JSDS: JOURNAL OF STATISTICS AND DATA SCIENCE
VOLUME 3, No 2, October 2024
e-ISSN: 2828-9986
https://ejournal.unib.ac.id/index.php/jsds/index

Analysis of the Quality of Health Service at the Air Haji Hearth Center Using the Ordinal Logistics Regression Method

Annisa Soleha[1*], Devni Prima Sari [2]
[1] Department of Mathematics, Universitas Negeri Padang, Indonesia
* Corresponding Author : Annisaagusti12345@gmail.com

Article Info

Article History:
Received: August, 19 2024
Accepted: September, 2 2024
Available Online: October, 30 2024

Key Words:
Patient satisfaction,
Service quality,
Health center,
Ordinal logistic regression.

Abstract

Improving the quality of public services has become a major concern in government agencies as an effort to provide optimal public services. The quality of service can be affected by various factors. Therefore, it is necessary to conduct an analysis to find out the relationship between factors that affect service quality and service quality itself. Efforts are made to analyze the relationship between factors that affect service quality and service quality itself by using the ordinal logistic regression method in analyzing the relationship between influencing factors and influencing factors. The purpose of this study is what factors significantly affect the quality of services at the Air Haji Health Center. Based on the results of data analysis, it was found that the variables that significantly affected the quality of service were direct evidence variables, guarantee variables, and empathy variables. This research is useful for the Air Haji health center in an effort to improve the quality of health services.

1. INTRODUCTION

Service quality is a comparison between the reality of the service received and the expectation of the service to be received [1]. Service quality can measure the success of government organizations or institutions as public service providers. Improving the quality of public services has become a major concern in government agencies as an effort to provide optimal public services [2]. Service quality can be influenced by several factors, to see what factors affect service quality, an analysis can be carried out to find out the relationship between influencing factors and service quality.

Another method that is often used in an effort to analyze the relationship between independent variables and dependent variables is the probit model. The probit model is one of the statistical methods used to analyze data with ordinal category dependent variables. This model uses the normal distribution function to determine the probability of each category [3]. However, there are drawbacks in the analysis using the probit model, namely limitations in interpretation, the probit model often produces parameters that are difficult to interpret [4]. The probit model is more complex compared to some other models, such as the ordinal logit model. In the context of health care quality analysis, it is important to consider these shortcomings and perhaps also explore other alternative methods such as the ordinal logit model, to ensure better and easier interpretation of the analysis results [5].

Ordinal logistic regression is a regression method that can be used to describe the relationship between a dependent variable (Y) and one or more independent variables (X), where the variable is bound to more than two categories and the measurement scale is hierarchical and aims to determine the factors that affect the bound variable Y and produce an odds ratio related to the value of each independent variable [6]. The goal of this method is to obtain a good and simple model that depicts a response variable with a set of predictor variables.

2. ORDINAL LOGISTIC REGRESSION

Logistic regression is a method that can be used to find the relationship between a response variable that is dichotomous (nominal/ordinal scale with two categories) or polychotomous (nominal/ordinal scale with more than two categories) with one or more categorical or continuous scale predictor variables [7].

The model form of logistic regression is:

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x_1}}{1 + e^{\beta_0+\beta_1 x_1}}$$

By using the logit transformation of (x) to make it easier to estimate the formulated regression parameters is:

$$g(x) = In\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1$$

Ordinal logistic regression is a statistical method that describes the relationship between a dependent variable (*Y*) and one or more independent variables (*X*), where the dependent variable has more than two categories and the measurement scale is hierarchical [8]. The model used for ordinal logistic regression is the cumulative logit model. The cumulative logit model is a model obtained by combining cumulative opportunities, namely less or equal opportunities than the response category with greater opportunities than the response category [9].

$$g_j(x) = In\left[\frac{P(Y \le j|X_i)}{P(Y > j|X_i)}\right] = \beta_{0j} + \beta_1 x_1 + \cdots + \beta_k x_i$$

Ordinal logistic regression has several assumptions, including: Bound variables or dependent variables must be ordinal scaled, Independent variables can be continuous, ordinal or categorical, There is no multionlinearity, and There is an odds ratio [10].

The maximum Likelihood Estimator (MLE) method is used to estimate parameters in logistic regression which basically provides an estimation value by maximizing its likelihood function. The likelihood function describes the probability of observation data as a function of unknown parameters, so it is suspected that the logistic parameter is known first as the likelihood function. Systematically the likelihood function is written with [11].

$$l(\boldsymbol{\beta}) = \prod_{i=1}^{n} [\pi_0(x_i)^{\,y_{0i}}\, \pi_1(x_i)^{\,y_{1i}} \pi_2(x_i)^{\,y_{2i}}]$$

The maximum likelihood method is the value of the parameter estimator by maximizing the likelihood log function. By differentiating the log likelihood form to the $\beta_1, \beta_2, \cdots, \beta_k$ and equalize it to zero.

