique used to project future population sizes and age structures by breaking down the projection process into its individual components, such as births, deaths, and migration. This method allows for a more detailed and nuanced understanding of how different demographic factors contribute to population changes over time. The method is named after its developer, demographer Tomas Frejka.

The Frejka component method involves the following key components:

Fertility Component: This component focuses on projecting future births based on assumptions about fertility rates. Demographers consider factors such as age-specific birth rates, total fertility rates (TFR), and changes in fertility behavior when making fertility projections.

Mortality Component: The mortality component deals with projecting future deaths by considering age-specific death rates, life expectancy, and other mortality-related factors. Changes in mortality rates, such as improvements in healthcare or increases in life expectancy, are incorporated into mortality projections.

Migration Component: Migration plays a significant role in population dynamics, especially in regions with significant international or internal migration. This component projects future migration patterns, including immigration, emigration, and net migration, based on historical data and assumptions about future migration trends.

Age Structure Component: This component considers the current age distribution of the population and how it changes over time. It takes into account both the natural age distribution changes (births and deaths) and the age distribution effects of migration.

Sex Ratio Component: Demographers may also consider changes in the sex ratio of the population when making projections. This involves projecting male and female populations separately and accounting for potential gender imbalances.

The key steps in the Frejka component method of population projection typically involve the following:

Base Year Data: Start with reliable and up-to-date demographic data for a base year. This data includes population counts, age structure, fertility rates, mortality rates, and migration data.

Projection Assumptions: Make assumptions about future trends in fertility, mortality, and migration. These assumptions can be based on historical trends, expert judgment, or specific policy scenarios.

Projection Model: Develop a mathematical model that incorporates the component-specific assumptions. This model calculates projected population sizes and age structures for future years.

Validation: Validate the projection model by comparing projected results with historical data for previous years to assess the model's accuracy and reliability.

Scenario Analysis: Consider different projection scenarios by varying assumptions about fertility, mortality, or migration to understand the range of possible population outcomes.

Sensitivity Analysis: Conduct sensitivity analyses to assess how changes in assumptions about one component (e.g., fertility) affect overall population projections.

Reporting and Communication: Present the projection results in a clear and understandable manner, often through tables, charts, and reports. Communicate the implications of the projections for policy planning and decision-making.

The Frejka component method is a flexible and widely used approach in demography for population projection. It allows policymakers, researchers, and planners to gain insights into how different demographic factors contribute to population change and can inform strategic decisions related to healthcare, education, and social services.

(A) ESTIMATING THE SURVIVORS OF THE PRESENT POPULATION AFTER t YEARS

The first Component 'Survivors of the present population' of the population is given by

$$ {}_nP_{x+t}^{(z+t)} = {}_nP_x^{(z)}\ {}_nS_x^{(t)} \tag{5.2.1} $$

$$ \begin{bmatrix} x\ to\ x+n\ age\ group \\ population\ projection \\ afterr\ t\ years \end{bmatrix} = \begin{bmatrix} x\ to\ x+n\ age\ group \\ population\ in\ the \\ present\ year\ (base\ year\ z) \end{bmatrix} \begin{bmatrix} x\ to\ x+n\ age\ group \\ survival\ rate\ at \\ year\ t \end{bmatrix} $$

The survival rates can be obtained from life tables. Generally, 5-year survival rates are used in population projections.

$$ \text{i.e. } {}_sS_x^5 = \frac{{}_5L_{x+5}}{{}_5L_x} \tag{5.2.2} $$

(B) ESTIMATING THE NEW BIRTHS IN (0,t) PERIOD AND THEIR SURVIVORS AT PROJECTION YEAR *t*

The estimated survivors of all the new births in (*0,t*) period recorded at t^{th} year is given by

$$ \Phi_{x+t} = \sum_{x=\lambda_1}^{\lambda_2} f_{p_x} \Pi_x \sum_{i=0}^{n} \sum_{x=n}^{kn} \left[\frac{L_{x-i}}{l_0} \right] \tag{5.2.3} $$

Where f_{p_x} =No. of females in the age group (x, x+1)

$\prod_x$ = Probability of giving birth in the year for a female in the age group (x, x+1).

