JOURNAL OF STATISTICS, OPTIMIZATION AND DATA SCIENCE
VOL: 1, ISSUE: 2, (DECEMBER 2023), PP: 10-21, ISSN: 2583-9462

Markov Modelling of T2DM Progression

Sarode Rekha [a], Tirupathi Rao Padi[b], and P. Sasikala[c]

[a] *Department of Statistics, Pondicherry University, Puducherry, India.*
[b] *Department of Statistics, Pondicherry University, Puducherry, India.*
[c] *Department of Statistics, Mount Carmel College (Autonomous), Bengaluru, India.*

ARTICLE HISTORY
Compiled January 26, 2024

ABSTRACT
This research focuses on applying Markov modeling to the progression of type 2 diabetes mellitus. The study involves constructing a transition probability matrix that represents various stages in the advancement of type 2 diabetes. By utilising this matrix, the probability distributions for the consecutive occurrences of the same state, looking one and two days ahead. Furthermore, the investigation formulates explicit mathematical expressions for different statistical measures, utilizing Pearson's coefficients. The developed model behaviour is examined through numerical examples and sensitivity analysis. The primary objective of this study is to create user-friendly tools, and developing appropriate software based on the derived mathematical formulations. The application of these findings can significantly enhance the management of type 2 diabetes in healthcare settings, potentially extending to decision support systems.

KEYWORDS
Type-2 Diabetes; Markov Model; Transition Probability Matrix; Disease Progression.

1. Introduction

Type 2 diabetes is a chronic metabolic disorder and one of the most common non-contagious diseases on the rise worldwide. Diabetes is the world's most prominent disease after cancer. There were two basic categories of diabetes. Type 1 Diabetes is mostly observed in early age and childhood onset. However, type 2 diabetes may be due to insulin insufficiency and insulin resistance. Hypoglycemia and hyperglycemia are two major causes where we can classify the diabetes problem into two categories. In hyperglycemia, the normal glucose levels are at a random point of time shall be above 140 mg/dl. Whereas, in hypoglycemia, the blood sugar levels at random point of time shall be less than 70 mg/dl.

Normally a healthy person's glucose levels will be in the threshold of 70 mg/dl on the lower side and 110 mg/dl on the higher side during the fasting time. Whenever food intake happens the ingredients will be broken into carbohydrates and further into glucose. As a result, the glucose levels will be increased in the blood serum. The

CONTACT: Sarode Rekha[a]. Email: statisticsrekha@gmail.com

indicators of increased glucose levels in the serum will be reached to the liver in turn, the liver will store a part of the glucose as glycogen.

Whenever glucose levels increase in the bloodstream, the indicators will reach the liver so that a part of the glucose will be stored as glycogen in the liver. Further, the liver with the help of the pancreas produces Beta cells and insulin granules which will be helpful for storing access levels of glucose Pintu the adipose tissues and muscles as a result the increased levels of glucose in the bloodstream will come down. Hence a healthy human body has a mechanism of regulating the levels of glucose intra-sold while consuming the glucose human body cells need insulin to accommodate the glucose molecule into the inner parts of the cell.

