Assessment of Fruit Consumption and Associated Factors: A Study in Bihar, India

Submitted by
RAJAGOPAL K.M
Department of Statistical sciences
Kannur University

Contents

1. INTRODUCTION

Fruit Consumption in our daily diet cannot be overstated, as it plays a critical role in maintaining and enhancing our overall health and well-being. Fruits are not only delicious and refreshing but also packed with essential nutrients, vitamins, minerals, and antioxidants that offer numerous health benefits [1]. Eating fruits regularly can greatly improve our health, emphasizing the importance of including them in our diet.

Firstly, fruits are a rich source of essential vitamins and minerals that are crucial for the proper functioning of the body. Vitamins such as vitamin C, found in oranges and strawberries, are vital for the growth and repair of tissues in all parts of the body. Secondly, fruits are packed with antioxidants, which are substances that help protect our bodies from oxidative stress and free radicals, which can damage cells and contribute to aging and diseases such as cancer. The consumption of fruits also has a positive impact on mental health. Studies have shown that diets high in fruits (and vegetables) are associated with lower levels of stress, anxiety, and depression. Furthermore, consuming fruits can aid in weight management [2,3].

According to the Indian Council of Medical Research the per capita intake in India is 120 grams per day, which is much lower than the average for other nations such as the United States (202 grams per day). Among fruit-rich regions, tropical countries often have low levels of fruit consumption due to various factors such as availability, affordability, and cultural preferences. In Rural areas, the household production will also increase the household consumption. There is a positive association between fruit production and consumption. Increased fruit production not only enhances the availability of fruits for own consumption but also provides opportunities for income generation through sales. This income can be used to purchase additional food items, further enhancing dietary diversity.

Among all Indian states, Bihar has the highest proportion (51.9%) of the country's population classified as "multidimensionally poor" by the Multidimensional Poverty Index [4]. The agriculture sector in Bihar is a significant contributor to the state's economy, providing livelihoods to a large portion of the population. The state's agricultural landscape is characterized by the cultivation of staple crops such as rice, wheat, maize, pulses, oilseeds, and vegetables. Despite having abundant water resources from rivers and groundwater, Bihar's agriculture sector faces challenges related to irrigation infrastructure and water management. The majority of farmers in Bihar hold small and marginal land holdings, which presents challenges in terms of mechanization and scale of operations. However, the state government has been proactive in promoting agriculture through subsidies, promoting high-yielding crop varieties, and investing in irrigation infrastructure. Despite these efforts, Bihar's agriculture sector still grapples with low productivity, lack of modernization, and vulnerability to weather-related risks.

This state relies heavily on agriculture, with fruit production playing a significant role in contributing to rural incomes.

Rank	States	Fruit Production (in '000 MT)
1	Andhra Pradesh	13939.1
2	Maharashtra	9785
3	Gujarat	8413.2
4	Tamil Nadu	6699.9
5	Karnataka	6619.6
6	Madhya Pradesh	5450
7	Uttar Pradesh	5176.1
8	Bihar	4249.2
9	West Bengal	3172.5
10	Kerala	2583.9

Table 1.1: Source: Horticultural Statistics at a Glance 2018, Horticulture Statistics Division, Department of Agriculture, Cooperation & Farmers Welfare (DAC&FW), Ministry of Agriculture & Farmers Welfare

From the table 1.1, Bihar ranks as the 6th largest fruit producer in India. A survey conducted in the state, particularly in the districts of Gaya and Nalanda, known for high rates of undernutrition, examined the consumption and production of various food items. When analyzing fruit consumption, it's been observed that the rate of consumption is notably lower in comparison to the production levels. This observation highlights a potential gap between production and consumption, which could have implications for various stakeholders in the fruit supply chain.

The study aims to understand **that what are the factors influencing fruit consumption and how these factors affect it.** By identifying factors influencing fruit consumption, this study can inform strategies to improve nutrition and reduce the risk of chronic diseases in Bihar. Possible factors influencing fruit consumption could include income levels, where higher income might lead to increased affordability of fruits, thus influencing consumption patterns. Awareness and education about the nutritional benefits of fruits and their importance in a healthy diet could influence consumption behavior. The gender of the head of the household, whether male or female, can also be a significant factor influencing fruit consumption patterns within the household.

