FEDERAL UNIVERSITY OF TECHNOLOGY
MINNA

PREDICTIVE STOCHASTIC MODELING OF LEPROSY DISEASE, HUMAN HEALTH CONDITION AND DESERTIFICATION IN NIGERIA: THE UNTAPPED OPEN SECRET INGREDIENTS

By

PROF. USMAN YUSUF ABUBAKAR
B.Ed (Zaria), M.Sc (Jos), PGD, PhD (Minna)
Professor of Mathematics

INAUGURAL LECTURE SERIES 57

26TH OCTOBER, 2017

FEDERAL UNIVERSITY OF TECHNOLOGY MINNA

PREDICTIVE STOCHASTIC MODELING OF LEPROSY DISEASE, HUMAN HEALTH CONDITION AND DESERTIFICATION IN NIGERIA: THE UNTAPPED OPEN SECRET INGREDIENTS

By

PROF. USMAN YUSUF ABUBAKAR

B.Ed (Zaria), M.Sc (Jos), PGD, PhD (Minna)

Professor of Mathematics

INAUGURAL LECTURE SERIES 57

26TH OCTOBER, 2017

This 57th Inaugural Lecture was delivered
under the Chairmanship of:

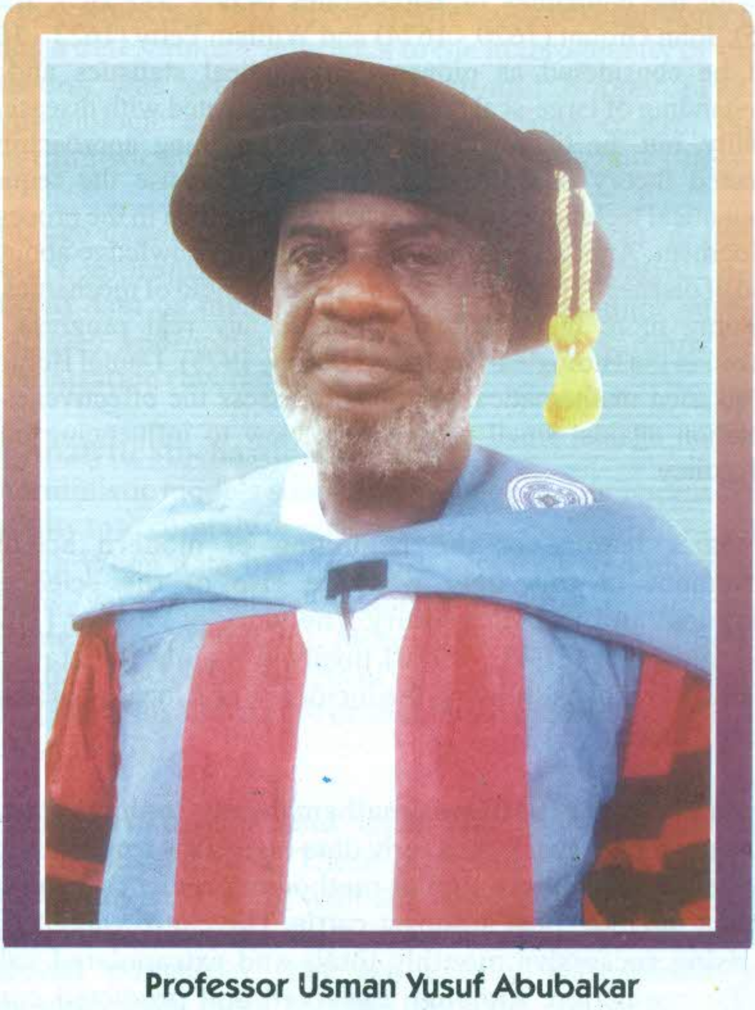
Professor M. A. Akanji, FNSMBM, FAS
Vice-Chancellor
Federal University of Technology, Minna

ISSN 2550 - 7087

Published by:
University Seminar and Colloquium Committee
Federal University of Technology, Minna.

26th October, 2017

Design + Print:
Global Links Communications, Nigeria
✆: 08056074844, 07036446818



Professor Usman Yusuf Abubakar

B.Ed (Zaria), M.Sc (Jos), PGD, PhD (Minna)

Professor of Mathematics

1. The Beginnings

The modeling of disease started as far back as the ancient Greeks, with the epidemics of Hippocrates (459 - 377 BC), Bailey (1975). John Grount (1620 - 1674) and William Petty (1623 - 1687) could be considered as pioneers of medical statistics and the understanding of large-scale phenomena connected with disease and mortality, but the time was not ripe for anything approaching a connected theory of epidemics. This was because the requisite mathematical techniques were themselves only then in the process of development. Another reason was insufficient knowledge about the spread of disease. A good start was made in the field of mechanics and astronomy more than 200 years before any real progress was achieved in the Biological Sciences (Barley, 1975). Daniel Bernoulli in 1760 used mathematical methods to assess the effectiveness of inoculation against small pox, with a view to influencing public health policy.

The major feature of the beginning of modern scientific achievement in this field was the rise of the science of bacteriology in the 19th century. The work of Pasteur (1822 - 1895) and Koch (1843 - 1910) involved mainly the statistical appraisal of records showing the incidence and locality of known cases of disease.

The work of Farr (1840) was mathematically sophisticated. He fitted a normal curve to quarterly data on deaths from smallpox. Brownlee (1866) used a similar method to predict the course of outbreak of rinderpest amongst cattle. The curve was fitted to four rising successive monthly totals and extrapolated values used for prediction. Although observed and predicted curves were both bell-shaped, agreement in detail was not very good.

The work of Farr and Brownlee involved more of curve fitting and prediction.

2. Mathematical Modelling

Ross (1911) developed a mathematical model for malaria, which attempted to take into account a set of measures describing various aspects of transmission. The study of respiratory disease using a deterministic approach to the heterogeneity of spread of infection was provided by Becker and Hopper (1983). An epidemiologic application of sophisticated control theoretic deterministic modeling was provided by Hethcote (1983).

The age-dependent immunization model was designed to predict appropriate strategies for disease control. Hethcote utilized data on measles and rubella to determine vaccination strategies appropriate for their control at various levels of immunization coverage.

2a. Analytic Stochastic Modeling

Deterministic models soon lost their popularity because of their inability to accurately describe recurrent cycles of disease, Bailey (1982). When data became more extensive and much smaller groups were considered, elements of "chance and variation" became more prominent. Mckendrick (1926) was the first to construct stochastic models of epidemic processes. Greenwood (1931) gave an alternative probability treatment five years later, Bailey (1975).

"Continuous infection" and "chain binomial" stochastic models were introduced next. These probability models were more appropriate for dealing with smaller groups in which random variation would play a larger role. Although these models achieved popularity they are usually mathematically and computationally more complex than the simple deterministic models, Korve (1993).

Stochastic models now appear more frequently in the study of diseases, Bailey (1975). Kimber and Crowder (1984) proposed a model to analyze resistance times to infection under treatment. A general stochastic model was proposed by Hillis (1979).

Several stochastic models have been presented to describe distributions of infectious disease over time and space, including Korve (1993) and Goldacre (1977) who attempted an analysis of meningitis using space-time clustering techniques introduced by Knox (1964) to detect the existence of factors associated with infection.

3. Modeling Approaches for Disease Control

Generally, there are three modeling approaches for disease control: Deterministic, Analytical stochastic, and Simulation, usually stochastic.

Deterministic and stochastic models were developed in the early part of the 20^{th} century. The etiology of disease is of primary concern to many epidemiologist and can be seen either in a deterministic or stochastic framework.

A deterministic perspective is one in which factor x cause y if (all other factors being held constant) a change in the value of x results in a change in the values of y, in a completely prescribed way tracing out a mathematical function of some form.

In practice, probability theory and statistical techniques are used to assess evidence regarding causality. In any causal analysis of data, the goal is to account for variation in the dependent variable.

Several models of this sort have been utilized to analyze data in studies of infectious diseases, including most commonly linear regression, log-linear analysis, logistic regression, discriminant analysis, and proportional hazards modeling, Korve(1993).

4. Probability

The approach to probability used in this work is the axiomatic, in which probabilities are defined as "mathematical objects" which behave according to certain well-defined rules. Then any other probability concepts, or interpretations, can be used, so long as it is consistent with these rules. Thus probabilities are values of a set function, also called a probability measure; this function assigns real numbers to the various subsets of the sample space. John et al (1980).

5. Random Variables

It frequently occurs that in performing an experiment, we are mainly interested in some function of the outcome as opposed to the actual outcome itself. For instance, in tossing dice we are often interested in the sum of the two dice and are not really concerned about the actual outcome. That is, we may be interested in knowing that the sum is three and not be concerned over whether the actual outcome was (1,2) or (2,1). These quantities of interest or formally, these real-valued functions defined on the sample space are known as random variables. Since the value of a random variable is determined by the outcome of an experiment, we may assign probabilities to the possible values of the random variable.

6. Discrete Random Variable

A random variable that can take on at most a countable number of possible values is said to be discrete. For a discrete random variable *X*, we define the probability mass function

$$p(a) = P\{X = a\}$$

The probability mass function P(a) is positive for at most a countable number of values of a. That is, if X must assume one of the values $x_1, x_2, \ldots\ldots$, then

$$p(x_i) > 0, \quad i = 1, 2, \ldots..$$

$$p(x) = 0, \quad all\ other\ values\ of\ x$$

Since X must take on one of the values x_i, we have

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

7. Continuous Random Variables

A random variable whose set of possible values is uncountable. Let X be such random variable. X is a continuous random variable if there exists a non-negative function $f(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set B of real numbers

$$P\{X \in B\} = \int_B f(x)dx \qquad (1)$$

The function $f(x)$ is called the probability density function of the random variable X.

In words, equation (1) states that the probability that X will be in B may be obtained by integrating the probability density function over the set B. Since X must assume some value, $f(x)$ must satisfy

$$1 = P\{X \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(x)dx$$

All probability statements about X can be answered in terms of $f(x)$. For instance, letting

B=[a,b], we obtain $P\{a \le X \le b\} = \int_a^b f(x)dx$

8. Stochastic Processes

The family of random variables $X(t)$, $t \geq 0$ indexed by the time parameter t. The values assumed by the process are called 'states' and the set of possible values is called the state space. The set of possible values of the indexing parameter is called the 'parameter space' which can be either continuous or discrete. In the discrete case, the process is represented as X_n, $n = 0,1,2, ...$

The stochastic process occurring in most real-life situations are such that for a discrete set of parameters t_1, t_2,........t_n t, T, the random variables $X(t_1)$, $X(t_2)$, ... $X(t_n)$ exhibit some sort of dependence. The simplest type of dependence is the first-order dependence underlying the stochastic process. This is called Markov dependence.