The extended hidden Markov model (HMM) was used to analyse spatial heterogeneity in rare phenomenon count data. It explored using Poisson and finite-mixture models in disease mapping [1]. The widespread use of HMMs in biological sequence analysis, understanding their applications in tasks like sequence alignments, gene annotation, and similarity searches across molecular biology, through a tutorial-style review [2]. The HMM used in predicting diabetes risk over time, along with the application of the IDF (International Diabetes Federation) risk score, introduced innovative medical approaches [3]. The Markov model demonstrated that fasting blood sugar trends and factors influencing type 2 diabetes progression and regression, estimating transition probabilities and mean sojourn times to reveal valuable insights into disease dynamics [4]. The Markov employed in predicting the reliability of haemoglobin A1c (HbA1c) as a biomarker for type 2 diabetes and assessed its diagnostic and prognostic significance in a three-state Markov model framework [5]. The HMMs effectively validate the Framingham Diabetes Risk Scoring Model (FDRSM), successfully identifying individuals at increased risk of diabetes within an 8-year period [6]. The HMM was applied to type 2 Diabetes patients' data set for estimating Diabetic levels, chance recovery etc., [7]. The diabetic risk prediction was done by integrating Newton's Divided Difference Method (NDDM) with an HMM [8]. The HMM was used for identifying the disease patterns from discrete-valued time series data derived from diabetic patients' treatment records [9]. A Markov model was used to analyse the progression of COVID-19 and established key health indicators for effective healthcare management and assessment of disease intensity [10]. A Markov model was employed to generate absent event details from data recorded by diabetes patients, resulting in enhanced data quality and continuity, with the outcomes underscoring the efficacy of this method in bolstering the overall data integrity [11].

After reviewing the existing literature, it is felt that the HMM has been used extensively for disease mapping. In order to reach the goal of achieving the dynamics of T2DM, this study has proposed a Markov Modelling to assess the dynamics of the disease progression of type 2 diabetes efficiently. The focussed literature is mostly emphasised using HMM for disease mapping, gene annotation etc., More often the researchers have developed models in HMM. However, in the context of disease progression can be studied by Markov Modelling too with more precision. The outcome has to be properly modelled with suitable Statistical tools.

Hence, the necessity of developing and predicting the parameters for the disease progression of type 2 diabetes motivated us to proceed with the work in this direction. This study has given focus on the development of constructing a Markov model based on transition probabilities. We have formulated the probability distributions for different states in the one length, two length sequences. The study also derived the statistical characteristics based on the probability distributions. Mathematical formulae for all Karl-Pearson's measures were derived from the formulated probability distribu-

tions. Numerical data from sources of Wikipedia is collected and analysed for a better understanding of the model behaviour in a common man's perception. Indicators of intensity of states like *Sub-normal*, *Normal* and *Ab-normal* are obtained from the data. Statistical summary reports are explored after writing the appropriate R-code and its execution. Data interpretation/statistical inferences are carried out from the numerical illustration of real-time.

2. Stochastic Model

This model intends to derive probability mass functions of the discrete distribution of a number of states. Let the states of transitions be of three categories, namely State – 1: *Sub-normal*; State-2: *Normal*; State-3: *Abnormal*. a_{ij} - the probability of transition from i^{th} state to j^{th} state.

$$a_{ij} : Pr\{X_n = j/X_{n-1} = i\}; i, j = 1, 2, 3; a_{ij} \geq 0; \text{and} \sum_{j=1}^{3} a_{ij} = 1; \forall i = 1, 2, 3$$

Let there be 'i' and 'j' states in which 'i' is the state of the previous trial and 'j' be the state of the current trial $i, j = 1, 2, 3$ where, 1, 2, 3 represents the identified positive cases in the states of *Sub-normal*, *Normal*, *Ab-normal* respectively.

π denotes the initial probability vector $\pi = (\pi_1 \quad \pi_2 \quad \pi_3)$ and $\sum_{i=1}^{3} \pi_i = 1$