2. METHODS

2.1 Data for the Study

A survey was conducted in two rounds among households in the Gaya and Nalanda districts of Bihar. Round 1 was carried out from July to August 2019, involving 2026 households, while Round 2, conducted as a follow-up from December 2019 to January 2020, included 2001 households. These rounds were timed to capture the seasonal variation in food production, especially fruit, which is typically higher in winter (Round 2) than during the rainy and summer seasons, represented in Round 1. These districts were selected due to their significant reported fruit production.

In the 2018 period, Gaya and Nalanda reported fruit production figures of 91125 MT and 91245 MT, respectively [5] and they produce quality foods, have good markets for agricultural produce, and households consume these foods in significant quantities. Additionally, these districts were not part of large-scale agriculture or nutrition interventions. They were selected based on advice from researchers and consultations with government departments. Data was collected through key informant interviews and focus group discussions to understand the pathways from production to consumption, as well as the facilitators and barriers in this process.

Households were selected using a multistage cluster sampling method, where each administrative village served as a cluster. The census served as the sampling frame, and 101 villages were randomly selected from each district. An approximately equal number of villages were chosen from different distance bands (0-5 km, 6-15 km, 16-30 km, and >30 km) from the district headquarters to account for varying distances to a large market. Within each village, 10 households were sampled using the Random Walk method. The village was divided into 5 segments, resembling a pie, based on the village map. In each segment, 4 landmarks were designated, and 1 was randomly chosen. Starting from that point, the sampling direction was determined by the mouth of a centrifuged bottle. This process was repeated for each segment. The sample included both landless households and landholding households engaged in producing nutrient-dense foods such as pulses, eggs, chicken, milk, and green leafy vegetables. Landless households producing such foods were included, as well as those who cultivated on others' farms. Landholding households involved in quality food production were further classified as small landholders (with less than 3 acres) and medium to large landholders (with more than 3 acres).

Household data on the consumption of food and non-food goods and services were collected using a household consumption questionnaire used in the national consumer expenditure survey [6]. Round 1 data were used to examine patterns and factors associated with Fruit consumption. They selected 8 varieties of fruit, such as banana, jackfruit, watermelon, pineapple, guava, orange (mausambi), papaya, mango, kharbooza, pears (naspati), berries, lychee, apple, grapes, pomegranate, sapota, custard apple, jamun, peach, plum, and other fresh fruits. They then determined their consumption rate, production rate, market price, and the monthly per capita income of 2026 households, which was recorded in the Round 1 Survey.

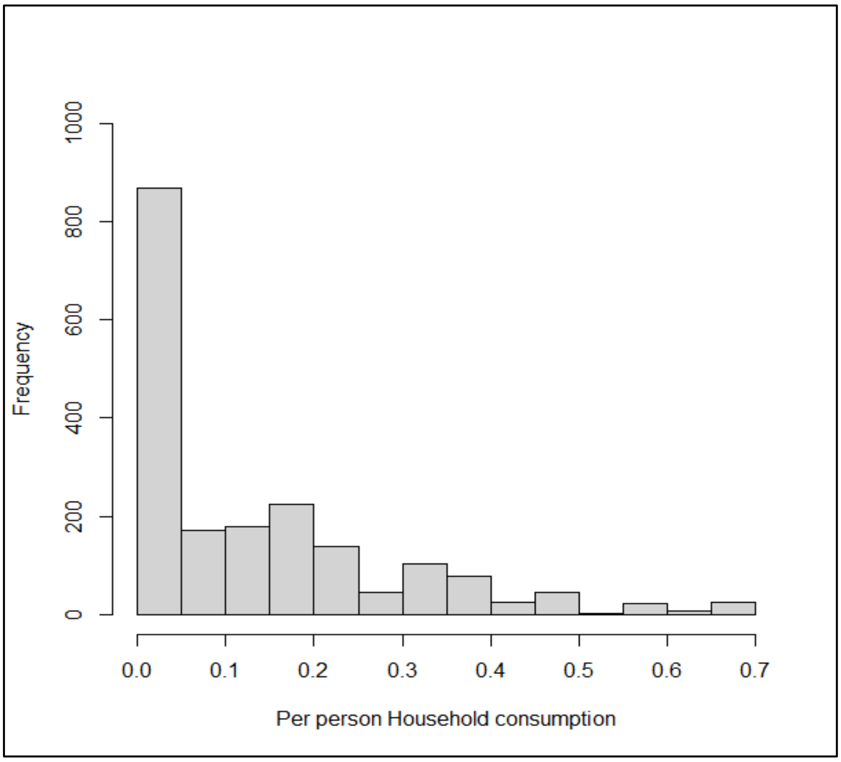

Fig 2.1: Distribution of the per person fruit consumption on the entire samples

Fig 2.1 shows the distribution of fruit consumption rate in the entire sample of size 2026 with a large proportion of zero and a continuous positive value.