Depending on the nature of the state space and the parameter space, we can divide Markov processes into four classes, which are given here in the form of a table. Wherever the parameter and state spaces are discrete the Markov process is called Markov chains. Otherwise the process is simply referred to as a Markov process or memoryless property.

Table 1: Classification of Markov process

PARAMETER SPACE	STATE	SPACE
	Discrete	Continuous
Discrete	Markov Chain	Markov Process
Continuous	Markov Process	Markov Process

The other typical Markov processes include Semi'Markov, Hidden Markov, Markov Decision and partially Observable Markov processes.

9. Markov Chains

Markov chains is the Markov process with discrete time and parameter spaces whose state space could be finite or countably infinite.

Let $\{X_n, n = 0, 1, 2,...\}$ be a Markov chain with a state space $S \subseteq Y = \{0,1,2, ...\}$. While discussing a finite m-state chain, we shall identify the state space S to be given by the set (1, 2, ...m). The element P_{ij}, means the probability that $x_1 = j$ if you know that $x_0 = i$. It is conditional probability

$P_{ij}, = p(X_1=j|X_0=i)$.

In the time homogeneous Markov chain the n – step transition probabilities is defined $p_{ij}^{(n)} = p(X_n=j|(X_0=i)$

The conditional probability $P(X_n=j|X_{n-1}=i)$ is referred to as one step transition probability from i to j at time n. If for all m and n, $P(X_n=j|X_{n-1}=i) = P(X_m=j|X_{m-1}=i)$ the Markov chain is said to be stationary. Stationary and time homogenous are synonymous.

10, Markov Chain and Classification of States

The value of x_n for a specific realization of the process is called the state of the process.

Definition (1): State j is said to be accessible from state i if j can be reached from i in a finite number of steps. If two states i and j are accessible to each other, then they are said to communicate.

It has been shown Bhat (1984) that all the states that communicate in a finite Markov chain form an equivalence relation.

Definition (2):- A state i is said to be recurrent if and only if starting from state i, eventual return to this state is certain.

In terms of probabilities $f^{*}ii$ this implies that the state i is recurrent if and only if $f^{*}ii = 1$

Definition (3): A state i is said to be transient if and only if, starting from the state, there is a positive probability that the process may not eventually return to this state. This implies that $f^{*}ii < 1$

Definition (4): A state i is said to be an absorbing state if when entered cannot be departed, also if and only if $P_{ii} = 1$. When i is absorbing $f^{*}ii^{(1)} = P_{ii} = 1$ and hence $f^{*}ii = 1$.

11. Non-Homogenous Markov Chain

In modeling the non-stationary transitions and time variation effect on the transitions,

Let transition matrice

$$M_k = f_{ij}(k), \quad i, j = 1, 2, 3 \ldots \quad and \quad k = 1, 2 \ldots$$

$$and \; P_k = P_{ij}(k)$$

$f_{ij}(k)$ denotes the transition count from state i to state j for the season k. $P_{ij}(k)$ is the transition probability from state i to state j for the season k.

Accordingly,

$$\hat{P}_{ij}(k) = \frac{f_{ij}(k)}{f_i(k)}, \quad k = 1, 2 \ldots \; and \quad i, j = 1, 2, 3 \ldots, n$$

$$where \; f_i = \sum_{j=1}^{n} f_{ij}(k)$$

12. Test for Stationarity of the Probability Matrices

To test for independence of P_K on K. the Null hypothesis is stated thus

H_0: $P_{ij}(k) = P_{ij}, for \; all \; i, j = 1,2,3$ and for all k

H_1: $P_{ij}(k)$ depends on K.

The likelihood ratio test for the above hypothesis is

$$M = \sum_{k=1} M_k = \left[f_{ij} \right]$$

$$where\ f_{ij} = \sum_{k=1} f_{ij}(k)$$

The maximum likelihood estimate of the stationary transition probability matrix is

$$P_{ij} = \frac{f_{ij}}{f_i}$$

$$where\ f_i = \sum_{j=1}^{n} f_{ij}$$

The λ, the likelihood ratio criterion is given by

$$\lambda = \prod_{i,j=1}^{n} \prod_{k=1} \left[\frac{P_{ij}}{P_{ij}(k)} \right]^{f_{ij}(k)}$$

$$-2In\lambda = \chi 2_{M(M-1)(T-1)}$$

where m is the number of states and T is the time parameter. We evaluate λ, and calculate $-2\,In\,\lambda$. We then get the critical value of χ^2 *at* α significance level and compare it with $-2\,In\,\lambda$. It is then decided to accept or reject the Null hypothesis. With the acceptance of H_0, we have a homogeneous Markov chain model. Otherwise we have the non-homogenous Markov chain model, Bhat (1984).

The stationarity assumption is one of 'constancy' over time. It suggests stability of the process, although of course it does not imply that the process remains in fixed state or even that there is a sluggishness in the rate at which transition occur. It is the probability mechanism that is assumed stable Ross(1989), Chung(1967).

n–step refers to the time interval between observations. In the matrix form

$$P = \left[P_{ij} \right]$$

and

$$P^{(n)} = \left[P_{ij}^{n} \right]$$

we have $\sum_{j \in S} p_{ij}^{(n)} = 1$

13. The n-step Transition Probability Matrix

The n-step transition probabilities $P_{ij}^{(n)}$ and the unconditional probabilities $P_j^{(n)}$, i, j, $\in$ S are determined by the following.

$$P^{(n)} = P^{n}$$

and $$P^{(n)} = P^{(0)} P^{n}$$

14. Ergodic Markov Chains

When the process is irreducible, recurrent, and aperiodic, we call the Markov chain ergodic. When the model is ergodic, several additional quantities, other than the transition probabilities can easily be calculated. Two of the most important of these are steady – state probabilities and mean first passage times.

Mathematically, $P^{(n)}$ and $P^{(n+1)}$ are essentially the same for large n,

$$P^{(n)} = P^{(n+1)} P$$

and $$\lim_{n \to \infty} P^{(n)} = \lim_{n \to \infty} P^{(n+1)} P$$

$$\pi = \pi P$$

and

$$\sum_{all\ i} \pi_i = 1$$

Feller (1971), Howard (1960), Stidham (2000) and Uche (2001).

15. A Review of the Methods of Determining the Steady State Probabilities in an Ergodic Markov Chain Models

It is observed that the most common type of Markov chain model in practical applications is of the ergodic category. When the model is ergodic, several quantities such as the transition probabilities, occupancy probabilities, first passage probabilities, mean first passage times and of course the steady-state probabilities can be calculated Ross (1989), Bhat (1984) and Howard (1960). The techniques include the matrix multiplication, the linear equations and substitution method and the transition diagram approach

16. Matrix Multiplication Technique

This is the most popular method. It involves the multiplication of two matrices defined thus: suppose $A = (a_{ij})$ and $B = (b_{ij})$ are matrices such that the number of columns of A is equal to the number of rows of B; say A is an M x P matrix and B is P x n matrix, then the product AB is the m x n matrix whose ij entry is obtained by multiplying the i^{th} row A_i of A by the j^{th} column B^j of B.

Symour (1981)

The n-step transition probabilities are defined by

$P_{ij}{}^{n} = P(X_n = j|X_{0=I}\ )$.

The equation can be represented in matrix form by

$$P^n = P^{n-1}P$$

where P^n is the matrix whose elements are the n-step transition probabilities. In general, the n-step transition materials equal the one-step transition matrix raised to the n^{th} power.

In introducing the steady state probabilities, we observed that the rows of the n-step transition matrix repeat themselves when n is sufficiently large. The most important advantage of this method is that it lends itself to a computer implementation.

17. The Linear Equations/Substitution Method

Let us define the steady state probabilities thus:

$$\pi_j = \lim_{n\to\infty} P_{ij}^n$$

as n grows large, $P^n = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \ldots\ldots\ldots \\ \pi_1 & \pi_2 & \pi_3 \ldots\ldots\ldots \\ \pi_1 & \pi_2 & \pi_3 \ldots\ldots\ldots \end{pmatrix}$

As long as the process is ergodic, it has been proved Ross (1989) that such limits exist.

It is tedious to find the steady-state probabilities by taking higher and higher powers of P, particularly when P is larger. However, we can generate large equations to determine them directly as follows:

$$P^n = P^{n-1}P$$

$$\lim_{n\to\infty} P^n = \lim_{n\to\infty} P^{n-1}P$$

This represents many replications of the same set of equations $\pi = \pi P$

This set of equations has many as unknown. It is a dependent set and therefore posses an infinite number of solutions. The dependence is derived from the fact that every row of P sum to unity. One of the infinite number solutions will qualify as a probability distribution. This one solution can be forced by requiring that π_j sum to 1, that is $\sum_{all j} \pi_j = 1$

To those previously expressed and the resulting set of linear equations will posses a unique solution satisfying all the requirements of a probability distribution.

This method although not very easy produces the exact steady-state probabilities.

18, The Transition Diagram Method

This is a kind of physical analog to the determination of the steady state probabilities. The approach is to think of the points as small reservoirs and the arcs as connecting pipes through which liquid can flow with valves to ensure that the flow goes only in the direction of the arrow. The probability P_{ij} associated with any arc is to be thought of as the fraction of the liquid in reservoir i that will pass to reservoir j in one transition time unit.

One unit of liquid is poured into the system. After a while, a dynamic equilibrium is attained; the liquid continues to flow, but the amount in every reservoir remains constant. When this happens, the amount in each reservoir gives the steady-state probability for the corresponding state, they are proper probabilities because they are nonnegative and sum to unity.

The necessary condition for steady-state is that the flow into any reservoir must equal the flow out, because if the two were not equal, the amount in the reservoir would be changing. If this condition is met for all reservoirs, it is sufficient.

For any *i*,

Flow out = $\sum_j \pi P_{ij} = \pi_j \sum_j P_{ij} = \pi_{ij}$

Since $\sum_j P_{ij} = 1$

Flow in = $\sum_k \pi k P_{kj}$

Hence $\pi_i = \sum \pi k P_{ki}$ for all i, or, in matrix form $\pi = \pi P$

If we think of the liquid in terms of its molecules. A trajectory of an individual molecule describes a realization of the stochastic process. The effect of pouring in many molecules is to consider many realizations simultaneously. Hence, we are using in effect, a statistical mechanics approach. This technique is used in chemical diffusion models in electronics and elsewhere.

19. Conclusion

The steady state probabilities are determined exactly by the linear equations method and are approximated by the matrix multiplication approach. The transition diagram method is applicable to objects that can flow through reservoir and pipes such as the liquid and gases.