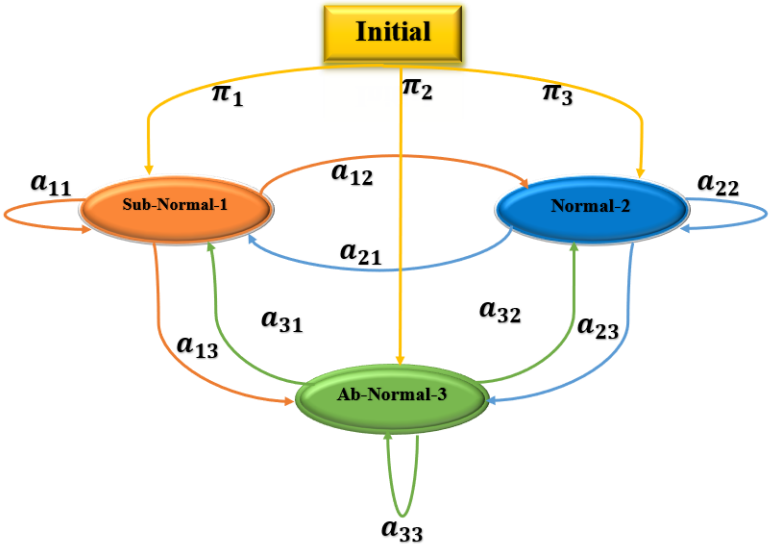

Figure 1. The schematic diagram for three state Markov model of Type 2 Diabetes progression

The notations presented below will be utilised in the upcoming sections, namely Section 3 and Section 4.

$$a(1) = \sum_{i=1}^{3} \pi_i a_{i1}; \quad a(2) = \sum_{i=1}^{3} \pi_i a_{i2}; \quad a(3) = \sum_{i=1}^{3} \pi_i a_{i3}$$

3. Probability distributions for one length of a sequence

$P[X(R)] = n$, Let R be a random variable, which denotes the number of times a specific state occurs in the study of one length. In this study, R can consist of three states namely, *Sub-Normal*, *Normal* and *Ab-Normal* i.e., R=S (or) N (or) A. 'n' will be the number of times the happening of the state, it can take values 0 to 1. Here, '0' indicates the non-happening, and '1' indicates the happening of the particular state under study.

3.1. Probability distribution of Sub-Normal state in one length sequence

$$P[X(S)] = \begin{cases} \sum_{j=2}^{3} a(j) & ;\text{for } X(S) = 0 \\ a(1) & ;\text{for } X(S) = 1 \\ 0 & ;\text{otherwise} \end{cases} \tag{1}$$

Characteristic function of probability distribution of *Sub-Normal* state in one length sequence is

$$\phi_{[X(S)]}(t) = \sum_{j=2}^{3} a(j) + e^{it}[a(1)];\ \text{Where}, a(j) = \sum_{i=1}^{3} \pi_i a_{ij}\ \forall j = 2, 3 \tag{2}$$

3.2. Probability distribution of Normal state in one length sequence

$$P[X(S)] = \begin{cases} \sum_{j=1, j\neq 2}^{3} a(j) & ;\text{for } X(N) = 0 \\ a(2) & ;\text{for } X(N) = 1 \\ 0 & ;\text{otherwise} \end{cases} \tag{3}$$

Characteristic function of probability distribution of *Normal* state in one length sequence is

$$\phi_{[X(N)]}(t) = \sum_{j=1, j\neq 2}^{3} a(j) + e^{it}[a(2)];\ \text{Where}, a(j) = \sum_{i=1}^{3} \pi_i a_{ij}\ \forall j = 1, 3 \tag{4}$$

3.3. Probability distribution of Ab-Normal state in one length sequence

$$P[X(A)] = \begin{cases} \sum_{j=1}^{2} a(j) & ;\text{for } X(A) = 0 \\ a(3) & ;\text{for } X(A) = 1 \\ 0 & ;\text{otherwise} \end{cases} \tag{5}$$

Characteristic function of probability distribution of *Ab-Normal* state in one length sequence is

$$\phi_{[X(A)]}(t) = \sum_{j=1}^{2} a(j) + e^{it}[a(3)];\ \text{Where}, a(j) = \sum_{i=1}^{3} \pi_i a_{ij}\ \forall j = 1, 2 \tag{6}$$

3.4. Statistical characteristics of probability distributions in one length sequence

In this section, some statistical characteristics are derived for the developed probability distributions given in equations 1, 3 and, 5 respectively. In this a(R) will be the probability for happening of the states where the values of 'R' will be '1' which indicates the probability of *Sub-normal* state,'2' which indicates the probability of *Normal* state and '3' indicates the probability of *Ab-Normal* state.

