

**FACTORS AFFECTING THE PREFERENCE OF LOCAL  
AND IMPORTED MILK CONSUMPTION IN MATARA  
DISTRICT OF SRI LANKA: A STATISTICAL APPROACH**

S. D.M. Dilshani

(158877D)

Degree of Master of Science in Business Statistics (MBS)

Department of Mathematics

University of Moratuwa

Sri Lanka

October 2019

**FACTORS AFFECTING THE PREFERENCE OF LOCAL  
AND IMPORTED MILK CONSUMPTION IN MATARA  
DISTRICT OF SRI LANKA: A STATISTICAL APPROACH**

S. D.M. Dilshani

(158877D)

Dissertation submitted in partial fulfillment of the requirements for the  
degree Master of Science in Business Statistics

Department of Mathematics

University of Moratuwa

Sri Lanka

October 2019

## DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date:

The above candidate has carried out research for the Masters Dissertation under our supervision.

Name of the Supervisor: Professor L.A.L.W. Jayasekara  
Department of Mathematics  
Faculty of Science  
University of Ruhuna  
Sri Lanka.

Signature of the supervisor:

Date:

Name of the Supervisor: Mrs. H. V. S. De Silva  
Department of Mathematics  
Faculty of Engineering  
University of Moratuwa  
Sri Lanka.

Signature of the supervisor:

Date:

## **ABSTRACT**

Milk is one of the most essential foods to humans and it contains many nutrients such as protein, calcium, phosphorus, vitamin B2 and vitamin B12. Intake of a sufficient amount of milk products is recommended for healthy lifestyle of humans. As an agricultural country, Sri Lanka had become self-sufficient in milk, before adopting the open economic policies in 1977. Because of that, imported milk products were highly consumed since 1977 with very lower prices. The government and private sector data indicated that currently in Sri Lanka, local milk production can supply around 42% of the demand and the country depend on the imported milk powder. Therefore, this study was focused on the socioeconomic and other factors (based on the consumer's attitudes) which are influencing consumer's milk pattern either local milk or imported milk. In this study the data were collected through a consumer survey questionnaire in Matara district. At the beginning of the data analysis study, descriptive statistic and chi-square test of independence have done to identify the significant factors which are related with customer's milk consumption behaviors. Then, the Logistic Regression model was fitted on data using R software. Results from fitted multiple logistic regression model show that Age, Monthly Income, price of the milk, Easy to melt, artificial ingredient and Advertisements are the key determinants of consumers milk type.

Keywords: Milk consumption, Binary Logistic Regression, ROC Curve, Hosmer Lemeshow Goodness-of-fit Test

## **ACKNOWLEDGEMENT**

I would like to express my sincere gratitude to my supervisor Prof. L.A.L.W. Jayasekara for the continuous support of my study and research for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. Besides my supervisor advisor, I would like to thank Mrs. H.V.S. De Silva, for her encouragement, insightful comments. My sincere thanks also go to Prof. T. S. G. Peiris, for dissemination great knowledge to us on Statistics and Statistically Softwares. His guidance and encouragement for the success of this report is very much appreciated. I would like to thank all the lectures, who giving great knowledge to us in Master of Science in Business Statistics Degree. I wish to thank all my friends for their support and encouragement.

## TABLE OF CONTENTS

DECLARATION .....	i
ABSTRACT.....	ii
ACKNOWLEDGEMENT .....	iii
LIST OF FIGURES .....	vii
LIST OF TABLES.....	viii
CHAPTER 1 .....	1
INTRODUCTION .....	1
1.1    Background of the Study.....	1
1.1.1 Importance of the Milk .....	1
1.1.2. History of Milk in Sri Lanka .....	1
1.1.3. Milk Consumption Patterns in Sri Lanka .....	2
1.1.4. Milk Production in Sri Lanka .....	3
1.1.5 World Milk Production.....	7
1.2    Objectives.....	8
1.3    Outline of the Dissertation.....	8
CHAPTER 2 .....	10
LITERATURE REVIEW .....	10
2.1 Milk Consumption Pattern.....	10
2.2 Effect of Socioeconomics Characteristics on Milk Consumption Pattern.....	10
CHAPTER 3 .....	14
MATERIAL AND METHODS .....	14
3.1 Data Collection.....	14
3.1.1 Involved Variables in the Model Building Process .....	14
3.2 Methodology.....	17
3.2.1 Contingency Table.....	17
3.2.2 Chi-Square Test of Independence.....	18

