IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 22, Issue 1, Ser. 1 (Jan. – Feb. 2026), PP 27-31
www.iosrjournals.org

Hidden Markov Modelling Of Food Consumption Types And Their Association With Seasons

Chiranjeevi Gudala
National Statistical Commission Secretariat, Govt. Of India, New Delhi, India

Shanti Priya Karri
AI/ML Researcher, Vishakhapatnam, Andhra Pradesh, India

Abstract
Understanding the seasonal dynamics of food consumption is essential for effective dietary assessment, agricultural planning, and food supply chain management. This study proposes a Hidden Markov Model (HMM) framework to analyse the relationship between food consumption types and seasons, treating seasons as latent (hidden) states and food types as observable states. Transition probability matrices are constructed to capture seasonal progression, while emission probabilities quantify the likelihood of observing specific food types within each season. Based on the proposed stochastic framework, explicit mathematical formulations are developed and key statistical measures-including mean, variance, coefficient of variation, and correlation coefficients-are derived. The study further emphasizes the development of user-friendly computational tools and software implementations based on these formulations. The proposed approach demonstrates the potential of HMMs as a flexible and robust statistical methodology for modelling seasonal food consumption patterns.
Keywords: *Hidden Markov Model, Food Consumption, Seasonal Dynamics, Transition Probability Matrix, Stochastic Modelling*

Date of Submission: 17-01-2026 Date of Acceptance: 27-01-2026

I. Introduction

Food consumption patterns are inherently dynamic and often influenced by seasonal, cultural, and environmental factors. Traditional analytical approaches frequently rely on simple Markov chains or static statistical models, which may fail to capture latent structures and temporal dependencies underlying dietary behaviour. Hidden Markov Models (HMMs), by contrast, provide a principled probabilistic framework for modelling systems in which observed outcomes are driven by unobservable states evolving over time.

The applicability of HMMs to food and health-related domains has been well documented. Ip et al. (2013) employed multi-profile HMMs to model mood, dietary intake, and physical activity in childhood obesity interventions, enabling the identification of latent health states from longitudinal data. Costa et al. (2016) demonstrated the effectiveness of HMMs in automatic meal intake monitoring for elderly individuals using non-intrusive sensing technologies. Dorairaj (2018) presented intuitive examples of HMMs to illustrate how hidden variables such as weather or seasons can be inferred from observable behaviour. More recent studies have extended HMM applications to food quality assessment (Liu et al., 2020), simulated food consumption data generation (Pan et al., 2022), ecological indicators (Rand et al., 2024), and seasonal livestock movement behaviour (Cheng et al., 2025).

Despite this growing body of literature, the explicit use of HMMs to model food consumption types as observable processes driven by latent seasonal states remains limited. Existing studies predominantly focus on Markov chains or classification-based approaches, often treating seasonality as an external or fully observable factor. This study addresses this gap by proposing an HMM-based framework in which seasons are modelled as hidden states, allowing for probabilistic inference, forecasting, and richer statistical characterisation of food consumption dynamics.

The primary objectives of this work are:

i. To formulate a Hidden Markov Model representing the relationship between food consumption types and seasons.
ii. To derive transition and emission probability structures that reflect seasonal progression and dietary variability.
iii. To develop statistical measures and computational tools based on the proposed model for practical applications in food system analysis.

II. Stochastic Model Framework

A Hidden Markov Model is defined by an underlying stochastic process that evolves through a finite set of hidden states over time, and an observable process whose distribution depends on the current hidden state.

In the present context:

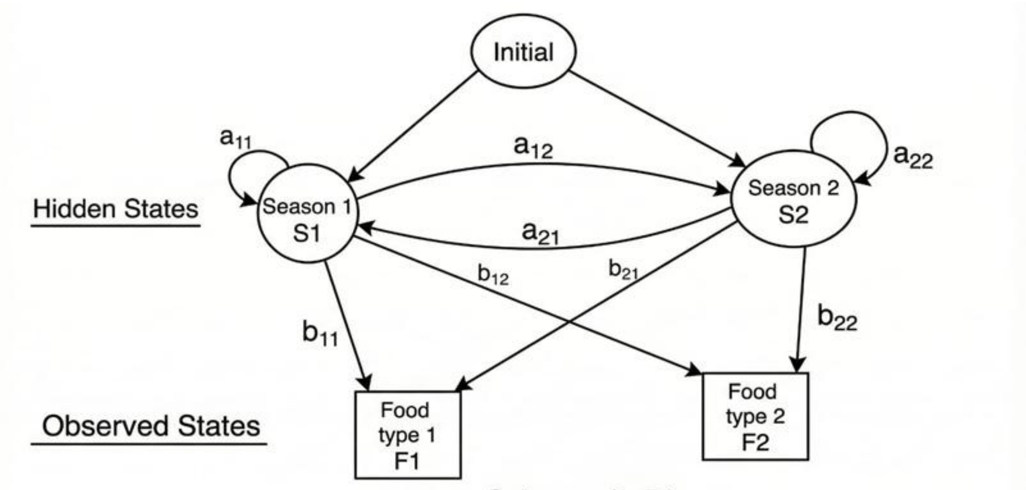
- **Hidden states (S):** Seasons (e.g., Season 1 – dry/summer, Season 2 – wet/winter), or season-driven phases such as harvest or planting periods.
- **Observed states (O):** Food consumption types (e.g., vegetables, fruits, cereals, animal-based foods).