3. METHOD

The type of research applied is applied research that begins with theoretical analysis and data collection, then ordinal logistic regression analysis is carried out. The ordinal logistic regression method is a method for determining the relationship between dependent variables and response variables. The data used is primary data obtained from the results of distributing questionnaires to visitors at the Air Haji health center. Factors that are suspected of influencing service quality are direct evidence, reliability, responsiveness, assurance, and empathy. In problem solving, ordinal logistic regression is assisted using R software.

The research procedures applied in this research are:

1. Sampling Techniques
   The sampling technique used in this study is a survey using a questionnaire containing a number of structured written questions to obtain information from respondents.
   The sampling steps are [12]:
   a. Making a sample framework in this study that will be the sample framework is name, age, gender.
   b. Determining the overall sample volume using the slovin formula
   c. Distributing questionnaires to visitors found at the time of the study were 200 people.
2. Research Instruments
   Research instruments are tools needed for researchers in collecting data. The research instrument used in this study is a questionnaire. The steps are:
   a. Making an instrument grid
   b. Designing research instruments in the form of questionnaires

c. Perform instrument validation and reality tests

3. Data Analysis Techniques
   The stages of the analysis:
   a. Collect questionnaire answers that have been filled out by respondents
   b. Checking the data and scoring the answers
   c. Calculate the average on each independent variable
   d. Estimating parameters using the Maximum Likelihood Estimator (MLE) method
   e. Forming a model of logistic regression speculation between the bound variable (*Y*) and the independent variable (*X*)
   f. Conducting a coefficient significance test of the logistic regression model using the G test
   g. Conducting an individual coefficient significance test to determine the independent variable that has a significant effect on the bound variable using the Wald test
   h. The formation of the best model is based on the coefficient significance test of the model that has been carried out.
   i. Conducting multilinearity assumption tests
   j. The purpose of the study was to test whether the regression model found a correlation between independent variables.
   k. Interpretation of the model using odds ratio
      The ordinal logistic regression model has been tested and the results of the model are good. The significance is real, then the data can be interpreted using the odds ratio

4. RESULTS AND DISCUSSION

4.1 Data Deskription

The following is a description based on the variables in tabulation obtained from the distribution of the questionnaire. Table 1 is the percentage of respondents to five independent variables out of the five variables studied, it can be seen that the highest level of respondents feel agree and strongly agree with the responsiveness variable of 94.6% and the lowest is the direct evidence variable of 81.50%.

Table 1. Description Research Results

No	Independen Variable (X)	Consent Level (%)			
		Strongly aggre	Agree	Disagree	Strongly disagree
1	Direct Evidence	13.4	75.3	9.6	1.6
2	Reliability	30.1	63.3	6.3	0
3	Responsiveness	29.9	64.7	5.5	0
4	Guarantee	17.5	74	8.5	0
5	Empathy	26.6	64.7	8.8	0

Table 2. Description Research Results

No	Y Category	Percentage
1	Very unsatisfactory	3.6
2	Unsatisfactory	12.1
3	Neutral/standard	29.6
4	Satisfying	36.7
5	Very satisfying	18.1

In the dependent variables studied, it can be seen that the highest level of service quality of respondents from the five categories above is satisfactory at 36.6% while the lowest is very unsatisfactory at 3.6%.

4.2 Data Analysis Result

1. Estimator Parameters

Parameter β in the ordinal logistic regression model is estimated using the maximum likelihood method (MLE) by maximizing the likelihood function. [13].

Table 3. Predicted Parameter Results

Variable	Variable	β
Dependent	[Y=1]	-4.325
	[Y=2]	2.095
	[Y=3]	5.585
	[Y=4]	8.484
Independent	$X_{1(2)}$	-8.202
	$X_{1(3)}$	-3.883
	$X_{1(4)}$	-0.853
	$X_{2(3)}$	1.663
	$X_{2(4)}$	1.255
	$X_{3(3)}$	-1.100
	$X_{3(4)}$	-0.636
	$X_{4(3)}$	3.486
	$X_{4(4)}$	6.254
	$X_{5(3)}$	5.424
	$X_{5(4)}$	5.444