3.4.1. Mean state of one length sequence

$$E[X(R)] = a(R);\ \ Where, a(R) = \sum_{i=1}^{3} \pi_i a_{iR}\ \ \forall R = 1, 2, 3 \tag{7}$$

3.4.2. Variance state of one length sequence

$$V[X(R)] = a(R)\Big[1 - a(R)\Big];\ \ \forall R = 1, 2, 3 \tag{8}$$

3.4.3. Third central moment's state of one length sequence

$$\mu_3[X(R)] = a(R)\Big[1 - a(R)\Big]\Big[1 - 2a(R)\Big];\ \ \forall R = 1, 2, 3 \tag{9}$$

3.4.4. Pearson's coefficient of Skewness of one length sequence

$$\beta_1[X(R)] = \left[a(R)\Big[1 - a(R)\Big]\Big[1 - 2a(R)\Big]\right]^2 \left[a(R)\Big[1 - a(R)\Big]\right]^{-3};\ \ \forall R = 1, 2, 3 \tag{10}$$

3.4.5. Pearson's coefficient of Kurtosis of one length sequence

$$\beta_2[X(R)] = \left[a(R)\Big[1 - a(R)\Big]^2\Big[1 - 3a(R)\Big]\Big[1 - a(R)\Big]\right]\left[a(R)\Big[1 - a(R)\Big]\right]^{-2};\ \ \forall R = 1, 2, 3 \tag{11}$$

4. Probability distributions for two length of sequence

On similar lines of one length in section 3, the sequence of two length is considered in this section. The occurrence of non-happening of the state in two length, happening of the state once in two length, and happening of the state twice in two length are to be modelled for the derivation of the probability distribution. The possibility of taking the value of 'n' is 0, 1, 2.

4.1. Probability distribution of Sub-Normal state in two length sequence

$$P[X(S)] = \begin{cases} \sum_{j=2}^{3}\Big(\sum_{i=2}^{3} a(i)a_{ij}^{(2)}\Big) & ;\text{for } X(S)=0 \\ \sum_{i=2}^{3} a(i)a_{i1}^{(2)} + a(1)\sum_{j=2}^{3} a_{1j}^{(2)} & ;\text{for } X(S)=1 \\ a(1)a_{11}^{(2)} & ;\text{for } X(S)=2 \\ 0 & ;\text{otherwise} \end{cases} \tag{12}$$

Characteristic function of probability distribution of *Sub-Normal* state in two length sequence is

$$\phi_{[X(S)]}(t) = \sum_{j=2}^{3}\Big(\sum_{i=2}^{3} a(i)a_{ij}^{(2)}\Big) + e^{it}\Big[\sum_{i=2}^{3} a(i)a_{i1}^{(2)} + a(1)\sum_{j=2}^{3} a_{1j}^{(2)}\Big] + e^{2it}\Big[a(1)a_{11}^{(2)}\Big] \tag{13}$$

4.2. Probability distribution of Normal state in two length sequence

$$P[X(N)] = \begin{cases} \sum_{j=1,j\neq 2}^{3}\Big(\sum_{i=1,i\neq 2}^{3} a(i)a_{ij}^{(2)}\Big) & ;\text{for } X(N)=0 \\ \sum_{i=1,i\neq 2}^{3} a(i)a_{i2}^{(2)} + a(2)\sum_{j=1,j\neq 2}^{3} a_{2j}^{(2)} & ;\text{for } X(N)=1 \\ a(2)a_{22}^{(2)} & ;\text{for } X(N)=2 \\ 0 & ;\text{otherwise} \end{cases} \tag{14}$$