The model is characterised by:

- **Initial state distribution** representing the probability of the system starting in a particular season.
- **Transition probability matrix** capturing the sequential and cyclical nature of seasons.
- **Emission probability matrix** representing the likelihood of observing specific food types given a season.

This formulation allows the inference of the most probable seasonal sequence underlying observed food consumption data using standard HMM algorithms such as the Forw ard–Backward and Viterbi algorithms.

Let us consider the scenario below, where the weather, the hidden variable, can be season1 (may be dry), season2 (may be wet) and the observed variables are the type of clothing worn. The arrows represent transitions from a hidden state to another hidden state or from a hidden state to an observed variable.



The schematic diagram for two state Hidden Markov model of food consumption types and their association with seasons

III. Methodology

The study develops a two-state Hidden Markov Model corresponding to two seasonal regimes (Season 1 and Season 2). Separate probability distributions are constructed for each hidden state. Using these distributions, key statistical characteristics—including mean, variance, coefficient of variation, and Pearson correlation coefficients—are analytically derived.

To demonstrate the behaviour of the proposed model, a numerical dataset is considered. Probability distributions are analysed for sequences of length one and two, enabling the examination of short-term and extended temporal dynamics. The methodology emphasises clarity and reproducibility, facilitating translation into software tools for applied use.

Mathematical Formulation of the Hidden Markov Model

Let $\{S_t\}_{t=1}^{T}$ denote the sequence of hidden seasonal states and $\{O_t\}_{t=1}^{T}$ the corresponding sequence of observed food consumption types at time t = 1,2,…,T.

The Hidden Markov Model is fully specified by the parameter set λ = (π, A, B) where:

Initial state distribution

$$\pi_i = P\{S_1 = i\},\ i = 1, 2, ..., N$$

State transition probabilities

$$a_{ij} = P\{S_{t+1} = j \mid S_t = i\},\ i, j = 1, 2, ..., N$$

subject to

$$\sum_{j=1}^{N} a_{ij} = 1,\quad a_{ij} \geq 0$$

Emission (observation) probabilities

$$b_j(k) = P(O_t = k \mid S_t = j),\ j = 1, 2, ..., N;\ k = 1, 2, ..., M$$

$$with$$

$$\sum_{k=1}^{M} b_i(k) = 1,\quad b_i(k) \geq 0$$

In this study, N=2 represents two seasonal regimes, while M denotes the number of food consumption categories.

Likelihood Function and Forward Algorithm

Given an observation sequence O = (O_1,O_2,…,O_T), the likelihood of the model is

$$P(O \mid \lambda) = \sum_{S} P(O, S \mid \lambda)$$

To compute this efficiently, the **Forward algorithm** is used. Define

$$\alpha_t(i) = P(O_1, O_2, ..., O_t, S_t = i \mid \lambda)$$

Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1)$$

Recursion:

$$\alpha_{t+1}(j) = \left(\sum_{i=1}^{N} \alpha_t(i) a_{ij} \right) b_j(O_{t+1})$$

Termination:

$$P(O \mid \lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

State Decoding Using the Viterbi Algorithm

To infer the most probable sequence of seasons underlying the observed food consumption, the Viterbi algorithm is employed.

$$\delta_t(i) = \max_{S_1, \ldots, S_{t-1}} P(S_1, \ldots, S_t = i, O_1, \ldots, O_t \mid \lambda)$$

Recursion:

$$\delta_{t+1}(j) = \max_i \left[\delta_t(i) a_{ij} \right] b_j(O_{t+1})$$

Backtracking yields the most likely seasonal path S^, allowing prediction of the season on day t based solely on observed food consumption.

Statistical Measures
For each season-specific emission distribution, the mean and variance are defined as

$$\mu_j = \sum_{k=1}^{M} k b_j(k)$$

$$\sigma_j^2 = \sum_{k=1}^{M} (k-\mu_j)^2 b_j(k)$$

The coefficient of variation is

$$CV_j = \frac{\sigma_j}{\mu_j}$$

These measures quantify the stability and variability of food consumption patterns across seasonal regimes.

IV. Simulated Data And Results

The analysis investigates:
- Probability distributions of food consumption for single-length sequences.
- State-specific probability distributions for Season 1 and Season 2.
- Statistical characteristics of these distributions.

A similar analysis is extended to two-length sequences, highlighting how transition probabilities influence observed consumption patterns over time. The results illustrate that the HMM framework effectively captures seasonal persistence, transition behaviour, and variability in food consumption types, offering advantages over simpler Markov chain models.

