The Factors Affecting Eye Patients (Cataract) In Jordan by Using the Logistic Regression Model

Adeeb Ahmed Ali AL Rahamneh[1] & Omar M. Hawamdeh[2]

[1] Associated Professor, Applied Statistics, Faculty of Business, Al-Balqa Applied University, Jordan

[2] Associated Professor, Mathematics, The dean of the Faculty of Business, Jordan

Correspondence: Adeeb Ahmed Ali AL Rahamneh, Associated Professor, Applied Statistics, Faculty of Business, Al-Balqa Applied University, Jordan. E-mail: moheeb982005@yahoo.co.uk, Omar_h@bau.edu.jo

Abstract

This study aims to use the logistic regression model to classify patients as infected and without cataracts. The independent variables were used to represent the gender, the age, the pressure in the right eye, the pressure in the left eye, HbA1C, and the anemia, representative variables for the study of Cataract disease affects the eyes, based on a random sample of (116) patients. The results proved that the used logistic regression model is an efficient and representative for data that shows through (Likelihood Ratio Test) and (Hosmer and Lemeshow test), and the study proved that the value of (R Square Nagelkerke=1) this means that 100% of the change in the occurred changes in the response variable explained through the Logistic regression model.

Keywords: cataract (white water), Jordan, logistic regression

1. Introduction

The human eye is a jewel, and it is the tool that enables us to see things around us. It is a priceless grant from God. Environmental factors, life patterns, behavior and daily practices of the individuals could cause damage to the eye, but the biggest danger on the eye is the inflammation so caution is a must. Eye diseases (inflammations) are trachoma, smallpox, leprosy, optic neuritis, cataract (white water), glaucoma (Blue water) and reproductive system diseases such as gallstones and syphilis. There are changes and complaints associated with eye diseases like pain, swelling, redness, a sudden change in vision as double vision, blurred vision, headache and intense sensitivity to light. The causes of partial vision loss are infections, hereditary diseases, congenital diseases, diseases of malnutrition, infections, tumors, glaucoma, short-sightedness and poor vision. Cataract is the opacity of the transparent lens in the eye, while glaucoma is a disease of the optic nerve due to high pressure in the eye as a result of damage to the optic nerve tissue, which leads to the formation of blind spots within the eye (partial loss of vision field). And if the disease is not treated, the optic nerve is completely damaged and the eye loses its ability to see. Glaucoma is commonly known as blue water, but in fact it is a common mistake as there is no blue water inside the eye, but this naming came from the Greek word "Glaucoma " which means blue waterfalls, because the patient sometimes sees blue halos around the light source giving the impression that there is blue water inside the eye. Glaucoma is the leading cause of blindness in the elderly which can be prevented if treatment is initiated early.

Many medical studies point out that there is a close relationship between the diagnostic aspects of the disease and some statistical studies and applications used in this field. And this helps in the speed of early diagnosis of difficult disease. However, we find that almost statistical data and the related applications at the Arabian region level in general and the local level in particular, which dealt with the use of the statistical application of the regression model in the medical field, focused on examining the factors influencing the incidence of the various diseases (the risk factors and the causes of the different diseases) and did not address the diagnostic side of the disease (and this is the search problem). Therefore, this study shows the relationship between the factors that assist in the diagnosis of cataract cases (white water), disease-specific disease conditions, and prioritization of factors using the statistical application of the regression model.

This Research aims to use the logistic regression model to classify patients as infected or without cataracts.

2. Literature Review

Some of the previous statistical reports dealing with the subject of the statistical application of the regression model in the medical field, in particular, are presented below:

Hassan and Ahmed (2014) employed the regression model to determine the factors which affect the incidence of anemia in children (2009 - 2013). Their study aimed to analyze several factors that affect anemia in children. Furthermore, Al-Sharout, Mohammed, and Mohsen, Amira Jabber (2013) used Regression analysis to study the Breast Cancer Risk, which focused on the use of Regression as a useful tool in studying the factors affecting breast cancer.