$\prod_x^j$ = Probability of giving birth in the year j for a female in the age group (x, x+1).

λ_i and λ_c are the lower and upper bounds of the child bearing period respectively.

5.3 REPRESENTATION OF THE COMPONENT METHOD BY THE USE OF LESLIE MATRIX (OR L. MATRIX).

P.H.Leslie (1945) represented the component method of population projections by a matrix which is known as leslie matrix or l-matrix. He made the following assumptions for representing this projection process:

1. The projection is made on every 5th year ie., t=5;

2. The representation process has been made with respect to the female population only;
3. Population is taken as stable population with the birth and death parameters as independent of time;
4. Child bearing age is taken as 15 to 45 years:

Notations:

${}_nP_x^{(t)}$ = Female population in the age group (*x, x+n*) at time *t*

${}_nF_x$ = Age specific fertility rate in the group (*x, x+n*) independent of time t.

With the above assumptions and notations, Leslie represented the component method of population projections as follows:

Consider the 'Survivors of the present population' component of the population projection as

$$ {}_nP_{x+t}^{(z+t)} = {}_nP_x^{(z)}\, {}_nS_x^{(t)} \qquad (5.3.1)$$

And

$$ {}_sS_x^{(t)} = \frac{{}_nL_{x+t}}{{}_nL_x} = survival\ Rate $$

Where ${}_nP_{x+t}^{(z+t)}$ = x to x+n age group population projection after t years

${}_nP_x^{(z)}$ = x to x+n age group population in the present year (base year)

${}_nS_x^{(t)}$ = x to x+n age group survivor rate at year t and ${}_sS_x^{(t)} = \frac{{}_nL_{x+t}}{{}_nL_x}$

We have

$$ {}_5P_5^{(t+5)} = {}_5P_0^{(t)}\, {}_5S_0 \ where\ {}_5S_0 = \frac{{}_5L_5}{{}_5L_0} \qquad (5.3.2)$$

$$ {}_5P_{10}^{(t+5)} = {}_5P_5^{(t)}\, {}_5S_5 \ where\ {}_5S_5 = \frac{{}_5L_{10}}{{}_5L_5} \qquad (5.3.3)$$

$$ {}_5P_{15}^{(t+5)} = {}_5P_{10}^{(t)}\, {}_5S_{10} \ where\ {}_5S_{10} = \frac{{}_5L_{15}}{{}_5L_{10}} \qquad (5.3.4)$$

$${}_5P_{45}^{(t+5)} = {}_5P_{40}^{(t)}\,{}_5S_{40} \text{ where } {}_5S_{40} = \frac{{}_5L_{45}}{{}_5L_{40}} \qquad (5.3.10)$$

Also

${}_5P_0^{(t+5)}$ Is given by

$${}_5P_0^{(t+5)} = \frac{1}{2}\left[{}_5P_{15}^{(t)} + {}_5P_{15}^{(t+5)}\right]{}_5F_{15}\Pi_0 + \frac{1}{2}\left[{}_5P_{20}^{(t)} + {}_5P_{20}^{(t+5)}\right]{}_5F_{20}\Pi_0 + ---------- + \frac{1}{2}\left[{}_5P_{40}^{(t)} + {}_5P_{40}^{(t+5)}\right]{}_5F_{40}\Pi_0$$

(5.3.11)

Where

$\Pi_0 = \frac{{}_5L_0}{l_0} =$ the probability of survival from birth to the age group 0 to 0+5

$\Pi_x = \frac{{}_5L_x}{l_0} =$ the probability of survival from birth to the age group x to x+5