Demographical variables such as Monthly per capita Expenditure, Education, Gender of the head of household, Does own any agricultural Land were used to examine the association of these variables with household fruit consumption. In Table 2.1, the types of variables and the levels of the independent variables are presented. This information is crucial for understanding the characteristics of the data and the factors being investigated in the study.

Independent Variables	Type of the Variable	Levels of Independent Variable
Education	Categorical	1. No formal education 2. Primary Education 3. Secondary Education 4. Post-secondary Education 5. Tertiary Education
Gender of the head of households	Categorical	1. Male 2. Female
Does own any Agricultural Lands	Categorical	0. No 1. Yes
Monthly per capita Income	Continuous	

Table 2.1 The Levels of Independent Variables Used in the Study

2.2 Preprocessing techniques

In preprocessing household fruit consumption data, we should handle missing values, encode the education level variable, deal with outliers, and standardize the data. These steps ensure the data is complete, properly represented, and on a consistent scale, leading to more accurate regression results.

2.2.1 Dealing with Outliers

Outliers can substantially impact clustering outcomes, and their identification is typically done using box plots or scatter plots, with removal as a potential step. However, it is crucial to ensure that outliers are genuinely extreme values and not removed arbitrarily. In our analysis, we focused on household fruit consumption rates ranging from 0 to 0.6 kg, beyond which outliers were identified and removed to ensure the integrity of our clustering results.

2.2.2 Missing Data

Data cleaning is a critical step in preparing data for analysis, as it ensures that the dataset is accurate, complete, and relevant to the research objectives. In our project, data cleaning involved several key processes. Firstly, we addressed missing data by either imputing values or removing incomplete records, ensuring that the dataset was complete and suitable for analysis. Secondly, we identified and corrected any inconsistencies or errors in the data, such as typos or incorrect values, to maintain data integrity.

3. Statistical Methods

The household fruit consumption distribution shown in Fig 2.1 has excess zeroes and a positively skewed distribution. We can deal with excess zeroes and positive skewness by applying a mixed discrete-continuous model for the total loss amount. Such an approach would involve the assumption that the dataset is stratified into two groups: the first group has excess zero amounts, and the second group has non-zero which are assumed to have a continuous distribution that accommodates heavy right skewness.

The mixed discrete-continuous probability function of y can then be written as:

$$f(y) = \begin{cases} \pi & if y = 0 \\ (1-\pi)g(y) & if\ y > 0 \end{cases} \quad (1)$$

where $g(y)$ is the density of positive continuous, right skewed distribution and π is the probability of getting zero consumption

3.1. Zero-adjusted Gamma Model

It is a statistical model used to analyze continuous data, where the outcome variable has a mixture of continuous positive values and excess zeros. This model extends the gamma distribution to account for excess zeros in the data by incorporating an additional component specifically for the structural zeros

Two Components: The zero-adjusted gamma model combines two components to capture the characteristics of the data:

1. **Gamma Component:** This component models the positive values of the data. The gamma distribution is commonly used to model positively skewed continuous data.

2. **Zero Component:** This component accounts for the excess zeros in the data. It assumes a mass at zero to capture the proportion of observations that are exactly zero.

Parameters: The model has parameters that need to be estimated:

1. **Gamma Parameters:** These parameters govern the shape and scale of the gamma distribution, capturing the characteristics of the positive values.

