The analysis of factors that affecting household food security of tuberculosis patients in Surabaya

D Susilaningrum*, B S S Ulama, S P Wulandari and M Salamah

Business Statistics Department, Faculty of Vocational, Institut Teknologi Sepuluh Nopember Kampus Keputih Sukolilo Surabaya, 60111, Indonesia
*Corresponding Author: destr.s@gmail.com

Abstract. Tuberculosis (TB) cases and household food security are fundamental problem which can be concerned with each other. Household food insecurity can affect the ability of a family member in counteracting an infectious disease such as TB. This research is aimed to study the factors that affecting household food security of tuberculosis patients in Surabaya. By Binary logistic regression model, there are four significant variables; Head of Household’s Occupation, Expenditure per Month, The Household Density and Type of Roof, with the correct classification 65.6%.

1. Introduction
Poverty, food security and malnutrition are major concerns among international health when all people have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life [1]. There are four food security indicators used to classify each household [2], those are:
- Food Availability
- Food Stability
- Food Access
- The Quality of Food Security

where, the categories of household food security can be assigned by combining two aspects, between continuity of food availability and quality of food [2], which is presented in Table 1.

| Continuity of Food Availability | Quality of Food         |
|---------------------------------|-------------------------|
| Continous                       | Food-secure             |
| Not continuos                   | Food-insecure           |

Table 1. Classified of Food Security for Each Households

Food security are fundamental problem which can be concerned with a public health. Household food insecurity has been associated with a bad ability in counteracting diseases, i.e. Tuberculosis. Tuberculosis is the type of disease that spread directly, caused by bacteria that spread through the air. TB has become the third highest cause of death in the world. According to data from the East Java Provincial Health Office, the largest TB cases in Indonesia were in East Java and Surabaya accounted for the largest number, which is 3093 out of 33999 cases in East Java in 2017 [3].

This research was conducted to study the factors that affecting variable on household food security of tuberculosis patients in Surabaya. Binary Logistic Regression was used to analyze the model in this research.
2. Material and Method

2.1 Binary logistic regression

Binary logistic regression is a method of analysis that is used to find out the relationship between the response variables (Binary or dichotomous) with the predictor variables that are polychotomous [4]. The response variable (y) consists of two categories: "success" (y = 1) and "failure" (y = 0). The model logistic regression given as:

$$
\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_j x_j + \ldots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \ldots + \beta_j x_j + \ldots + \beta_p x_p)}
$$

The Equation 2.1 can be explained as a logit model: by logit transformation of \( \pi(x) \) as:

$$
g(x_i) = \ln\left( \frac{\pi(x_i)}{1-\pi(x_i)} \right)
$$

$$
g(x_i) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_j x_{ij} + \ldots + \beta_p x_{ip} = \sum_{j=0}^{p} \beta_j x_{ij}
$$

The test is conducted to get the best model which was built by the significant parameters. Parameters were first tested simultaneously and then tested partially to get the significant parameters.

2.1.1 Simultaneously test of parameters

Hypothesis:

\( H_0 : \beta_i = \beta_2 = \ldots = \beta_p = 0 \)
\( H_1 : \) at least one of \( \beta_j \neq 0 ; j = 1, 2, \ldots, p \)

Test Statistic:

$$
G = -2 \ln \left[ \frac{\left( \frac{n_1}{n} \right)^{n_1} \left( \frac{n_0}{n} \right)^{n_0}}{\sum_{i=1}^{n} \hat{\pi}_i^{y_i} (1-\hat{\pi}_i)^{1-y_i}} \right]
$$

(3)

With \( n_1 = \sum_{i=1}^{n} y_i ; n_0 = \sum_{i=1}^{n} (1-y_i) ; n = n_1 + n_0 \)

Reject \( H_0 \) if \( G > \chi^2_{(\alpha, df)} \)

2.1.2 Partial test of parameters

Hypothesis:

\( H_0 : \beta_j = 0 \)
\( H_1 : \beta_j \neq 0 ; j = 1, 2, \ldots, p \)

Test Statistic:

$$
W_j^2 = \frac{\hat{\beta}_j^2}{[SE(\hat{\beta}_j)]^2}
$$

(4)

Reject \( H_0 \) if \( W_j^2 > \chi^2_{(\alpha, df)} \)

2.1.3 Goodness of fit model

The test for overall fit of a Binary logistic regression model with the hypothesis:

\( H_0 : \) Model Fit
\( H_1 : \) Model Not Fit
Test Statistic:

\[
\hat{C} = \sum_{i=1}^{g} \left( \frac{\sum_{j=1}^{n_i} y_{ij} - \sum_{j=1}^{n_i} \hat{p}_{ij}}{1 - \frac{\sum_{j=1}^{n_i} \hat{p}_{ij}}{n_i}} \right)^2 \tag{5}
\]

\(y_{ij}\) denote the Binary outcome for observation \(j\) in group \(i\) of the partition; \(i = 1, \ldots, g; j = 1, \ldots, n_i\)

\(\hat{p}_{ij}\) denote the corresponding fitted probability for the model fitted to the ungrouped data [5]

Reject \(H_0\) if \(\hat{C} > \chi^2_{g(2)}\) or P-value < \(\alpha\)

2.1.4 Classification procedure

This procedure was used to evaluate the result of prediction value given by the best model to compare with the observation value, [6] give an evaluation on classification procedure to see the probability of miss classification. It is measured by apparent error rate (APER). APER value stated the proportion value of miss classification sample by the function of classification. If the subject only classified as two groups, \(y_1\) and \(y_2\), then determination of classification errors can be known through the classification table described in Table 2.

| Table 2. Classification Table |
|-------------------------------|
| Observation | Prediction |
| --- | --- | --- |
| \(y_1\) | \(n_{11}\) | \(n_{12}\) |
| \(y_2\) | \(n_{21}\) | \(n_{22}\) |

Description:

\(n_{11}\): The number of the subject of \(y_1\) on the correct classified as \(y_1\)

\(n_{12}\): The number of the subject of \(y_1\) on the miss classified as \(y_2\)

\(n_{21}\): The number of the subject of \(y_2\) on the miss classified as \(y_1\)

\(n_{22}\): The number of the subject of \(y_2\) on the correct classified as \(y_2\)

\[
APER(\%) = \frac{n_{12} + n_{21}}{n_{11} + n_{12} + n_{21} + n_{22}} \times 100\% \tag{6}
\]

And the proportion of the correct classification = 1 - APER (\%) = \(\frac{n_{11} + n_{22}}{n_{11} + n_{12} + n_{21} + n_{22}}\) 100%

2.2 Source of data and research variables

This research used primary data which was collected by conducting survey using a random sampling technique [6] to 259 respondents taken from 3093 TB patients in Surabaya [3].

The research variables consist of three kind variables: Respond variable, Indicator variable of Food Security and Predictor variable.

2.2.1 Respond variable

Respond Variable (\(Y\)) in this research is Household food security Status of TB Patient in Surabaya as given in Table 3.

| Table 3. Respond Variable |
|---------------------------|
| Variable | Description |
| --- | --- |
| \(Y\) | \(Y = 0 =\) Secure or \(Y = 1 =\) Insecure |
2.2.2 Indicator variable of food security
The food security status to be determined by four indicators that gives in Table 4,
Table 4. Indicator Variable to determine Food Security Status

| Indicator               | Description                                      |
|------------------------|--------------------------------------------------|
| Sufficiency Food Availability | The Stocks of Rice for One Month               |
| Food Stability         | Meal Frequency Household Members per Day         |
| Accessibility of Food   | Market Location                                  |
|                        | The Number of Household Member                   |
|                        | Head of Household’s Education                    |
|                        | How to Obtain the Food                           |
| The Quality of Food Security | Protein Consumed                              |

2.2.3 Predictors variable (X)
Predictor variable (X) gives in Table 5.
Table 5. Predictor Variable

| Variable | Description                                      |
|----------|--------------------------------------------------|
| X₁       | Head of Household’s Occupation                   |
| X₂       | Work Status of Wife                              |
| X₃       | Number of School Age Children                    |
| X₄       | Number of Toddler                                |
| X₅       | The Family Income per Month                      |
| X₆       | Expenditure per Month                            |
| X₇       | The Ownership of House Status,                   |
| X₈       | The Household Density                            |
| X₉       | Type of Roof                                     |
| X₁₀      | Type of Wall                                     |
| X₁₁      | Type of Tile                                     |
| X₁₂      | Broad of Ventilation                             |
| X₁₃      | Toilet Ownership                                 |
| X₁₄      | Water Source                                     |
| X₁₅      | Littering Places                                 |
| X₁₆      | Landfill                                         |
| X₁₇      | Electricity Source                               |

3. Results and Discussion
3.1 The characteristics of household security of TB patient in Surabaya
Figure 1 shows that more than 50% households are food insecure in Surabaya. 58.7% of the TB patient household are food insecure and 41.3% of them are food secure.
3.2 Analysis binary logistic regression model
The full logit model of Binary Logistics Regression which include all variables gives as below.