20. Markov Processes

Consider a finite (or countably infinite) set of points $(t_0, t_1, ...t_n, t)$, $t_0 < t_1 < t_2 .... < t_n < t$ and t, t_r T (r = 1, 2, ...n). where T is the parameter space of the process $\{X(t)\}$. The dependence exhibited by the process $\{X(t)\}$, $t \in T$ is called "Markov - dependence" if the conditional distribution of X(t) for given values of $X(t_1)$, $X(t_2).....X(t_n)$ depends only on $X(t_n)$ which is the most recent known value of the process.

that is, if

$$\begin{aligned} P[X(t) \leq x \mid x(t_n) = x_n, x(t_{n-1}) = x_{n-1}, ...x(t_0) = x_0] \\ = P[x(t) \leq x \mid x(t_n) = x_n] \\ = F(x_n, x: t_n, t) \qquad (2) \end{aligned}$$

The stochastic process exhibiting this property is called a 'Markov Process'. In a Markov process, therefore, if the state is known for any specific value of the time parameter t, that information is sufficient to predict the next behavior of the process beyond that point.

Continuous time stochastic process is similar in many respect to discrete time stochastic processes. However, complexity does occur because each infinitesimal time is available as a possible transition time.

As a consequence of the property (2), we have the following relation:

$$F(x_0, x: t_0, t) = \int_{y\in s} F\,(y, x, t, t)\; dF(x_0, y, t_0) \qquad (3)$$

where $t_0 < t < t$ and S is the state space of the process $x(t)$.

When the stochastic process has a discrete state space and a discrete parameter space, (2) and (3) take the following forms:

Using this property, for $m < r < n$ we get

$$P_{ij} = \sum_{k\in S} P_{ik}^{(m,r)} P_{kj}^{(r,n)} \qquad (4)$$

we have again used S as the state space of the process.

Equations (3) and (4) are called the "Chapman-Kolmogorov equations" for the process. These are basic equations in the study of Markov processes. They enable us to build a convenient relationship for the transition probabilities between any points in T at which the process exhibits the property of Markov – dependence.

Above is the Chapman-Kolmogorov equation in general form, we shall use it in the special form of

$$P_{ij}(t+\Delta t) = \sum P_{ik}(t)\, P_{kj}(\Delta t)$$

This equation requires the Markov assumption to permit a multiplication of the probabilities referring to events during t and to events during Δt. it also requires stationary to permit use of the same probability functions for the interval t and for the later interval Δt.

$P_{ij}(t)$ functions are probabilities, for all t. they are non negative, bounded functions because they must lie between 0 and 1.

$$\frac{dP_{ij}(t)}{dt} = \sum_k P_{ik}(t)\lambda_{kj}$$

The result is an exact (not approximate) differential equation for $P_{ij}(t)$ in terms of the $P_{ik}(t)$. It is a linear, first – order differential equation with constant coefficients λ_{kj}'s.

Recognising the above sum as matrix multiplication, we may express all of the differential equation at once in the matrix form

$$\frac{dP(t)}{dt} = P(t)A$$

Where dP(t)/dt is the matrix whose $(i,j)^{th}$ element is $dP_{ij}(t)/dt$. P(t) is the matrix whose $(i,j)^{th}$ element is $P_{ij}(t)$, and A is the matrix whose $(i,j)^{th}$ element is λ_{ij}.

The elements of A may be further related by extending the properties of P(t); so that for each i, each row of A must sum to zero. Since off-diagonal element is non negative, the diagonal element λ_{ii}, must be equal in magnitude and opposite in sign to the sum of others in the same rows. That is

$$\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$$

Howard (1960)

21. Semi-Markov Processes

A semi – Markov is a process in which changes of state occur according to a Markov chain and which the time interval between two successive transitions is a random variable whose distribution may depend on the state from which the transition takes place.

22. Waiting Time/ Holding Time

Let the time spent by the process in state j before its next transition, given that the next transition is state K be a random variable T_{jk} having distribution function

$W_{jk}(t) = P\,[T_{jk} \leq t] = P\,(t_{n+1} - t_n \leq t / X_n = j, X_{n+1} = K)$ j, K = 1, 2,.....m. The random variable T_{jk} called the Surjourn time or waiting depends on the state X_n being visited and the state X_{n+1} to be entered in the very next transition Iwunor (2001).

23. Interval Transition Probability Matrix

For all i, j and for $t \geq 0$

$$\varphi_{ij}(t) = \delta_{ij}\, hi(t) + \sum_{K} P_{Kj} \int_{0}^{t} f_{ik}(x) \varphi_{kj}(t-x)_{dx}$$

$$\begin{aligned} where\ hi(t) \quad &= 1 - \sum_{k} \varphi_{ik}(t) \\ &= 1 - W_i(t) \\ &= P(T_i > t) \end{aligned}$$

$$and\ \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = J \end{cases}\ is\ the\ Kronec\text{ker}'s\ delta\ function.$$

In a summary, suppose that a process can be in any one of N states 1, 2,.........N and that each time it enters state i, it remains there for a random amount of time having mean M_i and then makes a transition into state j with probability P_{ij}. Such a process is called a semi – Markov process. We note that if the amount of time that the process spends in each state before making a transition is identically 1, then the semi – Markov process is just a Markov chain. Thus a Markov process is a semi – Markov process but the converse is not true Iwunor (2001) and Howard (1971).

24. Markov Decision Processes

Bhat in 1984 summarizes the definition of Markov decision processes thus; Markov decision processes bring together the study of sequential decision problems of statistics, and the dynamic programming technique of applied mathematics and operations research.

Consider a process that is observed at discrete time points to be in any one of m possible states, which we numbered by 1,2,3,...m. After observing the state of the process, an action must be chosen, and we let D, denote the set of all possible actions, we assume D is finite.

If the process is in state i at time n and action k is chosen, then the next state of the system is determined according to the transition probabilities P_{ij}.

Following Ross (1989), let X_n denote the state of the process at time n and K_n the action chosen at time n, then the above is equivalent to stating that

$$P(X_{n+1}=j/X_0,K_0,X_1,K_1,.....X_n=i,K_n=k) = {}^{K}p_{ij}$$

Thus the transition probabilities are dependent on the present state and subsequent action.

Definition (8): A policy - by a policy we mean a rule for choosing actions. A policy is a sequence of decisions, one for each state of the process.

Definition (9): Dynamic programming is an approach for optimizing multistage decision processes. It is based on Bellman's principle of optimality.

25. Bellman's Principle of Optimality

An optimal policy has the property that regardless of the decisions taken to enter a particular state in a particular stage, the remaining decisions must constitute an optimal policy for leaving the state, Bellman (1957).

Consider a Markov chain with state space S. Suppose with every state we associate a decision to be chosen out of a set D which depend on ${}^{K}P_{ij}$ be the probability of the transition epoch.

Let ${}^{K}P_{ij}$ be the probability of the one step transition from i to j. i,j $\in S$ under decision $K \in D$.

Also we associate a reward ${}^{K}R_{ij}$ with decision K and transition i to j . Knowing the set of alternatives in the decision set and the corresponding transition probabilities and rewards, the objective of the process is to select the optimal decision under certain criteria. When we associate rewards with every decision, maximization of expected reward over a given time horizon is the natural criterion.

If cost are associated with decisions; costs are essentially negative rewards and so minimization of expected costs is called for.

Let ${}^{K}v_i^{(n)}$ be the expected total earnings in n future transitions if decision K is made when the process is in state i. For the optimal decision K= 0 if it exists; we have

$$^0V_i^{(n)} = \max_{k\in D}\sum_{j\in S} {}^kP_{ij}\left[{}^kR_{ij} + {}^0V_j^{(n-1)}\right],\ n = 1, 2, 3 \ldots, i \in S$$

This is a functional equation satisfied by the expected reward.

26, Markov Reward Processes

Consider an aperiodic, irreducible Markov chain with m state ($m < \infty$ and a transition probability matrix. With every transition i to j associate a reward R_{ij}. If we let $V_i^{(n)}$ be the expected total earnings (reward) in the next n transitions, given that the system is in state i at present. A simple recurrence relation has been shown considering the transition probability matrix P and the reward matrix R as given.

Instead, suppose that the decision maker has other alternatives and so is able to alter elements of P and R. To incorporate this feature, define D as the decision set, and under decision $K \in D$, let $^KP_{ij}$ and $^KR_{ij}$ be the probability of the transition from i to j and the corresponding reward respectively for $^Kv_i^{(n)}$. The expected total earnings in n transitions under decision K; we have the recurrence relations (K=0 represents the optimal decision)

$$^0V_i^{(n)} = \max_{k\in D}\sum_{j=1}^{m} {}^kP_{ij}\left[{}^kR_{ij} + {}^0V_j^{(n-1)}\right]; n = 1,2,\ldots, i = 1,2,\ldots m$$

giving

$$^0V_i^{(n)} = \max_{k\in D}\left[{}^kQ_i + \sum_{j=1}^{m} {}^kP_{i_j}\, {}^0V_j^{(n-1)}\right] . i = 1,2,\ldots m; n = 1,2$$

where we have written $\sum_{j=1}^{m} P_{ij}\, {}^kR_{ij} = {}^kQ_j$

Recursive relation gives an iteration procedure to determine the optimum decision $di^{(n)} \in D$. For $i = 1, 2 \ldots m$ and $n = 1, 2 \ldots$

This is a Standard technique in Dynamic programming and it has been shown Bhat (1984) that this iteration process will converge on best alternative for each state as $n \to \infty$.

Since the procedure is based on the value of the policy (total earning) for any n, it is called the Value Iteration Method (VIM).

The method is based on recursively determining the optimum policy for every n, that would give the maximum value.

One major drawback of the method is that, there is no way to say when the policy converges into a stable policy; therefore, the value iteration procedure is useful only when n, is fairly small.

27. Markov Decision Processes and Linear Programming

A Mathematical program is considered to be a standard linear program if it is of the form.

Minimize $Z = CX$

Subject to

$$Ax = b$$

$$X \geq 0$$

where

$$\underset{(mxn)}{A} = \begin{bmatrix} a_{11} & a_{12} & . & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & . & a_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & . & .a_{nn} \end{bmatrix}, \underset{(mx1)}{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}, \underset{(mx1)}{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_m \end{bmatrix}$$

$C \quad = C_1, C_2, \ldots C_{n]},$ *Abubakar (2005c)*

Linear programming is an optimization technique. It receives so much attention in recent years due to the availability of the methods of solution to the general linear programming problems involving large variables Diego and German (2006) and Abubakar (2005). Linear programming formulation of Markov Decision processes has been reported also in Diego and German (2006) and Tijm (1988).