2. **Zero-Inflation Parameter:** This parameter controls the proportion of excess zeros in the data. It indicates the probability that an observation is a zero from the zero component rather than from the gamma component

The probability function of the zero-adjusted gamma distribution, denoted by ZAGA(μ, σ, π), is defined by Rigby and Stasinopoulos(2010) [8]:

$$f(y|\mu, \sigma, \pi) = \begin{cases} \pi & if\ y = 0 \\ (1-\pi)\left[\frac{1}{(\sigma^2\mu)^{\frac{1}{\sigma^2}}} \cdot \frac{y^{\frac{1}{\sigma^2}-1} e^{-\frac{y}{\sigma^2\mu}}}{\Gamma\left(\frac{1}{\sigma^2}\right)}\right] & if\ y > 0 \end{cases} \quad (2)$$

For $0 \leq y < \infty$, where $0 < \pi < 1$, mean $\mu > 0$, dispersion $\sigma > 0$ and π is the probability of getting zero
With:
$E(Y) = (1- \pi)\ \mu$ and Var(Y) = $= (1- \pi)\ \mu^2(\pi + \sigma^2)$.

The ZAGA model is implemented using the Generalized Additive Models for Location, Scale, and Shape (GAMLSS) framework, as described by Rigby and Stasinopoulos (2005) [7]. This framework enables the modeling of a wide range of skewed and kurtotic distributions by explicitly incorporating various distributional parameters, such as location/mean, scale/dispersion, skewness, and kurtosis, as functions of predictor variables. Unlike the Generalized Linear Model (GLM) and Generalized Additive Model (GAM) frameworks, which are limited to distributions from the exponential family, GAMLSS allows for the fitting of distributions that do not belong to this family.

3.1.1. Modelling the probability of zero and the positive continuous terms of predictor variables

The ZAGA model includes three components

Mu (Location) Model: The **Mu** part of the model focuses on the expected (mean) value of the response variable, which is household fruit consumption. The coefficients indicate the effect of each predictor variable on the mean consumption. The Mu component is modeled using a log link function

Sigma (Scale) Model: The **Sigma** part models the variability or dispersion of the response variable. The coefficients show how the predictor variables affect the dispersion. The Sigma component is modeled using a log link function

Nu (Shape) Model: The **Nu** part of the model deals with the shape of the distribution, especially for zero-inflated models. The coefficients indicate how the predictor variables influence the probability of observing zeros. The Nu component is modeled using the logit link function.

The mean, μ, and dispersion, σ, of a non-zero consumption amount, and the probability of a zero consumption, π, are modelled in terms of predictor variables using suitable link functions:

$$\log(\mu) = X_k^T \beta_k$$

$$\log(\sigma) = X_k^T \alpha_k \qquad (3)$$

$$\text{logit}(\pi) = X_k^T \gamma_k$$

Here X_k^T is the independent variables and β_k, α_k, γ_k represents the coefficients of the independent variables. The log link functions imply the existence of multiplicative effects on the response variable of the non-zero amount, and also ensure that the predictions will be non-negative.

3.1.2. Maximum likelihood estimation

According to the model, each account is associated with a probability of zero consumption, π, and positive value, y, given that probability of not getting zero, which produces a pair $(1-\pi, y)$. These pairs are then used to form the following log-likelihood function term:

$$lnf(y) = \ln \mathrm{f}(\pi) + ln\, f(y|(1-\pi)) \qquad (4)$$

The log-likelihood is then the sum of Eq. (4) over all accounts. The maximization of the likelihood proceeds in two separate maximizations, one for the component based on $f(\pi)$and one for the component based on $f(y|(1-\pi))$ We used an algorithm which was described by Rigby and Stasinopoulos (2005). The estimates of the probability of zero consumption, and of the mean and dispersion of $g(y)$, are used to compute an estimate of $f(y)$. The independent estimation for $f(\pi)$ and $f(y|(1-\pi))$ avoids the difficulty of having to estimate $f(y)$ directly.

3.1.3. Assumptions for ZAGA models

1. **Data Type**: The ZAG model is suitable for semicontinuous data, where the outcome variable is a mixture of a continuous distribution (gamma) and a point mass at zero. It assumes that the data can be accurately described by a combination of these two components.
2. **Distributional Assumption**: The continuous component of the ZAG model follows a gamma distribution. This assumption implies that the positive values in the data

are gamma-distributed. The gamma distribution is flexible and can accommodate a wide range of shapes, but it is important to ensure that the gamma distribution is a reasonable choice for the data being modeled.