Results of Table 3 an ordinal logistic regression model with all independent variables is obtained:

$$g_1(x) = -4.325 - 8.202x_{1(2)} - 3.883x_{1(3)} - 0.853x_{1(4)} + 1.663x_{2(3)} + 1.255x_{2(4)} - 1.100x_{3(3)} - 0.636x_{3(4)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(4)}$$

$$g_2(x) = 2.095 - 8.202x_{1(2)} - 3.883x_{1(3)} - 0.853x_{1(4)} + 1.663x_{2(3)} + 1.255x_{2(4)} - 1.100x_{3(3)} - 0.636x_{3(4)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(4)}$$

$$g_3(x) = 5.585 - 8.202x_{1(2)} - 3.883x_{1(3)} - 0.853x_{1(4)} + 1.663x_{2(3)} + 1.255x_{2(4)} - 1.100x_{3(3)} - 0.636x_{3(4)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(4)}$$

$$g_4(x) = 8.484 - 8.202x_{1(2)} - 3.883x_{1(3)} - 0.853x_{1(4)} + 1.663x_{2(3)} + 1.255x_{2(4)} - 1.100x_{3(3)} - 0.636x_{3(4)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(4)}$$

After the results of the parameter β estimation are obtained, they can then be tested using coeffecient significance testing model.

2. Coeffecient Significance Testing Model

The coefficient significance test of the model is useful to check whether the independent variable in the significant model has a real effect on the bound variable [14].

The significance test of the model by testing all independent variables on the model simultaneously using concurrent test.

Hypothesis:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_j = 0$
$H_1: there\ is\ at\ least\ one\ \beta_j \neq 0$

Table 4. Simultan Test

G	χ^2
426.501	11.070

At the Table 4, it can be seen that the logistics value of the G test obtained is greater than the value of $\chi^2 = 11{,}070$ and thus, for this hypothesis test, it is concluded that it is rejected H_0.

Testing the significance of logistic regression parameters by testing all independent variables against partially bound variables using the Wald test.

Hypotheses:

$H_0: \beta_j = 0, for\ j = 1,2,\cdots,k$ (xj modifier has no real effect)
$H_1: \beta_j \neq 0$ (xj modifier has real effect).

Table 5. Wald Test Results

Variable		β	P-value
Dependen	[Y=1]	-4.325	1.051
	[Y=2]	2.095	2.144
	[Y=3]	5.585	1.402
	[Y=4]	8.484	1.345
Independen	$X_{1(2)}$	-8.202	2.769364e-05
	$X_{1(3)}$	-3.883	4.630022e-02
	$X_{1(4)}$	-0.853	6.683753e-01
	$X_{2(3)}$	1.663	1.480498e-01
	$X_{2(4)}$	1.255	2.895705e-01
	$X_{3(3)}$	-1.100	6.029765e-02
	$X_{3(4)}$	-0.636	3.176876e-01
	$X_{4(3)}$	-0.853	1.656554e-04
	$X_{4(4)}$	1.663	6.698333e-10
	$X_{5(3)}$	5.424	2.509796e-10
	$X_{5(4)}$	5.444	2.663469e-09

By looking at the significant values in the table, it can be seen that there are several independent variables that have a significant effect on the bound variable, namely direct evidence variables for the disagreement category ($X_{1(2)}$), Direct Evidence Variables for Agree Category ($X_{1(3)}$), Guarantee variables for the Agree category ($X_{4(3)}$), the jeminvariable for the category strongly agrees ($X_{4(4)}$), Empathy variables for the Agree category ($X_{5(3)}$), and the empathy variable for the category strongly agrees ($X_{5(4)}$).

3. Selection of the Best Model of Ordinal Logistic Regression

The selection of the best model is which independent variables included in the model so that the model can explain the relationship with the dependent variable well. The selection of the best model using the backward method. The backward method is a method that includes only all significant independent variables. The procedure is stopped if there are no more significant independent variables. [15].

Table 6. Results of Coefficient Model Significance Testing

Independent Variable	β	Significance
$X_{1(2)}$	-8.202	2.769364e-05
$X_{1(3)}$	-3.883	4.630020e-02
$X_{4(3)}$	3.486	1.656554e-04
$X_{4(4)}$	6.254	6.698333e-10
$X_{5(3)}$	5.424	2.509796e-10
$X_{5(4)}$	5.444	2.663469e-09

Independent variables that have been reduced with a significance value smaller than α (0.05) will be independent variables for the best model. Next, a table can be made containing the results of the regression analysis after reduction.

$$g_1(x) = -4.325 - 8.202x_{1(2)} - 3.883x_{1(3)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(3)}$$

$$g_2(x) = 2.095 - 8.202x_{1(2)} - 3.883x_{1(3)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(3)}$$

$$g_3(x) = 5.585 - 8.202x_{1(2)} - 3.883x_{1(3)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(3)}$$

$$g_4(x) = 8.484 - 8.202x_{1(2)} - 3.883x_{1(3)} + 3.486x_{4(3)} + 6.254x_{4(4)} + 5.424x_{5(3)} + 5.44x_{5(3)}$$

All variables in the equation above have a small significance value of α (0.05) and the significance value of each variable can be seen in Table 6.