According to Kurkani (1999), Puterman (1994), Goto, *et al* (2004), Hillier and Lieberman (1980); we consider a Discrete Time Markov Chain (DTMC) $\{X_n, n = 0, 1, \ldots\}$, whose transition probability matrix depends on the action taken A_n. Additionally, the system incurs a cost $c(i, a)$ when an action a is chosen at some state i. Then the joint process $\{(X_n, A_n), n=0, 1, \ldots\}$, is a Discrete Time Markov Decision Process (DTMDP).

The policy-iteration algorithm solves the following average cost optimality equation in a finite number of steps by generating a sequence of improving policies.

It was observed in Abubakar (2011b) that the finite convergence of the policy-iteration algorithm implies that numbers g^* and v^*_i, $i \in I$, exist which satisfy the average cost optimality equation

$$v_i^* = \min_{a \in A(i)} \left\{c_i(a) - g^* + \sum_{j \in I} p_{ij}(a) v_j^*\right\}, i \in I \qquad (5)$$

I is the set of states. The constant g^* is uniquely determined as the minimum average cost per unit time, that is

$$g^* = \min_R g(R)$$

Moreover, each stationary policy R^* such that the action R^*_i minimizes the right side of (5) for all $i \in I$ is average cost optimal Tijm (1988).

Moreover, each stationary policy R^* such that the action R_i^* minimizes the right side of (5) for all $i \in I$ is average cost optimal Tijm (1988).

Another convenient way of solving the optimality equation is the application of a linear programming formulation for the average cost case.

$$\text{Min } \sum_{i \varepsilon S} \pi_i \sum_{a \varepsilon A(i)} f(i,a)c(i,a)$$

subject to

$$\pi_j = \sum_{a \varepsilon A(i)} \pi_i p_{ij}\ (a) \quad j \in S \quad \text{Balance equation}$$

$$\sum_{a \varepsilon A(i)} f(i,a) = 1$$ Normalization equation, S is the set of all allowable states.

This model is not linear. But if we define new decision variable, $x_{ia} = \pi_i\, f(i,a), i \in S, a \in A(i)$, then we can build an equivalent linear model. The meaning of $x_{i\alpha}$ is the long run fraction of the time that the system is in state i and action α is chosen.

$$\text{Min } \sum_{i \varepsilon S} \sum_{a \varepsilon A(i)} c(i,a) x_{ia}$$

subject to

$$\sum_{a \varepsilon A(i)} x_{ja} - \sum_{i \in S^-(j)} \sum_{a \varepsilon A(i)} p_{ij}\ (a)\, x_{ia} = 0 \quad j \in S$$

$$\sum_{i \varepsilon S} \sum_{a \varepsilon A(i)} x_{ia} = 1$$

$$x_{ia} \geq 0, \qquad i \in S, a \in A(i)$$

Where $S^{-}(j)$ is the set of possible predecessors of state j. i.e.
$S^{-}(j) = \{i: j \in S(i,a) \text{ for some } a \in A\}$.

Denardo and Fox (1968) and Goto, et al (2004).

28. The Exponential Distribution Functions and the Application to Markov Models

In making mathematical models for a real- world phenomenon it is always necessary to make certain simplifying assumptions so as to render the mathematics tractable. One simplifying assumption that is often made when modeling with Markov principle is to assume that certain random variables are exponentially distributed. The reason for this is that exponential distribution is both relatively easy to work with and is often a good approximation to the actual distribution, Ross (1989). An important simplifying assumption in making Markov chain models is that the time it takes to make a transition (random variable) be described by negative-exponential distribution.

29. The Exponential Distribution

The probability density function of the random variable *T* having the exponential distribution is

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0, \lambda > 0 \\ 0 & t < 0 \end{cases}$$

Kohlas (1982). The distribution has λ as a parameter. λ also determines the shape of the distribution.

The mean is $1/\lambda$ and the variance is $(1/\lambda)^2$. Thus the mean and variance are not separately adjustable, as one may frequently desire.

Figure1 plots this function for three values of λ. Notice that the function intercepts the vertical axis at λ, that it diminishes monotonically to zero (asymptotically), and that the rate of convergence is proportional to λ. The total area under the curve is, of course, always equal to 1, as it must be for any density function.

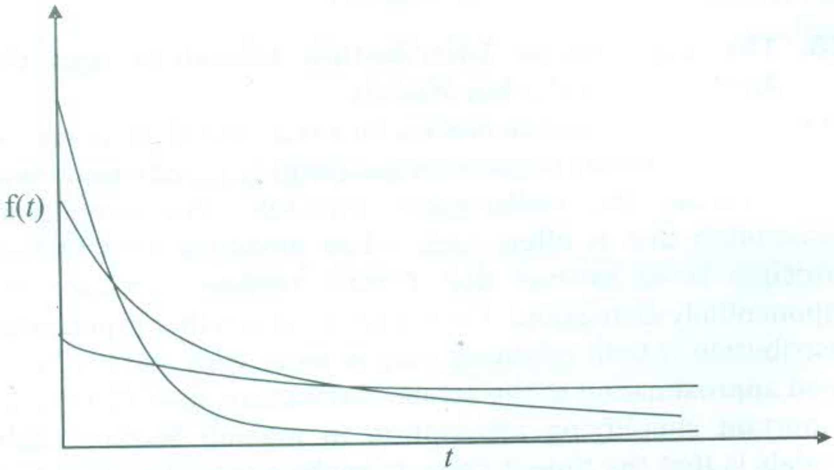


Figure1: The graph of Exponential

Most applications are based on its 'memory-less' property, when the measurement variable T has a time dimension. This property refers to the phenomenon in which the history of the past events does not influence the probability of occurrence of present or future events.

According to Ross (1989) A random variable X is said to be without memory, or memory-less, if

$$P\{X > s + t | X > t\} = p\{X > s\} \text{ for all } s, t \geq 0$$

Following Feller (1971), " We shall refer to this lack of memory as the Markov property of the exponential distribution".

30. The gamma distribution

The density $f(t)$ given by

$$f(t) = \frac{\lambda^{\alpha} t^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda t} \quad t \geq 0,$$

where the shape parameter α and the scale parameter λ are both positive. Here $\Gamma(\alpha)$ is the complete gamma function defined by $\Gamma(\propto) = \int_0^{\alpha} e^{-t} t^{\alpha-1} dt \qquad \alpha > 0,$

And having the property $\Gamma(\propto + 1) = \propto \Gamma(\propto)$ for any $\propto > 0$. The probability distribution function
F(t) may be written as

$$F(t) = \frac{1}{\Gamma(\propto)} \int_0^{\lambda t} e^{-u} u^{\alpha-1} du \quad t \geq 0.$$

The latter integral is known as the incomplete gamma function. If the shape parameter $\propto$ is a positive integer k, the gamma distribution is the well-known Erlang-K (E_k) distribution.

The Erlang-K distribution has a very useful interpretation. A random variable with an Erlang-K distribution can be represented as the sum of k independent random variables having a common exponential distribution.

31. The lognormal distribution

The density f(t)is given by

$$f(t) = \frac{1}{\alpha t \sqrt{2\Pi}} e^{-\{\ln(t)-\lambda\}^2/2\alpha^2}, \quad t > 0,$$

where the shape parameter α is a positive and the scale parameter λ may assume each real value. The probability distribution function F(t) equals

$$F(t) = \varphi\left(\frac{\ln(t) - \lambda}{\propto}\right), \quad t > 0,$$

Thus a unique lognormal distribution can be fitted to each positive random variable with given first two moments.

32. Weibull distribution

The density f(t) is given by

$$f(t) = \alpha\lambda(\lambda t)^{\alpha-1}e^{(\lambda t)^{\alpha}}, \quad t > 0,$$

With the shape parameter $\propto > 0$ and scale parameter $\lambda > 0$. It is observed that the Weibull distribution reduces to the exponential distribution when the shape parameter is unity.

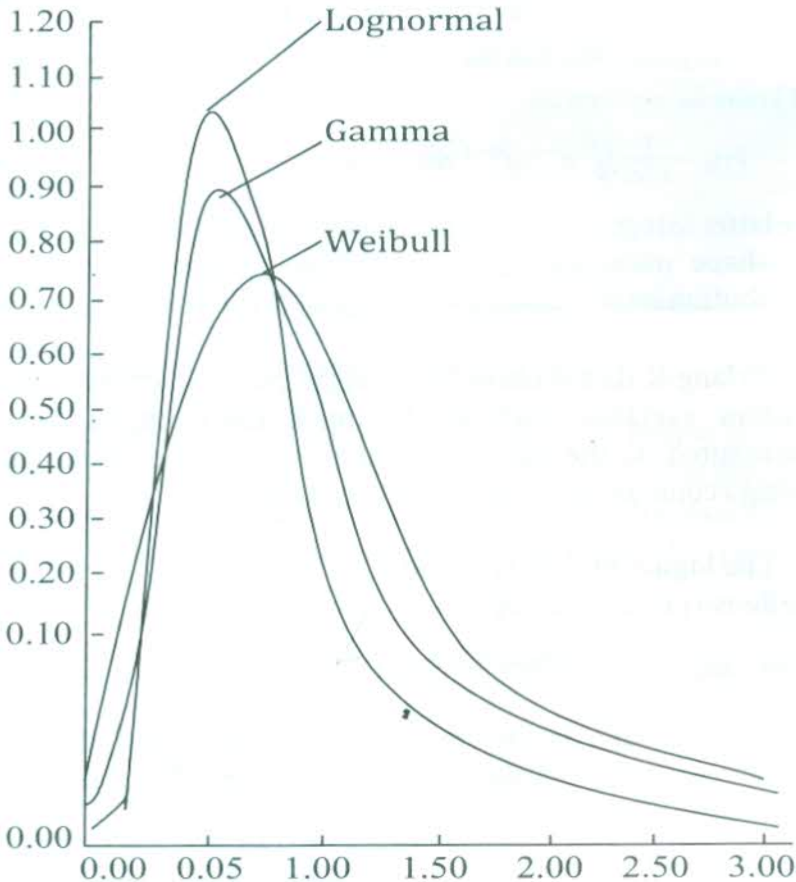


Fig. 2: The gamma, lognormal and weibull densities with E(x) = 1 and $c_x^2 = 0.25$.

Source: Tijms (1988).

It is observed that the Markov and stationary assumptions imply that the times between events must be negative-exponentially distributed. The parameters of these distributions, the λ_{ij}'s may be dependent on the state occupied, *i*, and the next state, *j*, but all of the distributions must be of the negative exponential form. No other distribution family can even be considered as a candidate for describing the times between events.