3. **Independence**: The ZAG model assumes that the observations are independent of each other. This means that the presence or absence of a zero value for one observation does not affect the presence or absence of zeros for other observations.
4. **Zero-Inflation**: The ZAG model assumes that the zeros in the data are not due solely to the underlying gamma distribution but are instead "extra" zeros that are added to the model. This is the key assumption that differentiates the ZAG model from a simple gamma model.
5. **Correct Specification**: As with any statistical model, it is important to ensure that the ZAG model is correctly specified for the data being analyzed. This includes choosing the appropriate form of the model (e.g., the correct distribution for the continuous component) and checking the assumptions of the model.

A univariate analysis was performed before proceeding to the final Zero-adjusted gamma regression model to discover predictors for the final model. It helps to streamline the analysis and ensures that only the most influential variables are included in the final model, thereby enhancing the validity and interpretability of the findings.

The analysis for this study was conducted using the R programming language, and several packages were utilized to perform the necessary statistical procedures. The 'gamlss' package is one of the main packages used for analysis.
The **gamlss** package in R, developed by Rigby and Stasinopoulos in 2007 [9], is used for fitting generalized additive models for location, scale, and shape (GAMLSS). This package is particularly useful for modeling complex relationships in data where the response variable may have non-normal distributions or where the variability of the response may be non-constant.
The **gamlss** package offers a flexible framework for fitting a wide range of distributions to the response variable, including the gamma distribution, which can be used for modeling non-negative continuous data. It also allows for the inclusion of smooth functions of predictors, enabling the modeling of non-linear relationships.

4. Results

4.1 Descriptive Analysis of fruit consumption survey

The table 4.1. provides detailed descriptive statistics for various characteristics of 1,936 households in the dataset. The median household size is 6, with an interquartile range (IQR) of 5 to 8. The median monthly per capita expenditure (MPCE) is 5,798 rupees, with an IQR of 4,097 to 8,164 rupees. Total consumption in kilograms has a median of 0.50 kg, with an IQR of 0.00 to 1.00 kg. Household consumption per capita in kilograms has a median of 0.08 kg, with an IQR of 0.00 to 0.21 kg. Household expenditure per capita in rupees has a median of 5 rupees, with an IQR of 0 to 16 rupees. The dataset is predominantly composed of male-headed households (88% male, 12% female), with a median age of 40 and an IQR of 30 to 55. In terms of education, 40% of households have no formal education, 22% have primary education, 24% have secondary education, 7.5% have post-secondary education, and 6.6% have tertiary education. Additionally, 56% of households own agricultural land, with a median size of 1 acre and an IQR of 0 to 6 acres. These statistics provide a comprehensive overview of the demographic, economic, and educational characteristics of the households in the dataset, offering valuable insights for further analysis and policy making

4.2 Univariate Analysis of household consumption of Fruits

The table 4.2. provides a detailed breakdown of median per capita household consumption in a week (HH consumption) across different demographic and socioeconomic groups. It shows that females have a slightly higher median HH consumption (0.092) compared to males (0.083). However, there is a clear pattern of increasing HH consumption with higher education levels, with those having tertiary education exhibiting the highest median consumption (0.16). Similarly, households that own land have significantly higher median HH consumption (0.111) compared to those that do not own land (0.026). Additionally, there is a positive correlation (rho = 0.285) between monthly per capita expenditure (MPCE) and HH consumption, indicating that as MPCE increases, so does HH consumption. These findings underscore the importance of education, land ownership, and income in influencing household consumption patterns, highlighting the need for targeted policies to improve access to education and resources for households to enhance their consumption levels.

Table 4.1 Characteristics of households in fruit consumption survey

Characteristic	**N = 1,936** Median, IQR / n, %
Household Size	6 (5, 8)
MPCE	5,798 (4,097, 8,164)
Total consumption (Kg)	0.50 (0.00, 1.00)
Household Consumption (kg/week)	0.08 (0.00, 0.21)
Household Expenditure (per capita) (INR)	5 (0, 16)
Gender	
Male	1,708 (88%)
Female	228 (12%)
Age	40 (30, 55)
Education	
No Formal Education	778 (40%)
Primary Education	417 (22%)
Secondary Education	469 (24%)
Post-Secondary Education	145 (7.5%)
Tertiary Education	127 (6.6%)
Own any Agriculture Land	
Yes	1,080 (56%)
No	856 (44%)
Size of the land (acres)	1 (0, 6)

All data represented as median (first quartile, third quartile) unless otherwise specified.