Characteristic function of probability distribution of *Normal* state in two length sequence is

$$\phi_{[X(N)]}(t) = \sum_{j=1,j\neq 2}^{3}\Big(\sum_{i=1,i\neq 2}^{3} a(i)a_{ij}^{(2)}\Big) + e^{it}\Big[\sum_{i=1,i\neq 2}^{3} a(i)a_{i2}^{(2)} + a(2)\sum_{j=1,j\neq 2}^{3} a_{2j}^{(2)}\Big] + e^{2it}\Big[a(2)a_{22}^{(2)}\Big] \tag{15}$$

4.3. Probability distribution of Ab-Normal state in two length sequence

$$P[X(A)] = \begin{cases} \sum_{j=1}^{2}\Big(\sum_{i=1}^{2} a(i)a_{ij}^{(2)}\Big) & ;\text{for } X(A)=0 \\ \sum_{i=1}^{2} a(i)a_{i3}^{(2)} + a(3)\sum_{j=1}^{2} a_{3j}^{(2)} & ;\text{for } X(A)=1 \\ a(3)a_{33}^{(2)} & ;\text{for } X(A)=2 \\ 0 & ;\text{otherwise} \end{cases} \tag{16}$$

Characteristic function of *Ab-Normal* state probability distribution in two length sequence is

$$\phi_{[X(A)]}(t) = \sum_{j=1}^{2}\Big(\sum_{i=1}^{2} a(i)a_{ij}^{(2)}\Big) + e^{it}\sum_{i=1}^{2} a(i)a_{i3}^{(2)} + a(3)\sum_{j=1}^{2} a_{3j}^{(2)} + e^{2it}\Big[a(3)a_{33}^{(2)}\Big] \tag{17}$$

4.4. Statistical characteristics of probability distributions in two length sequence

Some statistical characteristics are explored for the probability distributions shown in the above equations 12,14, and 16. Let us consider, γ_R is the probability of non-happening of the state, α_R is the probability of happening of the state once, β_R is the probability of happening of the state twice, were, R=1,2,3.
Let,
$\alpha_R = \sum_i a(i)a_{i1}^{(2)} + a(R)\sum_j a_{1j}^{(2)};\ \ \beta_R = a(R)a_{RR}^{(2)};\ \ \gamma_R = \sum_j\Big(\sum_i a(i)a_{ij}^{(2)}\Big)\forall R = 1,2,3.$
Where, in α_R and γ_R, if R=1 then i,j=2,3, if R=2 then i,j=1,3 and , if R=3 then i,j=1,2.

4.4.1. Mean state of two length sequence

$$E[X(R)] = \alpha_R + 2\beta_R \ \ \forall R = 1,2,3 \tag{18}$$

4.4.2. Variance state of two length sequence

$$V[X(R)] = \alpha_R(1-\alpha_R) + 4\beta_R(1-\alpha_R-\beta_R) \ \ \forall R = 1,2,3 \tag{19}$$

4.4.3. Third central moment's state of two length sequence

$$\mu_3[X(R)] = 2\alpha_R^3 + 16\beta_R^3 - 3\alpha_R^2(1-4\beta_R) - 24\beta_R^2(1-\alpha_R) + \alpha_R(1-18\beta_R) + 8\beta_R \quad \forall R = 1,2,3 \tag{20}$$

4.4.4. Person's coefficient of Skewness in two length sequence

$$\beta_1[X(R)] = \Big[2\alpha_R^3 + 16\beta_R^3 - 3\alpha_R^2(1-4\beta_R) - 24\beta_R^2(1-\alpha_R) + \alpha_R(1-18\beta_R) + 8\beta_R\Big]^2 \Big[\alpha_R(1-\alpha_R) + 4\beta_R(1-\alpha_R-\beta_R)\Big]^{-3} \quad \forall R = 1,2,3 \tag{21}$$