It was mentioned that in many applications, the times between events are most naturally conceived of as having a density function of the general form shown in Fig. 2 (perhaps a gamma or weibull or lognormal). That is, one tends to think in terms of some nominal value, the mean, plus or minus some relatively minor variation. Or, put another ways, the most likely values are considered to be clustered about the mean, and large deviations from the mean are viewed as increasingly unlike. However, the form of the negative exponential density functions implies that the most likely times are close to zero, and very long times are increasingly unlikely. If this characteristic of the negative exponential distribution seems incompatible with the application one has in mind, perhaps a Markov model is inappropriate.

This is an important understanding to be able to distinguish between those processes which might properly be modeled as stationary Markov process and those which should not, Abubakar (2011a).

33. Practical Applications by Me

The followings are some of my applications and contributions to the study of Markov Processes.

A Study of Leprosy Disease and Human Health Condition

33a. Definition of Leprosy

Many definitions of leprosy exist, but Hunter, *et al* (1966) defined leprosy as a chronic infectious disease primarily of the skin and nerves caused by *Mycobacterium leprae*. It is one of the least infections of all the infectious diseases. The incubation period varies from less than a year to many years, but probably averages three to five years.

33b. Types of Leprosy

Several variants of the disease are demonstratable, but the disease can be divided generally into two polar types; tuberculoid and lepromatous. A transitional or demorphous type may show a variable degree of similarity to the tuberculoid or the lepromatous types depending upon which pole it approximates.

33c. The Leprosy Model

We considered a leprosy patient. Let us assume that each year the leprosy patient is under treatment or has recovered from the disease or has relapsed or has died from the disease.

We therefore have a four state process thus:
State 1 - Under treatment, State 2 - Recovery, State 3 - Relapse, State 4 - Death due to leprosy.

These states are assumed to be mutually exclusive and exhaustive. The transition from one state to another is indicated in the transition diagram shown in figure 3.

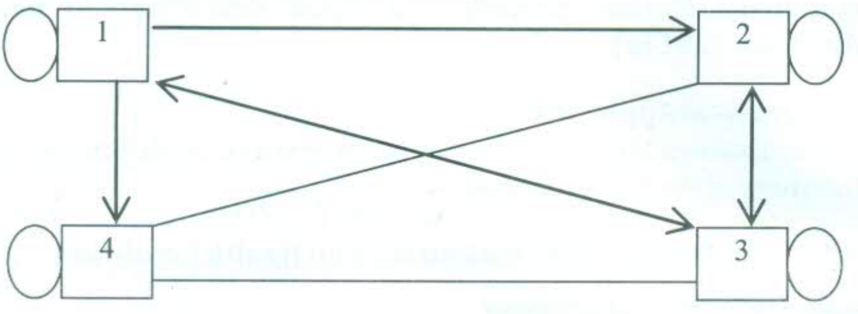


Figure 3: Transition diagram for leprosy

We observe that states 1, 2 and 3 are transient states and state 4 is an absorbing state. In other words, all possible transitions of the process are made between states 1, 2 and 3 but once a transition is made to state 4 the process terminates. We would like a transition to occur at a time the duration of stay in a state is completed, even if the new state is the same as the old. Such a transition is called virtual transition, and is represented by loops in the transition diagram.

From the above transition diagram we can record the transition probability matrix 'P' for the process as shown below:

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ O & P_{22} & P_{23} & O \\ P_{31} & P_{32} & P_{33} & O \\ O & O & O & P_{44} \end{bmatrix}$$

Let P_{ij} be the probability that the leprosy patient who is in state 'i' on his last transition will enter state 'j' on his next transition, i, j =1, 2, 3, 4. The transition probabilities must satisfy the following:

$$P_{ij} \geq 0, i,j = 1, 2, 3, 4.$$

and $\sum_{j=1}^{4} P_{ij} = 1, \quad i = 1, \; 2, \; 3, \; 4$

Whenever the patient enters state 'i' he remains there for a time T_{ij} in state i before making a transition to state 'j'. T_{ij} is called the 'holding time' in state i. The holding times are positive integer valued random variables each governed by a probability distribution function f_{ij} ( ) called the holding time distribution function for a transition from state i to state j.

Thus $P(T_{ij} = m) = f_{ij}(m)$. m = 1, 2, 3,…
i,j = 1, 2, 3, 4.

We assume that the means μ_{ij} of all holding time distribution are finite and that all holding times are at least one year in length. That is,

$$f_{ij}(0) = 0$$

To completely describe the semi-Markov process we must specify four holding time distribution functions in addition to the transition probabilities. For a fixed value of i T_{ij} is the same for each value of j, $(i, j = 1, 2, 3, 4)$.

Figure 4 shows a portion of a possible trajectory for the leprosy patient.

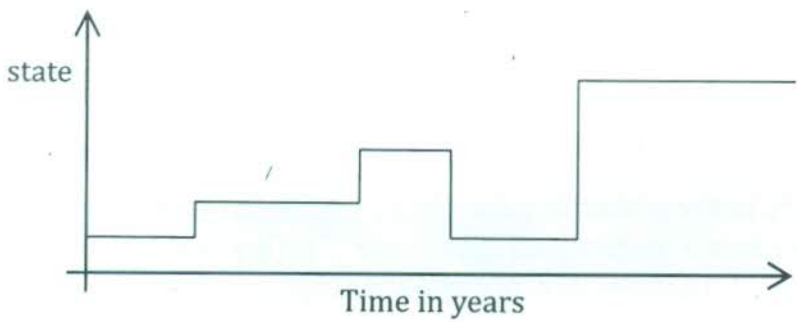


Figure 4: A possible trajectory for the process

33d. Effectiveness of the Treatment

When the leprosy patient undergoes treatment, it is expected that the treatment should have an effect on the disease. This effect should be noticed in the increase in probability of recovery, a decrease in the probabilities of death and having a relapse. An appropriate measure of this treatment effectiveness is obtained from the following expressions.

$$E_{12} = (1 + k)P_{12}$$

$$E_{1j} = (1 - \mathrm{k})P_{1j}, \ j = 3, 4$$

where k is a positive real number in the interval $[0, 1)$. Then

$$E_{11} = 1 - \sum_{j=2}^{4} E_{1j}$$

and the transition matrix is P with the first row replaced by E_{1j}, J = 1, 2, 3, 4, Abubakar (1995).

PROBLEM 1: A 10-year follow-up of a 6-week quadruple drug regimen for 136 multibacillary leprosy patients showed 13% relapses, seven late and the remaining (111) were considered to have recovered from leprosy, Pattyn, *et al* (2002).

The above is considered to be the situation at present; it is of interest therefore to determine recovery, relapse and death of Leprosy in the near/ distant future.

33e. Results

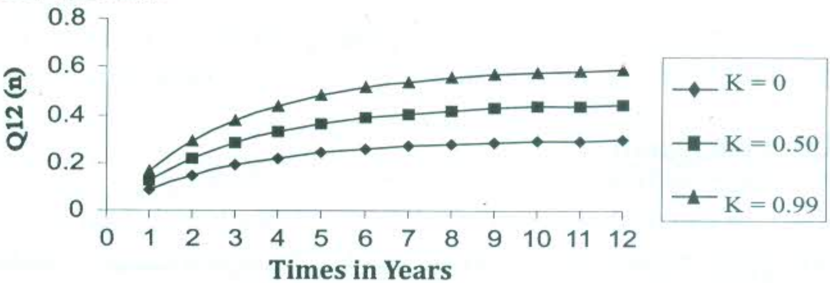


Figure 5: The Probability of being in state 2 having started from state 1

$Q_{12}(n)$ is the probability of recovery from leprosy at time n given that the patient started treatment at time zero.

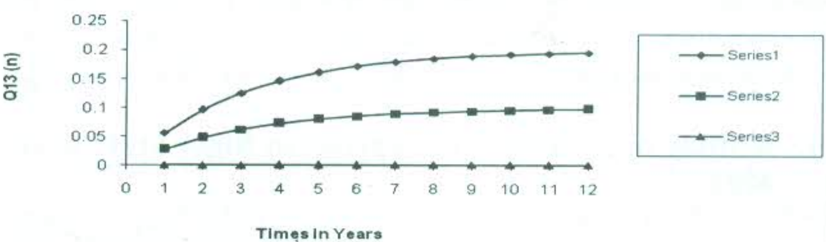


Figure 6: The Probability of being in state 3 having started from state 1

This is the probability that a leprosy patient will be a relapse in year *n* given that the patient was under treatment in year zero. It also represents the probability that a leprosy patient will develop a resistance to the drug.

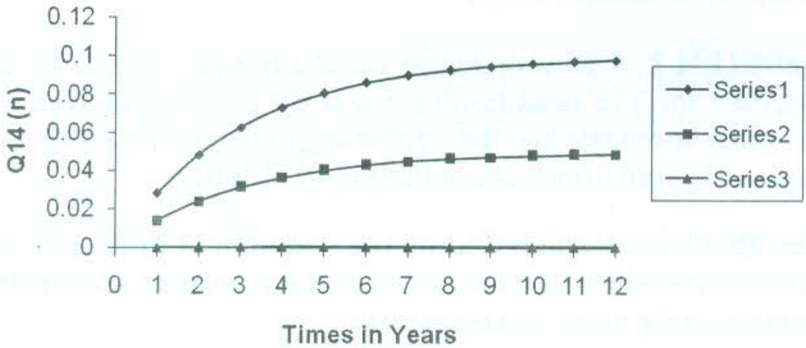


Figure 7: The Probability of being in state 4 having started from state 1

$Q_{14}(n)$ is the probability that a leprosy patient will die in year n given that the patient was under treatment in year zero.

33f. Conclusion

We observed from the graphs that when the treatment is 99% effective, the probabilities of relapse and death from leprosy has been reduced to zero and consequently increased the probability of recovery to unity.

Thus, we wish to submit that the Semi-Markov model could be used as a predictive device to study leprosy conditions. Such predictions could be useful to the government and non-governmental organization for the management of resources for the control of leprosy disease, Abubakar (1995, 2004a, 2007).

34. A Study of Disability in Leprosy in Niger State project Area

Let state 1 be a new case of leprosy; A case of leprosy is a person showing clinical signs of leprosy, with or without bacteriological confirmation of the diagnosis, and requiring MDT and has never been treated previously with anti-leprosy chemotherapy.

Let State 2 be a new case who has undergone a disability assessment with a WHO disability grade 1. (Hands and feet; anaesthesia present, no visible deformity or damage present. Eyes: eye problems due to leprosy present but vision not severely affected as a result, (vision 6/60 or better; ability to count fingers at 6 meters))

State 3 is a new case who has undergone a disability assessment with a WHO disability grade 2 (Hands and feet; visible deformity or damage present. Eyes: severe visual impairment. (vision worse than 6/60; inability to count fingers at 6 meters) WHO (1997) and ILEP (1998). We assumed that all the states communicate.