Abbreviation: INR, Indian rupee).

Table 4.2 : Distribution of per capita household fruit consumption in a week across independent variables and p-values

Characteristic		Per capita HH consumption (kg/week) Median (IQR)	P-value
Gender	Male	0.083 (0,0.2)	0.014
	Female	0.092 (0,0.25)	
Education	No formal education	0.038 (0,0.19)	.2e-16
	Primary Education	0.071 (0,0.2)	
	Secondary Education	0.125 (0,0.25)	
	Post-Secondary Education	0.17 (0,0.33)	
	Tertiary Education	0.16 (0,0.33)	
Owns any Land	Yes	0.111 (0,0.25)	2.7e-10
	No	0.026 (0,0.2)	
MPCE		rho= 0.285 (positive correlation)	2.2e-16

4.3. Interpretation of the Model

From the study's findings in the table 4.3., we can conclude that several factors influence fruit consumption patterns. On the interpretation of Mu Model Female headed households tend to consume more fruit than male headed household, as indicated by the 0.417 times higher consumption rate in female-headed households. Monthly Per Capita Expenditure (MPCE) has a minimal but positive association with fruit consumption, with a 0.000057-times increase for each unit increase in MPCE. This suggests that higher income levels may lead to slightly higher fruit consumption. Education also plays a role, with higher education levels associated with higher fruit consumption. Primary, secondary, post-secondary, and tertiary education were linked to 0.073, 0.22, 0.48, and 0.39-times higher fruit consumption, respectively, compared to no formal education.
Owning agricultural land did not have a significant impact on fruit consumption, suggesting that other factors may be more influential in determining fruit consumption patterns.

In Nu model interpretation households led by females, the chance of zero fruit consumption was 0.54 times lower than in households led by males, indicating that females are less likely to report zero fruit consumption. For each unit increase in MPCE, the likelihood of zero fruit consumption decreased by 0.00015 times, suggesting a small decrease in the odds of no fruit consumption with higher MPCE.
Education also played a role in zero consumption: households with primary, secondary, post-secondary, and tertiary education had 0.249, 0.572, 0.75, and 0.63-times lower odds of zero fruit consumption, respectively, compared to households with no formal education. Owning agricultural land did not significantly affect the odds of zero fruit consumption.

In the Sigma model interpretation, the variability in fruit consumption in female-headed households was 0.1269 times higher than in male-headed households, indicating higher variability. Each unit increase in MPCE was associated with a 0.000025-times increase in variability, suggesting a slight increase in variability with higher MPCE. Education levels and owning agricultural land did not significantly affect variability.

Table 4.3. Summary of zero adjusted gamma model for the excess zeroes and a positively continuous distribution

Mu Coefficients Link function: log			
Variable	**Estimate**	**Std Error**	**p-value**
(Intercept)	-1.879e+00	6.425e-02	< 2e-16
Gender(female)	4.167e-01	6.979e-02	2.79e-09
MPCE	5.689e-05	5.975e-06	< 2e-16
Education (Primary education)	7.163e-02	5.815e-02	0.21821
Education (Secondary education)	2.163e-01	5.642e-02	0.00013
Education (Post-secondary education)	4.803e-01	8.408e-02	1.28e-08
Education (Tertiary education)	3.922e-01	8.520e-02	4.41e-06
Own any Agriculture Land (No)	3.805e-02	4.554e-02	0.40350
Sigma Coefficients Link function: log			
(Intercept)	-5.241e-01	5.642e-02	< 2e-16
Gender(female)	1.269e-01	5.860e-02	0.0305
MPCE	2.517e-05	4.004e-06	3.98e-10
Education (Primary education)	-2.129e-02	5.657e-02	0.7067
Education (Secondary education)	4.095e-02	5.154e-02	0.4270
Education (Post-secondary education)	7.763e-02	7.241e-02	0.2838
Education (Tertiary education)	-5.150e-03	8.000e-02	0.9487
Own any Agriculture Land (No)	1.344e-02	4.097e-02	0.7429
Nu Coefficients Link function: logit			
(Intercept)	0.9308000	1.487e-01	< 2e-16
Gender(female)	-0.5401000	1.493e-01	0.000304
MPCE	-0.0001599	1.633e-05	< 2e-16
Education (Primary education)	-0.2497000	1.258e-01	0.047254
Education (Secondary education)	-0.5728000	1.262e-01	6.03e-06
Education (Post-secondary education)	-0.7519000	2.005e-01	0.000182
Education (Tertiary education)	-0.6326000	2.108e-01	0.002727
Own any Agriculture Land (No)	0.1645000	1.003e-01	0.101147

