Advances in Computer Science Research, volume 91

International Conference on Modeling, Simulation and Big Data Analysis (MSBDA 2019)

Analysis of Influence Factors of Dietary Structure of Urban Residents in Beijing and Heilongjiang Areas —A Case Study of Grain and Aquatic Products

Li Sun[1], Danni Ma[1] and Xiuli Liu[2,3,*]

[1]School of Information, Beijing Wuzi University, Beijing, China

[2]Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

[3]University of Chinese Academy of Sciences, Beijing, China

[*] Corresponding author

Keywords: Food consumption, Dietary structure, Regression analysis.

Abstract. To explore the impacting factors of dietary structure of urban residents in Beijing and Heilongjiang provinces in China. This paper initially selected nine influencing factors to establish multiple regression models. According to the correlation analysis between the variables and their economic significance, the factors were finally included in the regression model are the population of 15-64 years old and urban residents' disposable income. The consumption of grain and aquatic products by urban residents in the two regions was affected by the disposable income of urban residents. The Heilongjiang region was more positively affected by the disposable income of urban residents. In the development of urban economy in Heilongjiang, it is necessary to improve the education level of urban residents and strengthen their understanding and attention on the population structure, nutrition and health of the population aged 15 to 64 in Beijing.

Introduction

The national health level is an important indicator to measure the social development of a country and the welfare of its residents, and it is receiving increasing attention in many fields of the world. General Secretary Xi Jinping's "Implementing a Healthy China Strategy" at the 19th National Congress was an important part of the basic strategy of national development.

A reasonable diet is a guarantee for people's health and an important factor in measuring the health level of a country. Nutrient imbalance can lead to nutritional diseases [1-2], that is, due to abnormal metabolism of substances in the body, the incidence of certain nutrient diseases in the population is increased, and common nutritional diseases include obesity, diabetes, and cardiovascular diseases.

At present, the world's top medical journal "Lancet" released the first large-scale heavy research in the global diet field - the mortality and disease burden caused by the diet structure of 195 countries and regions, showing the current nutritional health caused by diet structure. The issue has received much more attention.

Different countries have issued their own national dietary guidelines, and their standard intake is also based on dietary survey data. Different dietary studies are conducted by different groups at home and abroad to make the research results more targeted. For example, Brazil pays more attention to the diet patterns and nutritional status of adolescents [3-4], and their relations with obesity and cardiovascular diseases. Okinawa [5], as a longevity treasure, concerns with the diet and health of the elderly, which advocated a reduction in the risk of cardiovascular disease, certain cancers and other chronic diseases through a balanced diet.

To investigate the factors impacting the dietary structure of urban residents in Beijing and Heilongjiang, this paper established multiple regression models. With results analysis, some suggestions were provided which will be helpful to promote the rational diet t of the residents.

Data Sources

The data used in the multiple regression analysis of this paper are from the 1998-2012 China Statistical Yearbook, the data on the price index of grain and aquatic products in Beijing and Heilongjiang, the age structure of the population, the data on the education level of the population and the per capita disposable income of urban residents. Data related to the consumption of grain and aquatic products, in which the annual per capita disposable income of urban residents is the per capita disposable income of the year multiplied by the consumer price index (CPI).

Since the data trend of some variables is not significant enough, it is standardized and then included in the model. Before the regression analysis, by analyzing the Pearson correlation coefficient between variables and the corresponding significance test, the model first contains the preliminary selection of the variables with strong correlation of the dependent variables. Then eliminate the variables by different methods and choose explanatory variables according to the actual economic meaning.

Dietary Structure Influencing Factors Analysis Results

Variable Selection

Due to the large differences in economic levels between Beijing and Heilongjiang, Beijing (2017 per capita income ranks 2, cold regions), Heilongjiang (2017 per capita income ranking 19, severe cold regions) belong to the same climatic zone (temperate monsoon climate), income of two cities the level difference is large. And in the China Statistical Yearbook, the relevant data on the disposable income and food consumption of urban residents in Beijing are higher than the national average, while the data in Heilongjiang are behind the national average, in order to understand the residents of the two regions during the socio-economic transition period. The influencing factors of dietary structure, combined with the results of the differences in dietary structure between the two regions, were chosen to study the factors affecting the consumption of grain and aquatic products by urban residents.