PROBLEM 2

The following is a summary statistics for WHO disability grade1 and grade2 in leprosy in Niger state project area for a period of nine years.

Table 2: WHO disability grade1 and grade2 in Niger state project area

Year	1993	1994	1995	1996	1997	1998	1999	2000	2001
New cases	58	344	570	324	452	83	56	79	158
Disability Grade l	10	193	342	272	348	78	47	62	136
Disability Grade 2	48	151	228	52	104	5	9	17	22

Given that the present conditions of disabilities in Leprosy are as shown in the table, it is important to investigate these conditions in the near future/long run.

34a., Results

We obtain the interval transition probability illustrated in figure 8.

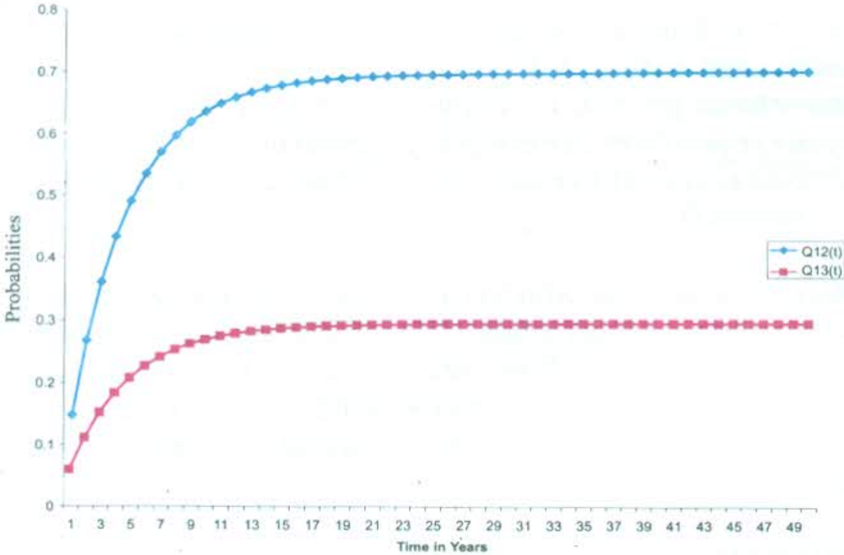


Figure 8: The Graph of Probabilities of Disability in Leprosy

34b. Conclusion

It could be seen from the graph that the probabilities of obtaining a patient with a disability grade1 attained an equilibrium of 0.7 in eighteen years and that of grade 2 stabilized at 0.3 in fifteen years. Thus, leprosy patients with grade 1 and grade 2 disability are not likely to exceed 70% and 30% in the next eighteen years and fifteen years, respectively. In view of the fact that the duration of treatment and types of drugs for the treatment of leprosy are dependent on the disability grade; this result could be useful to the government and Non-Governmental agencies in the planning of resources for the control of leprosy in Niger State, Abubakar (2005b).

35, A Discrete Time and State Case of A Study of Disability in Leprosy in Niger State project Area

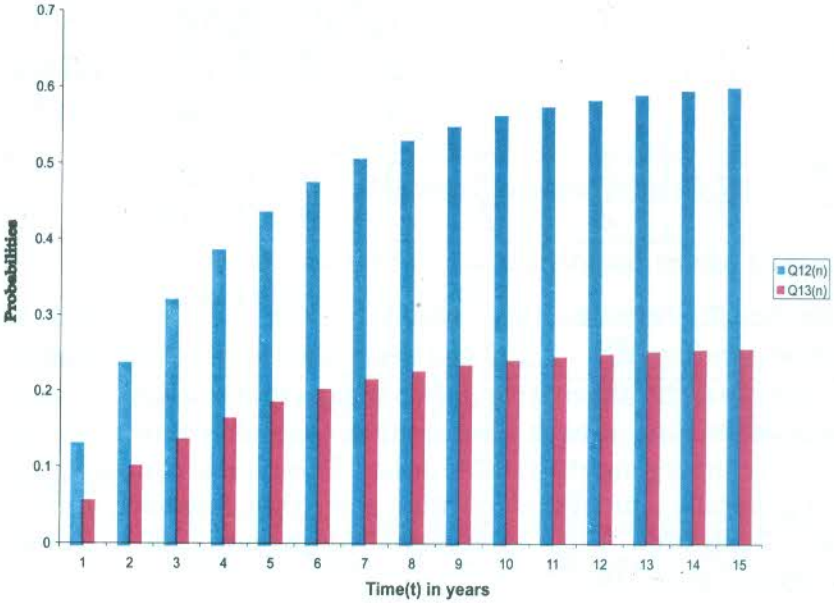


Figure 9: The Graph of Probabilities of Disability in Leprosy

35a. Conclusion

The result shows that the new cases of leprosy that can be identified to be of disability grade 1 and grade 2 will increase steadily and then attains a stable probability of 0.6 and 0.25 in fifteen years and twelve years respectively, Abubakar (2005c).

36. A Study of the Cost of Treatment of Leprosy Disease

PROBLEM 3

We determine the best alternative in terms of drug-type and a possible minimal cost in the treatment of leprosy disease.

Results

The result is presented in table 4

Table 3. A Summary result of the Optimal Policies and Costs

n	$d_1^{(n)}$	$d_2^{(n)}$	$d_3^{(n)}$	$^oV_1^{(n)}$	$^oV_2^{(n)}$	$^oV_3^{(n)}$
1	2	1	2	1,200	2,300	2,400
2	2	2	2	2,850	4,450	4,160
3	2	2	2	4,340	6,260	6,020
4	2	2	2	5,900	7,880	7,680
5	2	2	2	7,480	9,470	9,270

The results revealed that except for the $d_2^{(1)} = 1$ with the corresponding $^oV_2^{(1)} = 2{,}300$, the best policy for the other states at each time is the alternative 2. This means that the best policy for each other state at each time is the second alternative. This is a kind of convergence to a stable policy. This type of convergence is not generally true of this iterative algorithm, however, it makes this illustration very interesting Abubakar (2007, 2004a), Abubakar, *et al* (2007).

37. A study of Blood Inventory in General Hospital Minna

General Hospital Minna is a typical city hospital in Niger state Nigeria. It is a government owned medical out- fit that provides health serves to the populace at a subsidized price. Blood bank management is one of the important services rendered in the Hospital and it is highly patronized.

A major complication to the blood donation and transfusion is the existence of different blood types among humans and the matching. The eight types of blood that exist in human are: A+, A-, B+, B-, AB+, AB-, O+ and O-. These blood groups have complex substitutability pattern and could be regrouped into O, A, B, AB, Jevwia (2014).

Let the state space of the process be represented as follows:

State1: Blood group O, State 2: Blood group A, State 3: Blood group B, and State 4: Blood group AB.

It is assumed that the states are mutually exclusive and exhaustive.

A summary statistics for 4 years for the monthly transfused blood in the General Hospital Minna is presented in table 3.

Table 4: A frequency table of blood transfused in the General Hospital Minna between 2010 -2013

State	1	2	3	4
Frequency	3968	280	472	116

PROBLEM 4:

It is assumed that the probable supply and demand of blood in the General Hospital Minna (GHMX) changes in two major periods from January to August with a low supply/demand but with a high supply/demand from September to December of the same year. It is important to establish the above assertion and also to determine the probable transfusions of various bloods in the Hospital in the future.

37a. Results

The null hypothesis ($H_{(0)}$ of constant transition probability matrix is therefore rejected and we have non-homogenous Markov chain model. That is, the demand/supply of blood changes in the two periods.

The result indicates that blood group O is being transfused to blood group O, A, B, AB with probabilities of 0.81, 0.06, 0.11 and 0.02 respectively in January to August. This is in contrast to the probabilities of 0.75, 0.08, 0.13 and 0.04 respectively in September to December in the same year.

These set of probabilities indicates that the blood transfused from blood group O to blood groups A, B, and AB are about 0.4, 0.2 and 0.9 percent respectively. The result shows that the reception of blood by blood group AB from blood group A and blood group B is about 1.7% and 0.9%, respectively.

The result also indicates that there will be little variation in the blood needed for transfusion in the future on the basis of the present Jevwia (2014) and Abubakar *et al* (2014).

37b. Conclusion

The information about the blood needed for transfusion and blood bank inventory are very important for a successful health care delivery. If such information is available in every city hospital, it may be possible to move blood from one hospital to another during emergency.

38. A Study of Human Daily Health Condition

Suppose that at the beginning of each day the health condition of a man is observed and classified as good health or poor health. If he is found to have poor health, he is given either a first aid/preventive treatment or curative treatment so that the health condition could change to good health and could attend to his usual activities.

Suppose also that he could be found in good health conditions i = 1,2,...N. The good health condition i is better than i+1. That is the condition deteriorates in time. If the present condition is i and does not fall ill, then at the beginning of the next day then he has good health conditions j with probability P_{ij}. It is assumed that his body cannot improve on its own. That is $p_{ij} = 0$ for $j<i$ so that $\sum_{j \geq i} P_{ij} = 1$.

Let the health condition i = N represents a poor condition that requires treatment taking two days. For the intermediate states i with $1 < i < N$ there is a choice for him to preventively take treatment so that he could remain in good health condition for the present day. Let a first aid/preventive treatment takes only one day at most and a change from poor health to a good health (after treatment) has a good health condition i=1.

PROBLEM 5:

We wish to determine a rule which minimizes the long-term fraction of time the man will undergo treatment.

38a. Results

We obtain the minimum fraction of days that the staff is in a poor health condition equals 0.214; and to have assumed a cost of one unit for each time of treatment we therefore have that value as the average cost optimal for the treatment.

38b. Conclusion

The relative value associated with the policy obtained represent both the fraction of time in the long-run that the staff could be in a poor condition of health and perhaps absent from work, and the minimal cost incurred in the treatment. This could be determined for each staff so that for the staff whose value is a large contrast to the minimum of the other staff of the company could be considered as being in poor health condition quite often and therefore unproductive and may be retired. The cost obtained is not very realistic; it could be determined by other methods.