Using the expressions for ${}_5P_{15}^{(t+5)}, {}_5P_{20}^{(t+5)}, {}_5P_{20}^{(t+5)}, \ldots\ldots\ldots, {}_5P_{40}^{(t+5)}$ we rewrite ${}_5P_0^{(t+5)}$ as

$${}_5P_0^{(t+5)} = \frac{1}{2}\left[{}_5P_{15}^{(t)} + {}_5P_{10}^{(t)}\,{}_5S_{10}\right]{}_5F_{15}\Pi_0 + \frac{1}{2}\left[{}_5P_{20}^{(t)} + {}_5P_{15}^{(t)}\,{}_5S_{15}\right]{}_5F_{20}\Pi_0 + ---------- + \frac{1}{2}\left[{}_5P_{40}^{(t)} + {}_5P_{35}^{(t)}\,{}_5S_{35}\right]{}_5F_{40}\Pi_0$$

(5.3.12)

$$\begin{aligned}{}_5P_0^{(t+5)} &= \frac{1}{2}\left[{}_5S_{10}\,{}_5F_{15}\right]\Pi_0\,{}_5P_{10}^{(t)} + \frac{1}{2}\left[{}_5F_{15} + {}_5S_{15}\,{}_5F_{20}\right]\Pi_0\,{}_5P_{15}^{(t)} + \frac{1}{2}\left[{}_5F_{20} + {}_5S_{20}\,{}_5F_{25}\right]\Pi_0\,{}_5P_{20}^{(t)} + ---------\\ &+ \frac{1}{2}\left[{}_5F_{35} + {}_5S_{35}\,{}_5F_{40}\right]\Pi_0\,{}_5P_{35}^{(t)} + \frac{1}{2}\left[\Pi_0\,{}_5F_{40}\,{}_5P_{40}^{(t)}\right]\end{aligned}$$

(5.3.13)

The equations (5.3.2) to (5.3.13) can be expressed in the form of a matrix, known as Leslie matrix.

$$\begin{bmatrix} {}_5P_0^{(t+5)} \\ {}_5P_5^{(t+5)} \\ {}_5P_{10}^{(t+5)} \\ {}_5P_{15}^{(t+5)} \\ \vdots \\ \vdots \\ {}_5P_{45}^{(t+5)} \end{bmatrix}_{10X1} = \begin{bmatrix} 0 & 0 & \frac{\Pi_0}{2}\,{}_5S_{10}\,{}_5F_{15} & \frac{\Pi_0}{2}\left({}_5F_{15}+{}_5S_{15}\,{}_5F_{20}\right) & \cdots & \cdots & \cdots & \frac{\Pi_0}{2}\left({}_5F_{40}\right) \\ {}_5S_0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & {}_5S_5 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & {}_5S_{10} & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & {}_5S_{40} \end{bmatrix}_{10X9} \begin{bmatrix} {}_5P_0^{(t)} \\ {}_5P_5^{(t)} \\ {}_5P_{10}^{(t)} \\ {}_5P_{15}^{(t)} \\ \vdots \\ \vdots \\ \vdots \\ {}_5P_{40}^{(t)} \end{bmatrix}_{9X1}$$

(5.3.14)

$$\Rightarrow P_{10X1}^{(t+5)} = L_{10X9} P_{9X1}^{(t)} \qquad (5.3.15)$$

Where $P^{(t+5)}$ and $p^{(t)}$ represent the population age vector in year (t+5) and t respectively:

L is the Leslie matrix, consists of the elements which are functions of Fertility and mortality parameters which are independent of time.

Thus, with time independent Leslie matrix,

we have,

$$P^{(t+5)} = L\,P^{(t)}$$

$$P^{(t+10)} = L\,P^{(t+5)} = L\,L\,P^{(t)} = L^2\,P^{(t)}$$

$$P^{(t+15)} = L\,P^{(t+10)} = L^3\,P^{(t)} \quad :$$

:

:

$P^{(t+5k)} = L^k\,P^{(t)},$ for an integer k.

This shows the sequence { $P^{(t)}$, $P^{(t+5)}$, $P^{(t+10)}$,…… } gives a simple mark chain.