\[ g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_{17} x_{17} = \sum_{j=0}^{17} \beta_j x_j ; i = 1, 2, ... , 259 ; j = 1, 2, ... , 17 \]

3.2.1 Simultaneously test of parameter
To find out which parameters are significant, the simultaneously test with the hypothesis as in subsubsection 2.1.1 had done and give the result on Table 6.

| Table 6. Simultaneous test results |
|------------------------------------|
| \( \chi^2 \) | df | P-value |
|-----------------|---|-------|
| 26.479          | 4 | 0.000 |

For \( \alpha = 0.05 \), it showed that P-value = 0.000 less than \( \alpha \) indicated that reject \( H_0 \), the partial test was then conducted to find out the variables that significant in the model.

3.2.2 Partially test of parameter
The partial test with the hypothesis as in subsubsection 2.1.2, give the result as shown in Table 7.

| Table 7. Partial Test Results |
|-------------------------------|
| Variables | B   | Wald | df | P-value |
|-----------------|-----|------|----|---------|
| X_1: Head of Household’s Occupation | 0.210 | 5.484 | 1  | 0.019 |
| X_6: Expenditure per Month | -1.560 | 8.832 | 1  | 0.003 |
| X_8: The Household Density | -0.778 | 8.165 | 1  | 0.004 |
| X_9: Type of Roof | -0.719 | 6.334 | 1  | 0.012 |
| Constant | 0.714 | 6.658 | 1  | 0.010 |

Table 7 shows some of P-value less than \( \alpha = 0.05 \), that indicate the variables are significant to the model, and the variables are: Head of Household’s Occupation, Expenditure per Month, The Household Density, and Type of Roof. The logit Binary logistics Regression model gives as below.

\[ g(x) = 0.714 + 0.210 X_1 - 1.560 X_6 - 0.778 X_8 - 0.719 X_9 \]

3.2.3 Goodness of fit model
The hypothesis of goodness of fit model as in subsubsections 2.1.3, give the result as shown in Table 8.

| Table 8. Goodness Fit of Model Results |
|----------------------------------------|
| \( \chi^2 \) | df | P-value |
|------------------|----|--------|
| 7.486            | 8  | 0.485  |

This table showed that P-value more than \( \alpha = 0.05 \), indicated the test cannot reject \( H_0 \), and so the model fit.

3.2.4 Classification procedure
The subject only classified as two groups, with \( Y_1 \): Secure and \( Y_2 \): Insecure. The determination of classification errors can be known through the classification table described in Table 9.

| Table 9. Classification Table |
|-----------------------------|
| Observation | Prediction | Percentage |
|-------------|-----------|-------------|
| Secure      | 52        | 55          | 48.6        |
| Insecure    | 33        | 116         | 77.9        |
| Correct Classification | 65.6 |
Based on the results above, the factors affecting household food security are: Head of Household’s Occupation, Expenditure per Month, The Household Density and Type of Roof with the correct classification between predictor and observation 65.6%.

4. Conclusion
The factors that affecting household food security of TB patients in Surabaya are: Head of Household’s Occupation, Expenditure per Month, The Household Density and Type of Roof, with the correct classification between predictor and observation 65.6%.

Acknowledgement
The authors would like to express their gratitude to the LPPM -ITS and DRPM of The Ministry of Technology Research and Higher Education of Indonesia which has sponsored this research.

References
[1] FAO 2008 Food Security Information for Action Practical Guides (New York: The EC-FAO Food Security Programme)
[2] Puslit Kependudukan-LIPI 2009 Konsep dan Ukuran Ketahanan Pangan Rumah Tangga di Pedesaan (Jakarta: Puslit Kependudukan-LIPI)
[3] East Java Provincial Health Office 2018 Profil Kesehatan Provinsi Jawa Timur2018 (Surabaya: Dinkes Prov.Jatim)
[4] Hosmer D W and Lemeshow S 2013 Applied Logistic Regression (New York: John Wiley & Sons)
[5] Agresti, Alan 2002 Categorical Analysis Second Edition (New Jersey: John Wiley & Sons, Inc)
[6] Johnson R A and Winchern D W 2007 Applied Multivariate Statistical Analysis (6th ed.) (New Jersey: Pearson Prentice Hall)
[7] Susilaningrum D 2018 Teknik Pengambilan Sampel untuk Peneliti dan Pemula (Surabaya: Kanzun Books)