39. Linear Programming Approach to Markov Decision Model for Human Health Condition

PROBLEM 6:

There are three methods of solving Markov Decision problem;

the first is the Value iteration method used in problem 4. The second is the policy iteration method applied in problem 5. The third method is the linear programming formulation (of problem 5) and is the issue addressed in problem 6

According to Abubakar (2011b) the linear program has the optimal basic solution

$x_{10}^{*} = 0.6021,\ x_{20}^{*} = 0.1753,\ x_{31}^{*} = 0.0847,\ x_{41}^{*} = 0.0392,\ x_{51}^{*} = x_{61}^{*} = 0.0318$

This yields an average cost optimal policy R* = (0, 0, 1, 1, 1, 1)

with the minimal average cost $\sum_{i=2}^{N+1} x_{i1}^{*} = 0.206$, It is in agreement with the results obtained by the policy-iteration algorithm in problem 5.

39a. Conclusion

The policy-iteration formulation usually involves the writing of its own code. However, the two methods are very efficient and can be adopted for practical use by Doctors for the benefit of patients, Abubakar (2011c).

40. A Study of Desertification in Nigeria

Vegetation means the plant cover of the earth which includes trees and grasses of different kinds. Following Iloeje [1981] two broad belts of plant groups can be found in Nigeria (forest and savannah) and within each group it is possible to distinguish three sub-types; salt-water swamp, fresh-water swamp and high forest. The savannah comprises of guinea savannah, sudan savannah and sahel savannah. The transition between the last two grass regions is the subject of this study.

The area in desert is expanding, largely at the expense of grassland and cropland, Lester (2005).

An official federal government assessment placed the impact of desert encroachment as causing the visible sign of the shift in vegetation from grass and bushes, and in the final stages, expansive desert like sand. This same danger assessment estimated that between 50% and 75% of Bauchi, Borno, Gombe, Jigawa, Kano, Katsina, Kebbi, Sokoto, Yobe, Adamawa, and Zamfara states in Nigeria are being affected by desertification, Copyright [2003].

Acevedo *et al* [1995] have described and applied a correspondence between two major modelling approaches to forest dynamics: Transition Markovian models and Gap models or Jabowa-Foret type simulators. A transition model can be derived from a gap model by defining states on the basis of species, functional roles, vertical structure or other convenient cover types. A gap-size plot can be assigned to one state according to dominance of one of these cover types. A Semi-Markov framework is used for the transition model by considering not only the transition probabilities among the states, but also the holding times in each transition. Application in spatial are possible by considering collections of gap-size plots and the proportion of these plots occupied by each state.

The application of a semi-markov model to desert encroachment in the Northern part of Nigeria Abubakar (2005a), considered linear distance of land in kilometres; it is observed that linear distance is not very relevant to the land used for building and agriculture. This study therefore considered the land measured in square kilometres lost to desert in a year. Cost implication was considered using Markov reward structure similar to that reported in Abubakar *et al* (2007) in which value iteration method was used.

Following the classification of the vegetation of Nigeria Iloeje

(1981) presented in figure 10 and gap size model proposed by Acevedo *et al* (1995), we consider the following vegetation zones states for a Markov Process.

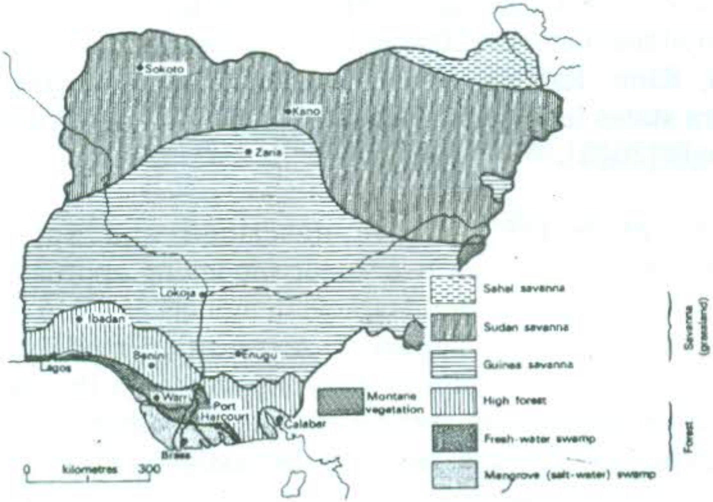


Figure 10: The Vegetation Belts in Nigeria

Source: Iloeje (1981

A possible transition between the states is presented in figure 11.

Figure11: A Transition diagram for the vegetation of Nigeria

PROBLEM 7:

Desertification and the associated persistent drought constitute the most serious environmental problem facing the Northern parts of Nigeria. The country is presently losing about 351,000 hectares of its landmass to desert conditions every year, which is advancing southwards at an estimated rate of 0.6 kilometers a year, (Government of Nigeria [2003]). Suppose that the cost of planting and taking care of the plants for the first year is estimated to be three million Naira and also one million and two hundred thousand Naira is required annually to maintain the vegetation. Our interest is to determine the rate of desertification in the long-run and also the probable cost of reclaiming a piece desert land (sahel) to sudan savannah in a year.

40a. Results

The result is presented in figure12

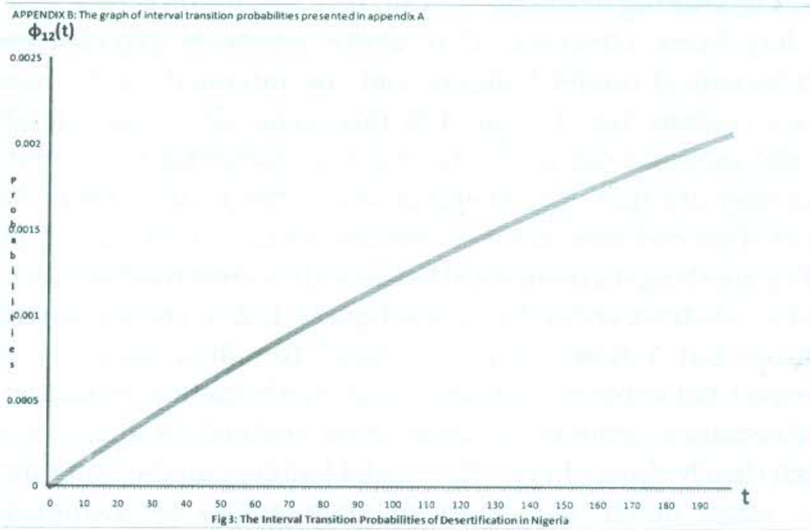


Figure12: The graph of interval transition probabilities

For the distance in kilometers, the result indicates a gradual loss of the Sudan savanna to desert encroachment with a probability of 0.00029 for the first year and converges to 0.0054 in about sixty years (about 55km in 60years), that is, about 0.92km in a year.

Further the result shows that some 92sq kilometres and 1,847sq kilometres of Nigeria cropland may be lost to desert in about 7 years and 192 years respectively. That is, nearly all the croplands of the front line states (Borno, Yobe, Sokoto and others may be claimed by desert.

The cost estimate shows that the initial cost of raising forest constitutes the significant amount and that the subsequent cost of maintenance may be secondary. Although, this is possible as long as there are no human disturbances such as farming, grazing and bush burning, Abubakar (2010, 2005a).

41. Concluding Remark

It has been observed that there exists a gap between mathematical model builders and the intended model users. Many reasons are attributed to this; generally many intended model users do not understand the fundamental mathematics and they are therefore skeptical about the results. Even then, most of the end users are non-mathematicians and are easily put-off by anything mathematical because they view mathematics as just an abstract entity. Of course figures 1, 2, 3, etc. are abstract entities but 1 tuber of yam, 3 men, 10million naira are not abstract but concrete entities. Thus, mathematical modeling is mathematics removed far away from abstraction and this has been clearly shown here. The model builders on the other hand do "water down" the rigors of mathematics by introducing appropriate 'simplifying assumption(s)'. The ideal practice is to take the midway between the two extremes and that is the basis for the development of the models discussed in these studies.

The models studied are therefore not dependent on data and therefore could be applied to study many other life processes with little or no modification. They are capable of generating relevant information that could be considered as very important ingredient for meaningful and purposeful planning for the future.

50 Recommendations

A call could therefore be made to the Government and Non-Governmental Organizations to extensively utilize the quantitative information emanating from scientific research particularly Mathematical models.

Further study is required to follow up and investigate the relationship between the predicted values and the observed values. This is possible only if the Government and stake holders could provide the necessary fund.

51, Acknowledgement

I am probably short of words to express my gratitude to the Almighty Allah; the Owner of every body and every thing. To Him belongs all glories and praises, (Subha naka wa bi Hamdika).

I will forever remain grateful to my late parents: Alhaji Abubakar Adamu and Hajia Hajara Abubakar. May Allah forgive and have mercy on you as you did bring me up.

To my Cousin Brothers; Alhaji Yusuf shaban Abubakar and Alhaji Saidu Abubakar, and of course all my sisters, I appreciate all of your financial, material and moral support for my education. Thank you very much.

I appreciate the supports (financial, material and moral) that I got from the big Daddy: Alhaji Abubakar Abubakar Bosso (Sarki Gabas, Minna). Thank you and the entire family that you provided for me.

My special thanks go to the Management(s) of the Federal University of Technology, Minna particularly the present administration under the leadership of Professor Musbau Adewunmi Akanji, FNSMBM, FAS. Thank you VC sir for among other things, you promoted me two times in your tenure. I appreciate the favor done to me by Professor M.A. Daniyan for giving me employment into this University. I experienced a rapid growth and development during the tenures of Professor Tukur H. Sa'ad and Professor M.S. Audu especially the latter, who appointed me as a Head of Department. I am very grateful to all of you.

I am highly indebted to all my Teachers from my primary school to my Academic Mentors in the University especially Dr. K.N. Korve (who retired recently as director of operations NHIS Abuja). He introduced me to Operations Research. Also Professor S.A. Reju (presently at University of Science and Technology, Namibia) who introduced me to classical Optimization. I give to you all a lot of thanks for doing so much to my life.

To the Dean of the School of Physical Sciences Professor A.S. Abubakar, I sincerely appreciate all that you have done for me. I am also very grateful to all the former Deans from Professor H. A. Suberu, Professor M. Galadima, Professor (Mrs) H.O. Akanya to Professor K.R. Adeboye that welcomed me to the then School of Science and Science Education.

I appreciate the contributions of my Senior Colleagues and Mentors in the Department of Mathematics including Professor K.R. Adeboye, Professor N.I. Akinwande, Professor Y.M. Aiyesimi, Professor Y.A. Yahaya, Dr. L.N. Ezeako and Dr. Audu Isah (HOD Statistics). I benefited immensely from your wealth of knowledge and experience. To other Colleagues Dr. D. Hakimi, Dr. A. Mohammed, Dr. A. Ndanusa, Dr. M. Jiya, Dr. M.D. Shehu, Dr. A.I. Enagi, Dr. M.U. Umar, Dr. (Mrs.) A.T. Cole, Dr. (Mrs.) R. Muhammed, Dr. N. Ngutor, Mr. Lawal Adamu, Mr. Festus Godfrey, Mr. O.R. Jimoh, Mr. S.O. Nasiru, Mr. Abubakar Samuel, Mr. Yusuf Adbulhakeem and finally the Head of Department Dr. R.O. Olayiwola, I am grateful to all of you for your kindness. I am also indebted to Dr. Yakubu Yisa, Dr. Usman Abubakar Zungeru and Mr. I. Onotu of Statistics Department.