3.2.3 Relative Risk .....	20
3.2.4 The Definition of the Odds.....	21
3.2.5 The Odds Ratio.....	21
3.2.5.1 Properties of Odds Ratio.....	22
3.2.6 Binary Logistic Regression.....	22
3.2.6.1 Use of the logistic curve.....	24
3.2.6.2 The Logistic regression Model .....	25
3.2.6.3 Significance of the Coefficients.....	27
3.2.7 Multiple Logistic Regression .....	30
3.2.7.1 The Multiple Logistic Regression Model.....	30
3.2.7.2 Fitting the Multiple Logistic Regression Model with Design Variables.....	32
3.2.7.3 Testing for the Significance of the Model.....	32
3.2.8 Assessing the Fitted Model.....	34
3.2.8.1 Hosmer Lemeshow Test.....	35
3.2.8.2 ROC Curve .....	36
3.2.9 Interpretation of the Fitted Logistic Regression Model.....	38
3.2.9.1 Interpretation Odds Ratio when Categorical Dichotomous Independent Variable .....	38
3.2.9.2 Interpretation of Odds Ratio when Categorical Polychotomous Independent Variable .....	41
3.2.9.3 Interpretation Odds Ratio when Continuous Independent Variables .....	42
CHAPTER 4 .....	43
RESULTS AND DISCUSSION .....	433
4.1 Descriptive Data Analysis .....	433
4.2 Univariate Analysis.....	49
4.3 Fitting a Logistic Regression Model .....	511
4.4 Model Selection Criteria .....	<b>Error! Bookmark not defined.</b> 8
4.5 Assessing the Fitted Model.....	<b>Error! Bookmark not defined.</b>
4.5.1. Hosmer Lemeshow Goodness-of-fit Test.....	59
4.5.2. ROC Curve .....	<b>Error! Bookmark not defined.</b>
4.6 Discussion.....	62
CONCLUSIONS .....	644

References.....	66
APPENDIX A: R Codes .....	68
APPENDIX B: Sample Questionnaire.....	744



## LIST OF FIGURES

Figure 1.1: Time Series Plot for Annual Milk Production in Sri Lanka from 1998 to 2017 .....	5
Figure 1.2: Population growth from 2007 to 2017(compared to previous year) in Sri Lanka. ....	5
Figure 1.3: Comparison of Prices of Imported milk powder and locally produced milk powder from 2011 to 2016 .....	6
Figure 1.4: World Milk Production in tonnes .....	7
Figure 3.1: Linear approximation to Logistic Regression Curve .....	24
Figure 3.2: Receiver Operating Characteristics (ROC) Curve .....	37
Figure 4.1: Bar Plot for type of Milk Consumption .....	433
Figure 4.2: Bar plot for Number of Family Members with Type of Milk .....	444
Figure 4.3: Bar plot for Monthly Income with Type of Milk .....	444
Figure 4.4: Bar plot for Type of Milk Consumption according to Education Level .....	455
Figure 4.5: Bar plot for Education Level of Household Head with Type of Milk Consumption .....	455
Figure 4.6: Bar plots for Type of Milk Consumption according to Consumer's opinion about their selected milk type .....	48
Figure 4.7: ROC curve .....	60

## LIST OF TABLES

Table 1.1: Annual Milk Production in Sri Lanka from 1998 to 2017 .....	4
Table 3.1: Description of the response variable and Predictor variables .....	186
Table 3.2: Contingency Table with Observed frequencies .....	18
Table 3.3: Expected frequencies .....	19
Table 3.4: Classification Table Based on the Logistic Regression Model .....	36
Table 3.5: Logistic Probabilities for the Dichotomous Independent Variable .....	39
Table 3.6: Coding of the Design variables for polychotomous independent variable using Reference Cell Coding with Level 1 as the reference group .....	411
Table 3.7: Specification of the Design variables for polychotomous independent variable using Reference Cell Coding with Level 1 as the reference group .....	42
Table 4.1: Results of Chi-Squared Test of Independence for Milk Type and Selected factors.....	49
Table 4.2: Binary Logistic Regression Model with all predictor variables.....	52
Table 4.3: Summary Table for the Model with Backward Elimination Method.....	57
Table 4.4: Comparison of the AIC values.....	59
Table 4.5: Summary measure of Hosmer-Lemeshow test.....	59
Table 4.6: Odds Ratios and 95% Confidence Interval to the odds ratio for the final fitted model...	61

## List of Appendix

APPENDIX A: R Codes .....	68
APPENDIX B: Sample Questionnaire.....	74