Remark:

Matrix has only one positive Eigen value and the other eigen values are either complex or negative

5.4 CHANDRASEKHARAN AND DEMING'S METHOD OF ESTIMATING VITAL RATES FROM REGISTRATION DATA

C. Chandrasekharan and W.I. Deming (1949) have proposed a mathematical technique which is useful to compose the Registrar's list of births and deaths with that if a list obtained from a house to house canvass. This method provides the estimates of the total no. of events over the area in a specified period.

Let

R: The Registrar's list

I: List obtained by Interviewers by a genyalete house to house canvass

C: The no. of events which are recorded in I as well as in R and these are correct events

N_1: Entries recorded in R but not in I and after investigation found to be correct

N_2: Entries recorded only in I but not in R and after investigation found to be correct

X: Entries recorded on one list r the other but not in both and found in correct after verification

N: The total no. of events

Y: No. of events which are missed by both R and I

Chandrasekharan and Deming Formula:

Chandrasekharan and Deming estimate of N is given by

$$\hat{N} = C + N_1 + N_2 + \hat{Y} \qquad (5.4.1)$$

We have, $$\hat{Y} = \frac{N_1 N_2}{C}$$

(5.4.2)

Proof:

Probability of R detecting and event is

$$p_1 = \frac{C + N_1}{N}$$

Probability of I detecting an event is

$$p_2 = \frac{C + N_2}{N}$$

Probability of an event being missed by both R and I as

$$(1 - p_1)(1 - p_2) = \left(1 - \frac{C + N_1}{N}\right)\left(1 - \frac{C + N_2}{N}\right)$$

$$\therefore (C + N_1 + N_2 + Y)\left(1 - \frac{C + N_1}{N}\right)\left(1 - \frac{C + N_2}{N}\right) = Y$$

$$\Rightarrow (C + N_1 + N_2 + Y)\frac{N_1 N_2}{(C + N_1)(C + N_2)} = Y$$

$$\because \hat{N} = C + N_1 + N_2 + \frac{N_1 N_2}{C} = \frac{(C + N_1)(C + N_2)}{C} \qquad (5.4.3)$$

Because $\hat{Y} = \dfrac{N_1 N_2}{C}$

It should be noted that

(i) $\hat{N} = C + N_1 + N_2 + \dfrac{N_1 N_2}{C}$ is an unbiased estimate of N

(ii) $\hat{N}$ is the maximum likelihood estimate of N

(iii) The standard error of $\hat{N}$ is given by $S.E(\hat{N}) = \sqrt{\dfrac{N q_1 q_2}{p_1 p_2}}$ where

$q_1 = 1 - p_1 \; and \; q_2 = 1 - p_2$

Remark: The Validity of this method depends on the correlation of events in the list of R and I.

5.5 LOTKA AND DUBLIN (1949) MODEL FOR STABLE POPULATION ANALYSIS

This analysis is based on the consideration that

(i) The population growth is independent of time when both fertility and the mortality

rates are also time independent and

(ii) The population is closed under migration (there is balance between migrants and

Immigrants)

The structural form of the stable population is characterized by

(i) Birth rate is independent of t

(ii) Death rate is independent of t

(iii) The age distribution between ages (x, x + δ x) is independent of t.

(iv) The population is closed to migration

Notations:-

(i) $C(x,t)\delta x$ = The proportion of population is the age group (x, x+ δ x) at time t

(ii) $B(t)$= Total no. of births at time t

(iii) $P(t)$= Population at time t

(iv) $b(t) = \frac{B(t)}{P(t)}$= Birthrate per individual

(v) $p(x)$= Probability of surviving up to age x. (Independent of t)

(vi) $f(x)\delta x$ = Probability of giving a birth between age (x, x + δ x) (independent of t)

(vii) $dx\delta x$ = Probability of giving between (x, x+ δ x)

This analysis is restricted to female cohorts, we have the basic identity

$$P(t)C(x,\ t)\delta x\ = B(t-x)\,p(x)\delta\,x \tag{5.5.1}$$

Where $B(t-x)$ is no. of births at time t by the age group (x, x + δ x)

$$\Rightarrow p(t)C(x,t)\ = B(t\ -\ x)\ p(x)$$

$$\Rightarrow \int_0^\infty p(t)C(x,t)f(x)dx\ =\ \int_0^\infty B(t\ -\ x)\ p(x)f(x)dx \tag{5.5.2}$$

The R.H.S of (5.5.2) gives the births at time t.