Moreover, Yusuff, N. Mohamad, U. K. Ngah & A. S. Yahaya (2011) had written a paper entitled Breast Cancer Analysis using logistic regression. This study intended to use regression as a finite tool by using Mammogram for breast image to diagnose breast cancer. Qasim and Abd al-Razzaq (2011) studied the effect of a number of variables on the periodontal disease by using the regression model. These variables are such as sedimentation, malnutrition and other factors. They found that the regression analysis is the appropriate statistical tool to be employed.

Mustafa and Khader (2011) studied the factors leading to traffic accidents and the casualties resulted from these accidents in Al-Jazerah province, The aim of this study was to identify the factors leading to traffic accidents and the casualties resulted from them in Al-Jazerah Province. The data were collected from all the victims of these accidents, including the cause and description of the accidents, vehicle type, the road and place of accident, in addition to other factors related to the accident circumstances for all accidents that took place in the province between January 2005 and December 2006. The source of the data is the traffic court. SPSS program was used to enter the data and analyze them. The Logistic regression model was the only method used for this purpose. The study showed that the logistic regression model is capable of predicting the casualty's factor. It was found that the factors related to the driver are the most influential in the increase of the casualties related to road accidents.

Ali and Mohammed (2010) aimed at building Logistic regression model for injuries of the Sudanese female breast cancer with reference to the state of Khartoum, and then used the method of logistic regression to study the status of breast cancer in Sudan.

Abdul Majed and Hasan (2009) used the multiple logistic models to determine the factors affecting glaucoma. The study aimed at determining the risk factors for glaucoma, where the multiple regression models were used. Five variables were used as potential risk factors.

Yassin ,Amal Hassan (2008) used regression load to determine the causes of breast cancer in females, studying the status of the National Center for Nuclear Medicine in Khartoum. This study aimed to identify the factors affecting breast cancer in 200 women from the patients of the National Center for Nuclear Medicine in Khartoum.

Al-Bayati, and Ibrahim (2005) Analyzed the path in the regression model with a particular application. This study addressed the direct and indirect effects of a group of factors influencing the incidence of anemia for two groups of populations most vulnerable to the disease in people who are under 18 years of age and pregnant women.

3. Method

The research is based on the following methodologies:

1) Explaining the statistical file of the regression model and focusing on the characteristics and how to estimate the parameters.

2) Adopting the multiple regression models and observing the results, using the SPSS, Minitab, and Software.

3.1 Variables Used In the Study

1.   Sex

2.   Age

3.   Intra-Ocular Pressure, physiological intraocular pressure is between 12 and 19 mm Hg, any increase in internal pressure of the eye exceeding 19 mm Hg is considered the beginning of the disease that should be treated and re-examined periodically every month for 3 or 6 months depending on the condition of the disease. If the pressure of the eye drops below normal range, it may lead to atrophy in Eye, which is very rare.

4.   Sugar: This variable is very important, and has a significant impact on the incidence of eye disease. The

Normal range of blood sugar in human is between 4.5 -6.5 mm mol per liter. When blood sugar drops below its normal levels, it causes fainting .While when it increases it leads to bleeding inside the eye and consequently causes blindness. Deterioration of the cataract is much faster in diabetics than in healthy people.

5. Packed cell volume or hemoglobin ratio in the blood: The normal ratio is between 40 – 45%. The patient suspected to general anesthesia should have normal blood ratio.

3.2 Statistical Tools Used in the Study

3.2.1 Logistic Regression

The statistical regression is defined as a statistical method for loading the (dataset). A variable is the response variable through one or more independent variables (One or more independent variable) where the response variable (dependent variable) is of the binary type[Logistic regression – MedCalc] (binary or dichotomous). Categorical data is also known as the statistical modeling method for class data is widely used in medical research, finance, business, marketing, economics and insurance (Health ... etc.) (El-Habi , Abdalla & El-Jazzar , Majed , 2014).

Logistic Regression is a linear regression in terms of the relationship between the (dependent Variable) and a set of extant variables (explanatory or explanatory variables) but differs because the dependent variable in the regression model must be of the binary type and this difference is reflected.

Additionally, the logistic regression is not required to be independent variables of the continuous type, and not normally distributed. And the relationship between the dependent variable and the independent variables is not Linearity. This makes the logistic regression model more flexible than the rest Prediction and classification models (Abbas, Khodair, 2012).

