IDENTIFICATION OF TRAIT MARKERS USING MARKOV CHAIN PROPERTIES

By

OLARINOYE, Segun Bamidele

Matric. No. 93/037563

Thesis submitted to the Department of Statistics, Faculty of Physical Sciences in Partial Fulfillment of the Requirements for the Award of the Degree of Doctor of Philosophy in Statistics

Supervisor: Prof. E.T. Jolayemi

MARCH, 2021

CERTIFICATION

This is to certify that this Thesis is an original work carried out by Segun Bamidele OLARINOYE With matriculation number 93/037563 in the Department of Statistics, Faculty of Physical Sciences, University of Ilorin. The Thesis has been read and approved as meeting the requirements for the Award of Doctor of Philosophy in Statistics of University of Ilorin, Ilorin, Nigeria

.....
Prof. E. T. Jolayemi
(Supervisor)

30/03/2021

.....
Prof. A. O. Adejumo
(Head of Department)

30/03/2021

.....
Dr. O. Job
(Postgraduate Coordinator)

30/3/2021

.....
Prof. M. S. Dada
(Internal External Examiner)

30/03/2021

.....
Prof. T. O. Obilade
External Examiner

30/03/2021

DEDICATION

My late Father, Chief Mathew Olawale Olarinoye and my beloved late brother, Opeyemi Timothy Olarinoye

DECLARATION

I, OLARINOYE, Segun Bamidele hereby declare that this thesis entitled Identification of trait markers using Markov chain properties has neither been presented nor accepted in any previous application for a higher degree. All sources of information have been specifically acknowledged. In addition, the research work has been ethically approved by the University Ethical Review Committee.

OLARINOYE, S.B. **Date**

ACKNOWLEDGEMENTS

I am most grateful to God Almighty for His mercy and goodness in my life. My profound gratitude goes to my supervisor, Prof E.T. Jolayemi for his fatherly care, love and correction right from my undergraduate days, I am most grateful. I similarly, appreciate the efforts of Prof A.O. Adejumo (HOD), Prof P.A.Osanaiye, Prof B.A.Oyejola, Prof.G.M. Oyeyemi, Dr. N.A. Ikoba, Dr. O. Job and all other lecturers in the Department of Statistics.

I wish to express my profound gratitude and appreciation to my uncles Dr. Rasheed Adegboye, Prince A.D. Adegboye and Mr. S.T. Abolarin for their love and care during the course of the study.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. I was funded by the Tertiary Education Trust Fund(TETFUND).This was made possible through the assistance of my employer, Kwara State Polytechnic, Ilorin.

Lastly, I would like to thank my family for all their love and encouragement. To my parents who raised me with a love for science and supported me in all my pursuits, immense appreciation. I appreciate my lovely children: Akinola,Akinloye,Ibukunoluwa and Adedamola for their immeasurable support, understanding and love. And most of all , many thanks to my loving, supportive, encouraging, and patient wife, Motunrayo, whose faithful support from the beginning to the final stage of this Ph.D. is so appreciated.

Olarinoye,Segun Bamidele

University of Ilorin
Ilorin, Nigeria

ABSTRACT

Lineage is characterized by some traits whose genetic marker is passed from any of the parents' genes to their children. Some of these traits may be physical. For example, hair, eye and/or skin colours. Laboratory tests some of which are costly and non accessible have been used for screening such traits whereas a non-laboratory Markov chain methodology which is cheaper and faster was proposed as an alternative. The study aimed at establishing a non-laboratory procedure using Markov Chain to identify useful markers to link traits. The objectives were to: (i) identify traits, the state spaces and classify them using probability transition matrix; (ii) determine the properties of the state space for each trait; (iii) determine which of the traits are the markers; and (iv) demonstrate the use of the procedure using real life data.

The process of transmitting traits from one generation to another was conceptualized as a Markov Chain. Markov chain methodology was employed to examine presence or absence of a given trait in an individual. The corresponding transition probability Matrix was studied to identify properties of the state spaces. The nature of Markov chain was examined to establish ergodicity and obtain the mean recurrence time. A test of homogeneity of the Markov chain was developed when at least three successive time points existed. These properties were used to characterize traits of Albinism as a real life example.

The findings of the study were that:

i. Markov chain concepts and techniques were successfully applied to classify various states of the available traits;
ii. all transition probability matrices were found to be recurrent with different periodicity;
iii. a best marker was identified to have the shortest period when state spaces are recurrent; and
iv. in the case of real life data involving albinism, the colour of the eye was the best marker.

The study concluded that Markov Chain is a vital tool to identify appropriate trait markers as an alternative to laboratory test. The study recommended that where information is available a Markov chain can be used to identify trait markers.

TABLE OF CONTENT

CHAPTER THREE: METHODOLOGY

CHAPTER FOUR: VALIDATION OF RESULTS

CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

LIST OF TABLES

FIGURES

LIST OF ABBREVIATIONS AND NOTATIONS

AFR	AVERAGE FACIAL RATIO
DNA	DEOXYRIBONUCLEIC ACID
FISH	FAMILY IDENTIFICATION OF SEQUENCE
FKP	FINGER KNUKLE PRINT
GESSA	GEODESIC ENSEMBLE SURFACE SAMPLING ALGORITHM
GHMM	GENERALIZED HIDDEN MARKOV MODEL
GLIMMER	GENE LOCATOR AND INTERPOLATED MARKOV MODELER
HINT	HAPLOTYPE INFERENCE TOOL
HMM	HIDDEN MARKOV MODEL
IMM	INTERPOLATED MARKOV MODEL
IBD	IDENTITY-BY-DESCENT
LDA	LINEAR DISCRIMINANT ANALYSIS
MC	MARKOV CHAIN
MFCC	MEL-FREQUENCY CEPSTRAL COEFFICIENTS
OA	OCULA
OCA	OCULOCUTANEOUS
SAHMM	STRUCTURE-ANCHORED HIDDEN MARKOV MODEL
SNP	SINGLE-NUCLEOTIDE POLYMORPHISM
VQ	VECTOR QUANTIZATION
WWW	WORLD WIDE WEB