Thus, we can write

$$B(t)\ =\ \int_0^\infty B(t\ -\ x)\ p(x)f(x)dx \tag{5.5.3}$$

(5.5.3) is an integral equation with time lag x.

Lotka and Dublin assumed a trial solution of the form

$$B(t)=\sum_{n=0}^{\infty} Q_n e^{r_n t} \tag{5.5.4}$$

Where Q0, Q1, Q2, ………….. are the populations at the beginning of each year under consideration and treated them as constants;

Substituting (5.5.4) in (5.5.3) gives

$$\sum_{n=0}^{\infty} Q_n e^{r_n t} = \int_0^{\infty} \sum_{n=0}^{\infty} Q_n e^{r_n(t-x)} p(x)\, f(x)dx$$

$$\sum_{n=0}^{\infty} Q_n e^{r_n t} = \sum_{n=0}^{\infty} Q_n e^{r_n t} \left[\int_0^{\infty} e^{-r_n t} p(x)\, f(x)dx \right] \tag{5.5.5}$$

It appears that $r_0, r_1, r_2, \dots\dots\dots r_n$ correspond to the roots of the integral

$$\int_0^{\infty} e^{-rx} p(x)\, f(x)dx = 1 \tag{5.5.6}$$

Equation(5.5.6) is known as Lotka's Integral equation.

$$(5.5.6) \Rightarrow \int_0^{\infty} e^{-rx} \phi(x)\, dx = 1 \tag{5.5.7}$$

Where $\phi(x) = p(x)\, f(x)$ is the net maternity function.

Remarks: $\int_0^{\infty} p(x)\, f(x)dx = R_0 =$ net reproduction rate per women

To obtain real root of the Lotka's integral equation (5.5.6),

We put

$$y = \int_0^{\infty} e^{-rx} p(x)\, f(x)dx \tag{5.5.8}$$

$$\Rightarrow \frac{dy}{dr} = \int_0^{\infty} \frac{d}{dr}\left(e^{-rx} p(x)\, f(x)dx\right)$$

$$\Rightarrow \frac{dy}{dr} = -\int_0^{\infty} x\left(e^{-rx} p(x)\, f(x)dx\right)$$

$$= -\left[\frac{\int_0^\infty x\left(e^{-rx} p\left(x\right) f(x) dx\right)}{\int_0^\infty e^{-rx} p\left(x\right) f(x) dx}\right]\left[\int_0^\infty e^{-rx} p\left(x\right) f(x) dx\right]$$

$$= -\{A(r)\}Y$$

Where

$$A(r) = \left[\frac{\int_0^\infty x\left(e^{-rx} p\left(x\right) f(x) dx\right)}{\int_0^\infty e^{-rx} p\left(x\right) f(x) dx}\right] \tag{5.5.9}$$

We have the differential equation

$$\frac{dy}{dr} = -Y.A(r) \tag{5.5.10}$$

By solving differential equation, a solution of r is given by

$$r = \frac{-2\left(\frac{R_1}{R_0}\right) \pm \sqrt{4\left(\frac{R_1}{R_0}\right)^2 - 8\left[\left(\frac{R_1}{R_0}\right)^2 - \frac{R_2}{R_0}\right]\log_e R_0}}{2\left[\left(\frac{R_1}{R_0}\right)^2 - \frac{R_2}{R_0}\right]} \tag{5.5.11}$$

Where r is the growth parameter of stable population.