I am grateful to Sheikh Nafiu Ibn Muazu of Sauki kahuta Zawiyya and Sheikh Mohammed Ibn Isah of Dutse Kura Gwari Zawiyya for your moral support and prayers.

I would like to appreciate the support that I receive from my beloved wife Hajia Ramatu Abubakar for taking care of me and our children. To my children, Abubakar D. Usman, Abdullahi U. Abubakar, Ahmad U. Abubakar and Yusuf U. Abubakar thank you for being good and obedient.

Finally, I am very grateful to the Chairman, Professor Y.A. Ayanwale and the members of the University Seminar and Colloquium Committee for making this lecture a huge success.

Thank you very Much.

REFERNCES

Abubakar U.Y., Hakimi D., Ndanusa A., Jevwia V.A. (2014), Analysis of Blood transfused in a City Hospital with the Principle of Markov Dependence. Applied Mathematical Sciences, Vol. 8, No.173, pp 8611-8621 Hikari Ltd, www.m-hikari.com

Abubakar U.Y. (2011a), The Exponential Distribution and The Application to Markov Models. Journal of Research in National Development, Vol. 9 No. 2a December, 29-35.

Abubakar, U.Y. (2011b), Markov Decision Model for Human Health with Policy Iteration. Journal of Science, Technology, Mathematics and Education (JOSTMED). Vol. 7 (2) pp 109-117.

Abubakar,U.Y. (2011c), Linear Programming Approach to Markov Decision Model For Human Health. Journal of Science, Technology, Mathematics and Education (JOSTMED). Vol. 7 (3) pp 191-197.

Abubakar, U.Y. (2010), A Stochastic Model with a Cost Structure and the Application to the Study of Desert in Nigeria. Journal of Mathematical Sciences. A Journal of National Mathematical Centre, Abuja, Nigeria. Vol. 1, No.1 pp 49-66.

Abubakar U.Y. Reju, S.A, Awojoyogbe B.O. (2007), Markov Decision Model and the Application to the Cost of Treatment of Leprosy Disease. Leonardo Journal of Sciences 11, pp 69-78.www ljs.acamicdiret.org.

Abubakar, U.Y. (2007), A Stochastic Model in Descrete States and

Descrete Time for the Control of Leprosy Disease. Leonardo Journal of Sciences, 11, July-December. Pp 51-60. www.ljs.academicdirect.org

Abubakar, U.Y. (2005a), Semi Markov Model and the Application to Desertification in Nigeria. Conference Proceedings September, 2005 Annual National Conference of the Mathematical Association of Nigeria Held in Awka, Anambra State Nigeria PP 69-78.

Abubakar U.Y. (2005b), Application of stochastic Model in Discrete States and Continuous Time to Disability in Leprosy in Niger State Project Area. Proceedings of the first annual Conference of the School of Science and Science Education (SSSE).Federal University of Technology, Minna Held on the 20th-23rd November, 2005. **Vol. 1** pp 48-52.

Abubakar, U.Y. (2005c), A computer Implementation of the Revised Simplex Algorithm. Abacus, The Journal of the Mathematical association of Nigeria, **32** No 2A 24 – 34.

Abubakar U.Y. (2005c), Application of Stochastic Model in Discrete States and Time to Study Disability in Leprosy Disease in Niger State Project Area. Proceedings of the first annual conference of the school of science and science education (SSSE), Federal University of Technology, Minna. Held on the 20th - 23rd November 2005. **Vol. 2** pp 14-19.

Abubakar U.Y. (2004a), The Application of Markov and Semi-Markov Models to the Control of Catarrh and Leprosy diseases, PhD Thesis (Unpublished), Federal University of Technology, Minna.

Abubakar,U.Y. (2004b),A Study of Relapses in Leprosy: A Semi-Markov Approach. Abacus, The Journal of the Mathematical Association of Nigeria, 31, No. 2B, 370-377.

Abubakar, U.Y. (1995), Semi-Markov Modelling for the Control of Leprosy Disease. Unpublished MSc. Thesis, Department of Mathematics University of Jos, Nigeria.

Acevado M.F. Urband D. L and Abiam M. (1995), Transition and Gap Models of Forest Dynamics, Ecological Applications 5(4) pp1040-1055.

Bailey N. T. J. (1975), "The Mathematical Theory of Infection Diseases and its application", second edition Griffin, London.

Bailey N. T. J. (1982), "The Biomathematics of Malaria", Griffin London.

Bellman R. (1957), Dynamic Programming. Princeton University Press, Princeton,.

Bhat U. N. (1984), Element of Applied Stochastic Processes, John Wiley New York.

Chung, K.L. (1967), Markov Chains with Stationary Transition Probabilities. 2nd ed. Springer-Verlag, Berlin.

Copyright 2003 Daily Trust (Abuja) Nigeria: Donor Agencies and the Threat of Desertification, January 8 Editorial.

Diego B., German R. (2006), Linear Programming Solvers for Markov Decision Processes. ww.sys.virgina.edu|sieds06|papers|FmorningSession5.1. Date accessed: Jan. 3rd 2011.

Feller, W. (1971). An Introduction to Probability Theory and its Applications. Vols. 1&11, 3rd ed, Wiley, New York.

Goto G.H., Lewis M.E. and Puterman M.L. (2004), "Coffee, tea, or...? A Markov Decision process Model for airline meal provisioning. Transportation Science, vol. 38 No.1 pp 107-118.

Government in action (2003), Overcoming environmental problems. January 13,

Hillier and Lieberman (1980), Introduction to operations Research, Holden Day.

Howard, R.A. (1971), "Dynamic probabilistic systems, Vols.1 & 2" John Wiley, New York.

Howard, R.A. (1960). Dynamic programming and Markov processes. The M.I.T. PRESS, Massachusetts.

Hunter/Frye/Swartzwelder (1966). A Manual of Tropical Medicine. W.B. Saunders Co. London

ILEP (1998) Medical Bulletin No 14 Operational Guidelines for the introduction of new MDT Regimens for the treatment of leprosy

Iloeje N. P. (1981), 'A New Geography of Nigeria', Longman Nigeria Limited.

Iwunor C.C.O. (2001), Markov and Semi-Markov Process, Research Oriented Course (ROC) in Probability and Stochastic Process. November 11-23, Lecture Notes. National Mathematical Centre, Abuja, Nigeria.

Jevwia V.A. (2014), Determining blood needed for transfusion in a general hospital with a Markov chain model. Unpublished M.Tech Thesis. Department of Mathematics/ Statistics, Federal University of Technology, Minna, Nigeria

John E. Freund, and Ronald E. Walpole (1980), Mathematical Statistics. Prentice/Hall International Edition, Third Edition

Kohlas, J. (1982). Stochastic Methods of Operations Research. Cambridge Univerity Press.

Korve K. N. (1993), "Models for the Management of Asthma" PhD Thesis, Faculty of Mathematical studies, University of Southampton England.

Kulkani V.G.(1999), Modelling, Analysis, Design, and Control of Stochastic System, Springer.

Lautedbacher C.J. and Stidham S. Jr. (1999), The Underlying Markov Decision Process in the single-leg airline yield Management problem. Transportation Science, vol. 33, No. 2 pp136-146.

Lester R. B. (2005), Outgrowing the earth: The Food Security Challenge in an Aged Falling Water Tables and Rising Temperature. W.W Norton and Co..., NY Nigeria

Pattyn S., and Grillone S. (2002). Relapse rates and a 10-year follow-up of a 6-week quadruple drug regimen for multibacillary leprosy. Leprosy Review 73, 245-247.

Puterman M. (1994), Markov Decision Processes: Discrete Stochastic Dynamic Programming, John Wiley New York.

Ross, S.M. (1989). Introduction to Probability Models. Academic Press, INC New York. Seymour L (1981), Schaum's Outline of Theory and Problems of Linear Algebra. McGraw-Hill, Singapore.

Stidham S. Jr. (2000), Optimal Control of Markov Chains; in Computational Probability. W.K.

Tijms H. C. (1988), Stochastic Modelling and Analysis: A Computational Approach. John Wiley & Sons Ltd. New York.

Uche P.I. (2001), Stochastic Models in Education and Manpower. Research Oriented Course (ROC) in Probability and Stochastic Processes. Unpublished Lecture notes. National Mathematical Centre, Abuja, Nigeria.

WHO (1997), WHO Expert Committee on leprosy. Seventh Report. Technical report Series 874.

A Brief Profile of the Inaugural Lecturer

Usman Yusuf Abubakar was born to the family of Alhaji Abubakar Adamu and Hajia Hajara Abubakar Adamu on the 10^{th} October 1961.

He attended St. Mary's Roman Catholic Primary School in Osomegbe Ekperi from 1969 to 1974; Arabic High School, Elekuro, Ibadan 1974/75; the School of Higher Islamic Studies, Shahunchi, Kano 1976/77. He later attended Nassarawa Teachers' Training College (NTTC), Kainji from 1977 to 1982; College of Education, Minna from 1982 to 1985 and Ahmadu Bello University, Zaria from 1987 to 1989.

He participated in the National Youth Service Scheme in 1989/90 in Bauchi State. He went for His Master's Programme in the University of Jos between 1991 and 1995; Postgraduate Diploma in Computer Science and PhD at the Federal University of Technology, Minna between 1995 and 1997; and 2000 to 2004, respectively.

Working Experience: He participated in the Niger State Voluntary Teaching Scheme (VTS) in 1986 and taught at Paiko Secondary School. He took up an appointment with Niger State Ministry of Education and served at Government Technical College, Kontagora in 1986/87. He lectured at the Niger State College of Advanced Studies (ZUCAS) now Niger State Polytechnic, Zungeru between 1991 and 1999, Federal Polytechnic, Bida in 2000 (January to May) and Federal University of Technology, Minna since June 2000 as an Assistant Lecturer. He was promoted to the rank of Professor of

Mathematics in 2015 and his research area is Stochastic Optimization.

He is actively involved in the Teaching and Research in the Department of Mathematics especially students project supervision at Undergraduate, Masters' and PhD level.

He has a sizeable number of credible publications in some reputable Local, National and International Journals.

He is married with children.