Simulated Data Generation
To demonstrate the applicability and behaviour of the proposed Hidden Markov Model, a simulated dataset was generated. A two-state HMM was considered, corresponding to **Season 1** and **Season 2**. The hidden seasonal process was assumed to follow a first-order Markov chain with the following transition probability matrix:

From / To	Season 1	Season 2
Season 1	0.75	0.25
Season 2	0.30	0.70

This structure reflects realistic seasonal persistence, where remaining in the same season is more likely than an abrupt transition.

Food consumption types were treated as observable states. For illustrative purposes, three food categories were considered: **Food A** (e.g., cereals), **Food B** (e.g., vegetables), and **Food C** (e.g., animal-based foods). The emission probability distributions were defined as:

Season	Food A	Food B	Food C
Season 1	0.50	0.35	0.15
Season 2	0.25	0.30	0.45

Using the above parameters, a synthetic observation sequence was generated over multiple time steps, ensuring adequate representation of both seasonal regimes and food choices.

Statistical Analysis of Simulated Results
Based on the simulated sequences, probability distributions of food consumption were analysed for single-length and two-length observation sequences. For each season-specific distribution, key statistical measures were computed, including the mean occurrence probability, variance, and coefficient of variation.

The results indicate that **Season 1** is characterised by higher average probabilities for plant-based food categories, whereas **Season 2** shows increased likelihood of animal-based food consumption. Variance measures reveal greater volatility in food choices during seasonal transition periods, which is consistent with real-world dietary adjustments across seasons.

For two-length sequences, transition effects become more pronounced. The joint probability distributions reflect the influence of seasonal persistence, with food consumption patterns showing stronger dependence on the previous hidden state. These findings demonstrate the advantage of the HMM framework in capturing temporal dependencies that are not observable using independent or static models.

Discussion

The simulated results confirm that the proposed Hidden Markov Model effectively captures both seasonal dynamics and variability in food consumption behaviour. The model allows for probabilistic inference of the underlying season based solely on observed food choices, enabling forecasting of future seasonal states and associated dietary patterns. Compared to traditional Markov chain approaches, the HMM provides richer interpretability by separating observable behaviour from latent seasonal drivers.

V. Conclusion

This study demonstrates the potential of Hidden Markov Models as a robust statistical tool for modelling and predicting food consumption patterns influenced by seasonal dynamics. By treating seasons as latent states, the proposed approach enables probabilistic inference, forecasting of future seasonal states from observed food choices, and comprehensive statistical characterisation of the system.

The findings are relevant to applications in food supply chain management, dietary planning, and agricultural forecasting. The framework can be extended to compute higher-order statistical measures such as skewness and kurtosis, and to incorporate additional hidden states or covariates.

VI. Future Scope

Future research may extend this framework to:

- Multi-season and multi-region models.
- Integration of nutrient-level data and dietary quality indicators.
- Real-time forecasting of food intake and nutritional risks.
- Applications in ecological foraging studies and public health nutrition.

Compliance with Ethical Standards

Conflict of Interest

The author declares that there is no conflict of interest.

Disclaimer

The views expressed in this paper are purely personal views and not the views of the institution/ Department to which the author(s) is affiliated i.e., Government of India.

Research Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1]. E. H. Ip Et Al., 2013, “Multi-Profile Hidden Markov Model For Mood, Dietary Intake, And Physical Activity In An Intervention Study Of Childhood Obesity”, Stat Med. 2013 August 30; 32(19): 3314–3331. Doi:10.1002/Sim.5719.

[2]. Luís Costa Et Al., 2016, “Automatic Meal Intake Monitoring Using Hidden Markov Models”, Procedia Computer Science 100, 110 – 117.

[3]. Sanjay Dorairaj, 2018, “Hidden Markov Models Simplified” Https://Medium.Com/@Postsanjay/Hidden-Markov-Models-Simplified-C3f58728caab.

[4]. Taoping Liu, Wentian Zhang, Mitchell Yuwono, Miao Zhang, Maiken Ueland, Shari L. Forbes, Steven W. Su, 2020, “A Data-Driven Meat Freshness Monitoring And Evaluation Method Using Rapid Centroid Estimation And Hidden Markov Models”, Sensors And Actuators B: Chemical, Volume 311, 2020, 127868, Issn 0925-4005, Https://Doi.Org/10.1016/J.Snb.2020.127868

[5]. Xinyue Pan Et Al., 2022, “Simulating Personal Food Consumption Patterns Using A Modified Markov Chain”, Madima ’22, October 10, 2022, Lisboa, Portugal Arxiv:2208.06709v1 [Cs.Cv] 13 Aug 2022.

[6]. Zoe R. Rand Et Al., 2024 “Using Hidden Markov Models To Develop Ecosystem Indicators From Non-Stationary Time Series”, Ecological Modelling, Volume 495, 2024, 110800, Issn 0304-3800, Https://Doi.Org/10.1016/J.Ecolmodel.2024.110800.

[7]. Cheng, H., Et Al. 2025, “Seasonal Movement Behavior Of Domestic Goats In Response To Environmental Variability And Time Of Day Using Hidden Markov Models”. Mov Ecol 13, 28 (2025). Https://Doi.Org/10.1186/S40462-025-00557-2.