Here R_0, R_1 and R_2 are estimated as

$$\hat{R}_0 = \sum p\left(x\right) f(x) = N.R.R \tag{5.5.12}$$

$$\hat{R}_1 = \frac{\sum xp(x)\, f(x)}{\sum p(x)\, f(x)} = \text{ Mean age of child bearing} \qquad (5.5.13)$$

$$\hat{R}_2 = \frac{\sum x^2 p(x)\, f(x)}{\sum p(x)\, f(x)} \qquad (5.5.14)$$

By substituting $\hat{R}_0, \hat{R}_1$ and $\hat{R}_2$ for R_0, R_1 and R_2 in (5.5.11), the two real roots of *r* can be obtained as one root is positive and the other is negative.

5.6 CONCEPT OF STATIONARY POPULATION

A hypothetical model of a population, based on unchanging conditions of fertility, mortality and the total size, is called a stationary population. A population generated by any given life table is essentially a stationary population. The crude birth and death rates of the stationary population are equal.

The main characteristics of stationary population are:

1. The total size of the population is constant
2. The annual no. of births (and also deaths) is constant.
3. The age composition of population is invariable
4. Crude birth rate is equal to crude death rate.

i.e., These characteristics are fixed in time.

5.7 CONCEPT OF STABLE POPULATION

Any population with a constant age distribution and which is increasing at a constant rate is called a stable population. A stationary population is then a special case of a stable population in which the rate of growth is zero and the age distribution is same as the life table age distribution.

Certain Results in Stable Population Analysis

(i) Birth and death rates are independent of time.

(ii) $C(x,\ t) = C(x)$ ie., Age distribution is independent of 't'.

Thus birth and death rates as well as age distribution may undergo changes but these changes are only of random nature.

Certain Important Deductions of stable population Analysis

In the stable population analysis, we have

(i) $B(t) = B(t\ -\ x)e^{r\,x}$

(5.7.1)

(ii) $$C(x;t) = \frac{B(t\ -\ x)\,p(x)}{p(t)} = b(t)e^{-r\,x}\,p(x) \tag{5.7.2}$$

(iii) $$b(t) = \frac{1}{\int_0^{\infty} e^{-rx}\,p(x)\,dx} = \text{birth rate (independent of time)} \tag{5.7.3}$$

(iv) $$C(x;t) = \frac{e^{-rx}\,p(x)}{\int_0^{\infty} e^{-rx}\,p(x)\,dx} = \text{ age distribution (independent of time)} \tag{5.7.4}$$

REFERENCES

1. Burnham, K.P. & Anderson, D.R. 1998 Model selection and inference. A Practical information – Theoretic approach. New York. Springer.
2. Draper, N.R. & Smith H. 1981. Applied Regression Analysis. 2nd edition. John Wiley & Sons Inc. New York 709P.
3. Hamilton, J.D., 1994. Time series Analysis, Princeton University press, Princeton.
4. Hosmer,D.w., and Lemeshow,S. (2000); "Applied Logistic Regression", Second Edition, John Wiley & Sons, INC., New York.
5. Johnson, N.L., Kotz. S. and BalaKrishna, N. (1995) "Continuous Univariate Distributions", Vol.1, 2nd ed., John Wiley and Sons.
6. Kerbs, C.J 2002 Two Complementary Paradigms for analyzing population dynamics. Phil. Trans. R. Soc. Land. B 357, 1211 – 1219. (DOI 10. 1098/rstb. 2002.1122)/
7. Kleinbaunm,D.G., L.L Kupper and Muller, K.E.(1998),"Applied Regression Analysis and Multivariate Methods", Second Edition, Brooks/Cole Publishing Company, Pacific Grove, California.
8. Montgomery,D.C., Peck,E.A., and Vining,C.G.(2003), "Introduction to Linear Regression Analysis", Third Edition, John Wiley & Sons, INC., New York.
9. Nelson, W. (1990). "Accelerated Life Testing: Statistical models, Data analysis and Test plans", John Wiley and sons, New York.
10. Newton, I. 1998 population limitation in birds, London: Academic press.