4.4.5. Person's coefficient of Kurtosis in two length sequence

$$\beta_2[X(R)] = \Big[\alpha_R + 4\beta_R(4-10\alpha_R) - 4\alpha_R^2(1-12\beta_R) - 8\beta_R^2(8-15\alpha_R+9\alpha_R^2) + 6\alpha_R^3(1-4\beta_R) 96\beta_R^3(1-\alpha_R) - 3\alpha_R^4 - 48\beta_R^4\Big]\Big[\alpha_R(1-\alpha_R) + 4\beta_R(1-\alpha_R-\beta_R)\Big]^{-2} \forall R = 1,2,3 \tag{22}$$

5. Description of Methodology

The current study is on developing Markov model for three (*Sub-Normal*, *Normal*, *Ab-Normal*) states of type 2 Diabetes. Separate probability distribution developed for each state (*Sub-Normal*, *Normal*, *Ab-Normal*). Statistical measures like mean, variance, coefficient variation, and Pearson's coefficients are derived for each state using a corresponding developed probability distribution. For understanding the behaviour of developed model numerical data set is to be considered for this study. The secondary data on Diabetes of procured from the internet sources which consists of 190 observations/records of 2 patients. The observations have been recorded every 15 minutes from 8 hours to 20 hours from the patients. The number of incidences per 15 min is categorised as three transitions: *Sub-Normal*, *Normal*, and *Ab-normal*.

The classification of states in this study is as follows: *Sub-normal* transition arrives if the current reading score is less than 90 mg/dl (FBS level), and the *Normal* comes if the reading score is between 90 mg/dl and 120 mg/dl. The state of *Ab-normal* has arrived when the current reading's score is more than 120 mg/dl.

6. Results and Discussion

6.1. Probability distribution for Sub-Normal state

The probability distribution for one and two length sequence of *Sub-Normal* state is placed in table-1 and table-2 respectively.

Table 1. Probability distribution for *Sub-Normal* state of one length sequence

X(S)	0	1
P[X(S)]	0.885	0.115

From table-1, it is observed that the non-happening of the *Sub-Normal* state has a chance of 0.885 and the happening of the condition has a chance of 0.115. Hence, it is inferred that the non-happening of the *Sub-Normal* is more likely than the happening in the *Sub-Normal* state.

Table 2. Probability distribution for *Sub-Normal* state of two length sequence

X(S)	0	1	2
P[X(S)]	0.84691	0.0690166	0.084073

From table-2, it is observed that the non-happening of the state has a chance of 0.84691 and happening of the state once in two length is 0.0690166 and the occurrence of the condition twice in two length has a chance of 0.084073. Hence, we may interpret the result of non-occurrence of the *Sub-Normal* state as more likely when compared to others.

6.1.1. Statistical measures for Sub-Normal state

The statistical characteristics for the *Sub-Normal* state are placed in the table-3 for understanding the behaviour of the developed model.

Table 3. Statistical results for *Sub-Normal* State

Statistical measure	1 length	2 length
Mean	0.115	0.237163
Variance	0.101775	0.349063
Third Cental moment	0.078367	0.479908
Beta 1	5.825596	5.415072
Beta 2	6.825596	6.877236

6.2. Probability distribution for Normal state

The probability distribution for one and two length of *Normal* state is placed in table-4 and table-5 respectively.

Table 4. Probability distribution for *Normal* state of one length sequence

X(N)	0	1
P[X(N)]	0.630789474	0.369210526

From table-4, it is observed that the non-happening of the state is having a chance of 0.630789474 and the happening of the state has a chance of 0.369210526. Hence, it may be inferred that the non-happening of the *Normal* state has more likely than the happening of the *Normal* state.

Table 5. Probability distribution for *Normal* state of two length sequence

X(S)	0	1	2
P[X(S)]	0.512818977	0.234412236	0.252768787

From table-5, it is observed that the non-happening of the state has a chance of 0.512818977 and happening of the state once is 0.234412236 and the occurrence of the state twice in a two length has a chance of 0.252768787. Hence, we may interpret the result of non-occurrence of the *Normal* state as more likely when compared to others.

6.2.1. Statistical measures for Normal state

The statistical characteristics for the *Normal* state are placed in the table-6 for understanding the behaviour of the developed model.