In order to study the influencing factors of grain consumption of urban residents, this paper initially selected nine influencing factors X1-X9, and the variable Y1, carried out multiple regression analysis. The variables are defined in TABLE I. and TABLE II. The independent variables and dependent variables are defined, and all the variables are related data in Beijing area. The annual disposable income of urban per capita is the per capita disposable income of the year multiplied by the consumer price index (CPI). The data of X9 and Y1 are all standardized data. Data from Beijing or Heilongjiang are selected for analysis.

Table 1. Definition of independent variables and dependent variables (Grain)

x	Variable definition content	x	Variable definition content
X1	The proportion of grain consumption of urban residents in T-2 (%)	**X2**	Grain price index (%)
X3	0-14 years old (%)	**X4**	15-64 years old (%)
X5	Proportion of over 65 years old (%)	**X6**	Junior high school and below rate (%)
X7	High school rate (%)	**X8**	College and above rate (%)
X9	Annual per capita disposable income of urban (yuan)	**Y1**	Urban residents' grain consumption (yuan)

Since the data sample of Beijing urban residents is small, the independent variables of 2-3 are the most suitable. It is necessary to combine the economic meaning and the correlation analysis between variables to select the independent variables. It is not difficult to see that X3-X5 is the relevant variable of the age structure of the population. X6-X8 is the relevant variable of education level. According to its corresponding economic significance, it is more scientific to select one of the variables when constructing the regression model.

Table 2. Definition of independent variables and dependent variables (Aquatic Products)

x	Variable definition content	x	Variable definition content
X1	The proportion of Aquatic Products consumption of urban residents in T-2 (%)	**X2**	Aquatic Products price index (%)
X3	0-14 years old (%)	**X4**	15-64 years old (%)
X5	Proportion of over 65 years old (%)	**X6**	Junior high school and below rate (%)
X7	High school rate (%)	**X8**	College and above rate (%)
X9	Annual per capita disposable income of urban (yuan)	**Y1**	Urban residents' grain consumption (yuan)

Table 3. Pearson correlation analysis between variables

		X1	X2	X3	X4	X5	X6	.	Y1
X1	**Pearson correlation**	1	-.733**	.713**	-.309	-.451	.561*		-.393
	Significant (two-tailed)		.004	.006	.304	.122	.046		.184
X2	**Pearson correlation**	-.733**	1	-.758**	.501	.200	-.429		.507
	Significant (two-tailed)	.004		.003	.081	.513	.143		.077
X3	**Pearson correlation**	.713**	-.758**	1	-.791**	-.051	.717**		-.808**
	Significant (two-tailed)	.006	.003		.001	.869	.006		.001
X4	**Pearson correlation**	-.309	.501	-.791**	1	-.571*	-.782**		.963**
	Significant (two-tailed)	.304	.081	.001		.042	.002		.000
X5	**Pearson correlation**	-.451	.200	-.051	-.571*	1	.315		-.487
	Significant (two-tailed)	.122	.513	.869	.042		.295		.091
X6	**Pearson correlation**	.561*	-.429	.717**	-.782**	.315	1		-.795**
	Significant (two-tailed)	.046	.143	.006	.002	.295			.001
…									
Y1	**Pearson correlation**	-.393	.507	-.808**	.963**	-.487	-.795**		1
	Significant (two-tailed)	.184	.077	.001	.000	.091	.001		

**. At the 0.01 level (two-tailed), the correlation is significant.
*. At the 0.05 level (two-tailed), the correlation is significant.

In order to determine whether the model uses multiple linear regression models, the Pearson correlation analysis of the nine influencing factors and the dependent variable Y1 is carried out to explore the relationship between them. The results are shown in TABLE III. below.

Correlation analysis obtained the Pearson correlation coefficient and the corresponding significance test, it can be seen that there is a significant correlation between X1, X4, X6, X7, X8, X9 in Y1 and 9 factors, and thus linear regression analysis can be performed. However, due to the high correlation within the factor, there is a multicollinearity problem, and the sample size is less than 30. The regression method selects stepwise regression, and the whole model is eliminated the variables one by one to obtain the optimal multiple regression equation.