Moreover, the categories must be specific and sealed so that each item belongs to only one category. Finally, the sample size used in the regression should be larger than the size of the sample used in Linear regression because the coefficients of the regression model are estimated by using the maximum function method (maximum likelihood method). The way of working out the sample size needed is relatively large.

3.2.1.1 Estimation of Logistic Regression Parameters

The regression model is based on a fundamental assumption that the dependent variable (Y)response variable which is marked by a binary variable following a distribution (Bernoulli)rank (1)with probability (p) &rank (0)with probability(q = 1- p)the occurrence of the response and non-occurrence (Ghanem & Fareed, 2011).

The regression model can be expressed in the case of only one independent variable containing the equation (Hosmer *et, al.,* 2013).

$$f(x) = E(y / x) = \frac{e^{\beta_o + \beta_1 x}}{1 + e^{\beta_o + \beta_1 x}} \tag{1}$$

Such that:

$f(x) = E(y / x)$ represents the conditional average (conditional mean) for response variable (Y) in determined value for (x).

ß$_0$, ß1: logistic regression parameters.

If the regression model contains more than one independent variable, it can be: k, from the dependent variables then they express the following mathematical formula:

$$f(x) = E(y / x) = \frac{e^{\beta_o + \beta_1 x + \beta_2 x_2 + \ldots \beta_k x_k}}{1 + e^{\beta_o + \beta_1 x + \beta_2 x_2 + \ldots \beta_k x_k}} \tag{2}$$

Such that:

Logistic regression parameters [Hosmer , David W. & Lemeshow , Stanley , (2013)] $\beta_o, \beta_1, \beta_2, \ldots \beta_k$:

$$T(x) = Ln(\frac{f(x)}{1-f(x)}) == \beta_0 + \beta_1 X_1 + \beta_2 X_2. + ...\beta_k X_k \tag{3}$$

Such that:

When $y = 1...then...p(y = 1/x) = f(x)$

When $y = 0...then...p(y = 0/x) = 1 - f(x)$

To estimate the parameters ($\beta_o, \beta_1, \beta_2, ...\beta_k$), by using the Maximum Likelihood Method as (Hosmer, David W.

& Lemeshow, Stanley, 2013)

When we have sample size (n)

When we have sample of (n) independent observations, each pair (xi,yi), i=1, 2.,,,,, , n , such that (Hosmer *et, al.,* 2013).

yi: represent the rank of a binary variable for observation (i)

xi: the value of independent variable for observation (i)

Now when y=1, then ...$p(yi = 1/x) = f(x)^{yi}$

When y=o, then ...$p(yi = o/x) = [1 - f(x)]^{1-yi}$

The maximum likelihood function can be expressed as the following mathematical formula (Hosmer *et, al.,* 2013).

$$L(\beta) = L \prod_{k=1}^{n} f(xi)^{yi} [1 - f(xi)^{1-yi} \tag{4}$$

By taking the natural logarithm of both sides of the equation (4) we get (Hosmer *et, al.,* 2013).

$$L(\beta) = Ln[Ln(\beta)] = \sum_{i=1}^{n} [yiLn(f(xi)] + (1 - yi)Ln[1 - f(xi) \tag{5}$$

Now we derive the equation no (5). And then equated with zero, get a number of equations can be solved only through the iterative algorithm (Qasim &Abdel Razaq, 2011). The so-called (Iteratively Weighted Least Squares Algorithm) (Marlene , 2004).

3.2.1.2 The Explanatory Force of the Model Evaluation

In this model in instead of the coefficient of determination by ($R^2 cox \& snell$) and ($R^2 Nagel \& Lklerake$)

Such as (Burns , Robert & Burns , Richard , 2008**)**

$0 \le R^2 cox \& snell) < 1$

$0 \le R^2 Nagel \& Lklerake \le 1$

Can be calculated ($R^2 cox \& snell$) **) as: [**Qasim, Bahaa Abdel Razaq, (2011) **]**

$$R^2 cox \& snell = 1 - [\frac{L_0}{L_1}]^{(2/n)}. \tag{6}$$

L_0: Maximum likelihood function in the case of the inclusion in the Model constant