4. Model Interpretation

Interpret the model using odds ratio. Odds ratio is a measure that estimates how much the tendency of independent variables to dependent variables is. Odds ratio is a measure to determine the risk of a tendency to experience an event between one category and another category in a variable notated with (OR). To interpret how much chance of free variables in the best model that has been obtained, you can see the odds ratio table below :

Table 7. Nilai odds ratio dari model

Independent Variables (X)	Exp (β)
$X_{1(2)}$	2.740284e-04
$X_{1(3)}$	2.057391e-03
$X_{4(3)}$	32.66686
$X_{4(4)}$	520.5527
$X_{5(3)}$	226.9482
$X_{5(4)}$	231.4855

In table 7 above, we can see that the guarantee variable for the strongly agree category has the highest chance of 520.5527 times the patient feels very satisfied compared to the direct evidence variable and the empathy variable.

5. CONCLUSION

The independent variables that affect the quality of health services are direct evidence, guarantees, and empathy. The opportunity for the quality of health services can be seen from the odds ratio. The odds ratio value of the direct evidence variable for the disagreement category is 2.740284e-04. The odds ratio value of the direct proof variable for the agree category is 2.057391e-03. The value of the odds ratio of the collateral variable for the strongly agree category is 32.66686. The odds ratio value in the guarantee variable for the strongly agree category is 520.5527. The odds ratio value in the empathy variable for the agree category is 226.9482. The odds ratio value in the empathy variable for the strongly agree category is 231.4855.

REFERENCES

[1] Fitrianti, Rachma. 2012. Goverment Support in Triple Helix Collaboration to Provide Health Service Delivery: Case Study Government Hospital in Bengkulu Hospital. Universitas Indonesia.
[2] Arifah, Umi, 2013. Pengaruh Kualitas Pelayanan Terhadap Kepuasan Rawat Jalan di Balai Besar Kesehatan Paru Masyarakat (BBKPM) Surakarta. Program Studi Magister Manajemen. Universitas Muhammadiyah Surakarta.
[3] Brady, Michael K & J. Joseph. 2001. Some Thought on Conceptualizing Perceived Service Quality: A Hierarchical Approach.
[4] Purnami, D. & Gandhiadi, G. K. (2015). Penerapan Regresi Logistik Ordinal Untuk Menganalisis Tingkat Keparahan Korban Kecelakaan Lalu Lintas Kabupaten Buleleng.
[5] Hosmer and Lemeshow. (2000). Applied Logistic Regression. Canada. A Wiley Interscience Publication.
[6] Hosmer, D., & Lemeshow, S. (2013). Applied Logistic Regression Edisi ke-3 New Jersey, Canada: John Willey and Sons
[7] Agresti, A. (1990). Categorical Data Analysis. New York: John Wiley and Sons, Inc.
[8] Agresti, J. W. & S. (2009). An introduction to categorical data analysis. New York: John Wiley and Sons Inc
[9] Kleinbaum, D. G. (1994). Logistic Regression. Springer, New York.
[10] Adam, L & Mark, L. (2018). Ordinal Regression Using SPSS Statistics. UK: Lund Research Ltd.
[11] Susilaningrum, D., & Megahardiyani, C. E. (2009). Analisis Regresi Logistik Ordinal untuk Mengetahui Faktor-Faktor yang Mempengaruhi Status Gizi Balitas Nelayan Kecamatan Bulak Surabaya. PYTHAGORAS.
[12] Umar, Husain. (2011). Metodologi Penelitian untuk Tesis dan Bisnis. Jakarta: PT. Raja Grafindo.
[13] Rolko, K., & Friedrich, H. (2017). Locations of Logistics Service Providers in Germany. The basis for a new freight transport generation model. Transportation Research Procedia.
[14] Pentury, T., Aulele, S. N., & Wattimena, R. (2016). Analisis Regresi Logistik Ordinal. BAREKENG: Jurnal Ilmu Matematika Dan Terapan.
[15] Fathurahman. M. (2010). Pemilihan Model Terbaik Menggunakan Akaike's Information Criteria. Universitas Mulawarman.