The Establishment and Testing of Multiple Linear Regression Models (Taking the Grain Consumption of Urban Residents in Beijing as an Example)

According to the correlation analysis between the variables and the economic significance of the research, X4: 15-64 years old (%), X8: college and above (%), X9: urban residents disposable income (yuan) as regression. The independent variables of the model, and Y1: urban residents' grain consumption (yuan) in the SPSS statistical software for stepwise regression analysis, the elimination of variables and the established model are summarized in TABLE III.

Table 4. Input/Removal Variablesa

Model	Input variable	Removed variable	Method
1	Table 1-3 X9: Urban disposable income (yuan) - standardization, X4: 15-64 years old (%), X8: college and above rate (%)b	.	Input
2	.	X8: College and above rate (%)	Backward (condition: probability of F to be removed >= .100).

a. Dependent variable: Y1: grain consumption of urban residents (yuan) - standardization

b. All variables requested have been entered.

TABLE IV. Shows the modeling process. It can be seen that Model 1 is the full model of explanatory variables X4, X8, and X9. After stepwise regression analysis, the explanatory variable X8: the probability of college and above (%) is ≥ 0.1. According to the model rules, it was removed and obtained model 2.

Then the multiple linear regression model with X4 and X9 as explanatory variables and Y1 as the dependent variable. TABLE V. shows that the goodness of fit of Model 2 is R^2=0.977, which works well.

Table 5. Model Summary

Model	R	R^2	Adjusted R^2	Standard estimated error
1	.988a	.977	.969	.03661
2	.988b	.977	.972	.03498

a. Predictors: (constant), X9: Urban disposable income (yuan) - Standardization, X4: 15-64 years old (%), X8: College and above (%)

b. Predictors: (constant), X9: Urban disposable income (yuan) - Standardization, X4: 15-64 years old (%)

It can be seen from TABLE VI. That the F-test value of Model 2 is 207.952, and the P value is close to 0. The significance test indicates that the linear relationship of the model is significant. The p-value of the regression coefficient t-test in model 2 is <0.05 (α=0.05). The significance test shows that the regression coefficient of the model has a linear relationship with the dependent variable. When the multi-collinearity test is performed, the tolerance of the model 2 is Above 0.1 and VIF below 10, it is possible to pass the test without serious multicollinearity.

Table 6. Analysis of variance

Model		sum of squares	Degree of freedom	Mean square	F	Sign.
1	**regress**	.509	3	.170	126.542	.000[b]
	Residual	.012	9	.001		
	total	.521	12			
2	**regress**	.509	2	.254	207.952	.000[c]
	Residual	.012	10	.001		
	total	.521	12			

a. Dependent variable: Y1: grain consumption of urban residents (yuan) – standardization

b. Predictors: (constant), X9: Urban disposable income (yuan) - standardization, X4: 15-64 years old (%), X8: College and above (%)

c. Predictors: (constant), X9: Urban disposable income (yuan) - standardization, X4: 15-64 years old (%)

Summary of Multivariate Linear Regression Models for Grain and Aquatic Products

Analysis of Influencing Factors on Grain Consumption of Urban Residents in Beijing. The final factors included in the regression model were the proportion of X4:15-64 years old (%) and X9: disposable income of urban residents (yuan), with regression coefficients of 0.041 and 0.586, respectively.

The optimal multiple linear regression equation was obtained: $y_1=-2.903+0.041x_4+0.586x_9$.

The consumption of grain by urban residents in Beijing is affected by the proportion of people aged 15-64 and the disposable income of urban residents, both of which are positive. And the positive impact of disposable income of urban residents is greater.

Analysis of Influencing Factors on Grain Consumption of Urban Residents in Heilongjiang Province. The final factors included in the regression model were X: proportion over 65 years old (%) and X9: disposable income of urban residents (yuan). The regression coefficients were -0.023 and 1.063, respectively.

The optimal multiple linear regression equation was obtained: $y_3 = 0.186-0.023x_5 +1.063x_9$.

The consumption of grain by urban residents in Heilongjiang Province is affected by the proportion of over 65 years old and the disposable income of urban residents. The disposable income of urban residents has a positive impact, while the proportion of over 65 years old has a negative impact, and the positive impact is greater than the negative impact.