Table 6. Statistical results for *Normal* State

Statistical measure	1 length	2 length
Mean	0.369210526	0.73994981
Variance	0.232894114	0.697961663
Third Cental moment	0.060920197	0.302051179
Beta 1	0.293796802	0.268328229
Beta 2	1.293796802	1.625792875

6.3. Probability distribution for Ab-Normal state

The probability distribution for one and two length of the *Ab-Normal* state is placed in table-7and table-8 respectively.

Table 7. Probability distribution for *Ab-Normal* state of one length

X(A)	0	1
P[X(A)]	0.484211	0.515789474

From table-7, it is observed that the non-happening of the state is having a chance of 0.484211 and the happening of the state has a chance of 0.515789474. Hence, it may be inferred that the happening of the *Ab-Normal* is more likely than the non-happening of the *Ab-Normal* state.

Table 8. Probability distribution for *Ab-Normal* state of two day length

X(A)	0	1	2
P[X(A)]	0.400664	0.17578563	0.423551

From table-8, it is observed that the non-happening of the state has a chance of 0.400664 and the occurrence of the state once is 0.17578563 and the occurrence of the state twice in a two length has a chance of 0.423551. Hence, it interprets the result that the occurrence of the *Ab-normal* state on twice in two length is more likely when compared to other states.

6.3.1. Statistical measures for Ab-Normal state

The statistical characteristics for the *Ab-Normal* state are placed in the table-9 for understanding the behaviour of the developed model.

Table 9. Statistical results for *Ab-Normal* State

Statistical measure	1 length	2 length
Mean	0.515789	1.022887
Variance	0.249751	0.823691
Third Cental moment	-0.00789	-0.03368
Beta 1	0.003993	0.00203
Beta 2	1.003993	1.215549

6.4. Recommendations

From table 3,6, and 9 it is observed that the average occurrence of one length for *Sub-Normal*, *Normal*, and *Ab-Normal* states are 0.115, 0.369210526, and 0.515789 respectively. Hence, it reveals that the average occurrence of an *Ab-Normal* state in one length is more likely than other states (*Sub-Normal*, and *Normal*). Similarly, in the sequence of two length the average occurrence for *Sub-Normal*, *Normal*, and *Ab-Normal* states are 0.237163, 0.73994981, and 1.022887 respectively. Hence, it reveals that the occurrence of an *Ab-Normal* state ($1.022887 \simeq 1$) in two length is more likely than other states (*Sub-Normal*, and *Normal*). Healthcare professionals may find these findings useful in controlling and monitoring patient diabetes levels.

The variance of one length for *Sub-Normal*, *Normal*, and *Ab-Normal* states are 0.101775, 0.232894114, and 0.249751 respectively. In two length for *Sub-Normal*, *Normal*, and *Ab-Normal* states are 0.349063, 0.697961663, and 0.823691 respectively. Hence, it is observed that there is less volatility observed in the *Sub-Normal* and more in the *Ab-Normal* state of both one and two length. It is indicating that there are fluctuations in blood sugar levels. These results may be helpful to the healthcare professionals giving good treatment and assessing glycemic control stability.

Regarding Skewness, the third central in one and two sequence length of *Sub-Normal*, and *Normal* states are non-negative and positive in the *Ab-Normal* state. It is interpreted that the distribution is positively skewed in *Sub-Normal* and *Normal* states, and negatively skewed in *Ab-Normal* state. These results may help doctors in giving optimal treatment to patients in controlling diabetic levels.

The coefficient of Kurtosis in both one and two length of *Normal* and *Ab-Normal* states is less than 3, and greater than 3 in the *Sub-Normal* state. It reveals that the distribution is platy Kurtic in *Normal*, and *Ab-Normal* states, and lepto Kurtic in Sub-*Normal* state. These are helpful to healthcare professionals in managing hypoglycemic or hyperglycemic incidents.