5. Conclusion

The study's findings provide valuable insights into the factors influencing fruit consumption patterns among households, highlighting the roles of gender, income, education, and land ownership. Female-headed households tend to consume more fruit than male-headed households, suggesting a gender disparity in fruit consumption habits. This could be attributed to differences in food preferences, cultural norms, or access to resources. The higher fruit consumption rate in female-headed households underscores the importance of considering gender dynamics in health and nutrition interventions. Monthly per capita expenditure (MPCE) shows a positive association with fruit consumption, indicating that higher income levels are associated with higher fruit consumption. While the increase in consumption for each unit increase in MPCE is minimal, it suggests that economic factors play a role in shaping fruit consumption patterns. This finding highlights the importance of economic empowerment in improving fruit consumption and overall dietary diversity. Education emerges as a significant factor influencing fruit consumption patterns. Higher education levels are associated with higher fruit consumption, with a clear trend of increasing consumption with higher education levels. This suggests that education can positively influence dietary choices and nutritional outcomes. Policies aimed at improving access to education, particularly for women and marginalized groups, could help improve fruit consumption patterns and overall health. Contrary to expectations, owning agricultural land did not have a significant impact on fruit consumption. This suggests that factors other than land ownership, such as access to markets, knowledge about nutrition, and cultural practices, may play a more significant role in determining fruit consumption patterns. Understanding these factors is crucial for designing effective interventions to promote fruit consumption among households.

In conclusion, the study's findings suggest that gender, income, and education are key determinants of fruit consumption patterns among households. Policies and interventions aimed at promoting fruit consumption should consider these factors to effectively target populations and improve dietary diversity and nutrition outcomes. Further research is needed to explore the underlying mechanisms driving these associations and to develop tailored interventions to address fruit consumption disparities.

6. Appendix

```
library(gamless)
Model=gamlss(Fruit_data$Household.Consumption~Fruit_data$Gender+Fruit_data$MPCE+Fruit_d
ata$Education+Fruit_data$Own.any.Agriculture.Land,

sigma.formula=~Fruit_data$Gender+Fruit_data$MPCE+Fruit_data$Education+Fruit_data$Own.an
y.Agriculture.Land,

nu.formula=~Fruit_data$Gender+Fruit_data$MPCE+Fruit_data$Education+Fruit_data$Own.any.A
griculture.Land ,  family =ZAGA)
```

REFERENCE

1. NUTRIENT VALUES OF FRUIT AND VEGETABLES. Risk Assessment Studies Report No. 28 June 2007, Centre for Food Safety Food and Environmental Hygiene Department, The Government of the Hong Kong Special Administrative Region

2. Slavin JL, Lloyd B. (2012). Health Benefits of fruits and vegetables. *Adv Nutr*, 3(4): 506–16.

3. Dhandevi PEM, Rajesh JEEWON*2015 Oct; 44(10): 1309–1321. Fruit and Vegetable Intake: Benefits and Progress of Nutrition Education Interventions- Narrative Review Article

4. NITI Aayog National Multidimensional Poverty Index Baseline Report (2021) Delhi, India,

5. https://nalanda.kvk4.in/district-profile.php

6. National Sample Survey Office (NSSO) Level and Pattern of Consumer Expenditure 20112. NSS 68th Round (2014) July 2011-July 2012. NSS Report No. 555 (68/1.0/1), New Delhi, India

7. Rigby and Stasinopoulos, 2005R.A. Rigby, D.M. Stasinopoulos, Generalized additive models for location, scale and shape, Journal of the Royal Statistical Society, Series C, 54 (2005), pp. 507-554

8. Rigby, R. A., & Stasinopoulos, D. M. (2010). *A flexible regression approach using GAMLSS in R*. London Metropolitan University.

9. R.A. Rigby, D.M. Stasinopoulos, Generalized additive models for location scale and shape (GAMLSS) in R, Journal of Statistical Software, 23 (2007)