Analysis of Influencing Factors on Aquatic Products Consumption of Urban Residents in Heilongjiang Province. The final factors included in the regression model were X2: aquatic product price index (%) and X9: disposable income of urban residents (%). The regression coefficients were -0.02 and 0.729, respectively.

The optimal multiple linear regression equation was obtained: $y_2 = 0.461-0.002x_2 + 0.729x_9$.

In conclusion, the consumption of aquatic products by urban residents in Beijing is affected by the price index of aquatic products and disposable income of urban residents. Aquatic product price index has a negative impact, and the positive impact is greater than the negative impact.

Analysis of Influencing Factors on Aquatic Products Consumption of Urban Residents in Heilongjiang Province. The final factors included in the regression model were X6: the rate below junior middle school (%) and X9: the disposable income of urban residents (yuan). The regression coefficients were -0.007 and 0.969, respectively.

The optimal multiple linear regression equation was obtained: $y_4 = 0.559-0.007x_6 + 0.969x_9$.

The consumption of aquatic products by urban residents in Heilongjiang Province is affected by the rate below junior middle school and the disposable income of urban residents. The rate below junior middle school has a negative impact, and the positive impact is greater than the negative impact.

Suggestion

From the regression analysis, it is found that the consumption of grain and aquatic products by urban residents in the two regions is affected by the disposable income of urban residents. The grain consumption of urban residents in Heilongjiang Province is the most positively affected by the disposable income of urban residents, followed by the consumption of aquatic products by urban residents in Heilongjiang. It is recommended to develop the urban economy and improve the educational level of urban residents. At the same time, we will increase the awareness of nutrition and diet, and promote the intake of aquatic products in Heilongjiang.

The consumption of grain by urban residents in Beijing is affected by the proportion of people aged 15-64 and the disposable income of urban residents, and both have a positive impact. The consumption of aquatic products by urban residents in Beijing is affected by the price index of aquatic products and the disposable income of urban residents. The disposable income of urban residents has a positive impact, and the price index of aquatic products has a negative impact. It is recommended to develop the urban economy while paying attention to the dietary structure and

eating habits of the 15-64-year-old population, and to enhance the propaganda of relevant dietary nutrition knowledge.

Acknowledgment

This project was funded by 2018 "Practical Training" Project of High-level Talents Cross-cultivation in Beijing Universities and Colleges, the Key Project of the Chinese Academy of Sciences (ZDBS-SSW-DQC): National Nutrition and Food Security, Beijing Intelligent Logistics System Collaborative Innovation Center (PXM2018_014214_000009), Beijing Key Laboratory (BZ0211), Quality Construction of Talents Training in 2018-Specialty Construction-Specialty Group Construction in Beijing Universities (Municipal Level) (PXM2018_014214_000005).

References

[1] Y. L. Hong, F. E. Jiang, X. Jin, F. Lv. “Research status and prospect of dietary nutrition of residents at home and abroad,” Processing of Agricultural Products, vol.02, pp.66-70, Jan 2019.

[2] J. M. BOURRE, “Effects of nutrients (in food) on the structure and function of the nervous system: update on dietary requirements for brain. Part 1: micronutrients,” Journal of Nutrition Health & Aging, vol.10(5), pp.386–399. Sep-Oct 2006.

[3] E. C. Cristina, S. Betzabeth, “Dietary intake of adolescents compared with the Brazilian Food Guide and their differences according to anthropometric data and physical activity,” Revista Brasileira de Epidemiolo, vol.18(4), pp. 798–808, Jan 2015.

[4] Y. Matsumura, “Nutrition trends in Japan,” Asia Pacific Journal of Clinical Nutrition, vol.10, pp.40-47, Apr 2015.

[5] J. D. Wright and C. Y. Wang, “Trends in intake of energy and macronutrients in adults from 1999-2000 through 2007-2008,” NCHS Data Brief, vol.49, pp.1-8, Nov 2010.

[6] Q. Liu, X. L. Liu, S. Y. Wang, “Estimating China’s food grains demand from 2020 to 2050 based on reasonable dietary pattern,” Systems Engineering — Theory & Practice, vol.38(03), pp.615-622, Mar 2018.