7. Conclusion

The provided analysis delves into the probabilities of *Sub-Normal*, *Normal*, and *Ab-Normal* states, revealing that the non-occurrence of *Sub-Normal* and *Normal* states is more likely. This trend persists across both one and two length sequences, where the *Ab-Normal* state's occurrence stands out. The data highlights the *Ab-Normal* state's prevalence, indicating its significance in diabetes management.

Moreover, the examination of variance underscores the volatility in blood sugar levels across these states. Notably, the *Sub-Normal* state showcases lower volatility, while the *Ab-Normal* state demonstrates higher fluctuations. This insight into variability can guide healthcare professionals in devising strategies for stable glycemic control.

Skewness and kurtosis analyses offer further insights. *Sub-Normal* and *Normal*

states exhibit positive skewness and platykurtic distribution, signifying their distribution shape. In contrast, the *Ab-Normal* state displays negative skewness and leptokurtic distribution, revealing its distinct distribution characteristics.

In summation, this multifaceted analysis provides comprehensive insights crucial for healthcare practitioners managing diabetes. The probabilities, average occurrences, volatility, and distribution properties of different states collectively contribute to an enhanced understanding of blood sugar dynamics, fostering more effective and tailored patient care strategies.

Acknowledgemnet

We are thankful to the DST INSPIRE. Ms. Sarode Rekha is supported by INSPIRE fellowship (IF190741) of the Department of Science and Technology (DST), Government of India.

References

[1] Green, P. J., and Richardson, S. (2002). Hidden Markov models and disease mapping. Journal of the American statistical association, 97(460), 1055-1070.

[2] Yoon, B. J. (2009). Hidden Markov models and their applications in biological sequence analysis. Current genomics, 10(6), 402-415.

[3] Radheya Suresh Afre, (2015). Tracking diabetes using Hidden Markov Model, J Diabetes Metab, 7th Indo Global Diabetes Summit and Medicare Expo, 6(11), http://dx.doi.org/10.4172/2155-6156.C1.040

[4] Nazari, M., Nazari, S. H., Zayeri, F., Dehaki, M. G., and Baghban, A. A. (2018). Estimating transition probability of different states of type 2 diabetes and its associated factors using Markov model. Primary Care Diabetes, 12(3), 245-253.

[5] Goel, K., Grover, G., Sharma, A., and Bae, S. (2018). Multistate Markov model for predicting the natural disease progression of type 2 diabetes based on hemoglobin A1c. Journal of Nephropharmacology, 8(1), e04-e04.

[6] Perveen, S., Shahbaz, M., Keshavjee, K., and Guergachi, A. (2019). Prognostic modeling and prevention of diabetes using machine learning technique. Scientific reports, 9(1), 13805. https://doi.org/10.1038/s41598-019-49563-6.

[7] Varshney, M. K., Sharma, A., Goel, K., Ravi, V., and Grover, G. (2020). Estimation of transition probabilities for diabetic patients using hidden Markov model. International Journal of System Assurance Engineering and Management, 11, 329-334.

[8] Perveen, S., Shahbaz, M., Ansari, M. S., Keshavjee, K., and Guergachi, A. (2020). A hybrid approach for modeling type 2 diabetes mellitus progression. Frontiers in Genetics, 10, 1076.

[9] Lin, W., Orgun, M. A., and Williams, G. J. (2001, August). Multilevels hidden markov models for temporal data mining. In Proceedings of the KDD 2001 Workshop on Temporal Data Mining (held in conjunction with the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2001)).

[10] T. R. Padi, Kanimozhi. V, and Sakkeel P.T., (2021). Markov Model of Covid-19 Disease Progression, Stochastic Modeling and Applications, 25(1), 101-114.

[11] De Carvalho, D. F., Kaymak, U., Van Gorp, P., and van Riel, N. (2022). A Markov model for inferring event types on diabetes patients data. Healthcare Analytics, 2, 100024.