L_1: Maximum likelihood function in the case of the inclusion in the Model all independent variables

And Can be calculated ($R^2 Nagel \& Lklerake$)as: Qasim &Abdel Razaq, 2011).

$$R^2 Nagel \& Lklerake = \frac{R^2 Cox \& Snell}{1 - [L_o]^{(\frac{2}{n})}}. \tag{7}$$

L_0: Maximum likelihood function in the case of the inclusion in the Model constant

n: Sample Size

3.2.1.3 Logistic Regression Tests

3.2.1.3.1 Wald Test

The test hypothesis:

$H_o : \beta_j = o$ Vs. $H_1 : \beta_j \neq o$ for all j=1,2,3, ..., k (Ghanem & Fareed, 2011).

$$Wald = [\frac{\tilde{\beta}_j}{S.E(\tilde{\beta}_j)}]^2 \tag{8}$$

Distributed with : chi- square with degrees of freedom (1) for all parameter

$\tilde{\beta}_j$: The estimated parameter of the rank (j)

SE ($\tilde{\beta}_j$) The Standard Error for the estimated parameter of the rank (j)

3.2.1.3.2 Homser & Lemeshow Test

This test is used to determine whether well-fitting model or not, through the difference between observed values and expected values and it model is distributed chi- square distribution. (Qasim & Abdel Razaq, 2011), and used to test the following hypothesis (Wuensch , 2014).

H_o :   There is no significant difference between observed values and expected values.

H_1 :   There is a significant difference between observed values and expected values.

Then if (Homser & Lemeshow) static is > (0.05, then the model representatives will be a good model (Wuensch , 2014).

3.3 The Application Side

Table (1) shows the number of the cases inserted in the model, which is 116 cases. While the number of the missing cases was zero.

Table 1. Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	116	100.0
	Missing Cases	0	.0
	Total	116	100.0
Unselected Cases		0	.0
Total		116	100.0

Table (2) shows the values of the variable binary response, as the code zero stands for the injured people, while the code number one stands for the uninjured people.

Table 2. Dependent Variable Encoding

Original Value	Internal Value
Infected (injured people	0
Not infected	1

Table (3) shows the number of iterations for achieving the least value (-2 Log likelihood). This would be achieved in the iteration (20) with value of (0.0). Notice that the initial value in the indicated model equals to (143.695), Estimates were ($\tilde{\beta}_1, \tilde{\beta}_2, \tilde{\beta}_3, \tilde{\beta}_4, \tilde{\beta}_5, \tilde{\beta}_6$) achieved in the iteration (20) are the best estimated in this model.

Table 3. Iteration History[a,b,c,d]

Iteration		-2 Log likelihood	Constant	Coefficients					
				X1	X2	X3	X4	X5	X6
Step 1	1	32.782	14.247	-.095-	-.216-	-.069-	-.108-	-.066-	.011
	2	11.649	23.246	-.186-	-.345-	-.096-	-.205-	-.133-	.016
	3	4.365	32.107	-.280-	-.468-	-.118-	-.322-	-.204-	.020
	4	1.661	41.224	-.376-	-.591-	-.137-	-.458-	-.279-	.022
	5	.634	50.719	-.472-	-.715-	-.155-	-.612-	-.359-	.024
	6	.242	60.640	-.568-	-.842-	-.175-	-.784-	-.445-	.024
	7	.092	70.997	-.663-	-.971-	-.197-	-.973-	-.539-	.024
	8	.035	81.780	-.756-	-1.103-	-.221-	-1.176-	-.642-	.023
	9	.013	92.962	-.849-	-1.237-	-.249-	-1.393-	-.753-	.021
	10	.005	104.504	-.941-	-1.373-	-.281-	-1.620-	-.873-	.018
	11	.002	116.364	-1.032-	-1.510-	-.318-	-1.855-	-.999-	.015
	12	.001	128.498	-1.125-	-1.649-	-.360-	-2.097-	-1.132-	.011
	13	.000	140.866	-1.219-	-1.790-	-.407-	-2.345-	-1.269-	.006
	14	.000	153.434	-1.314-	-1.931-	-.459-	-2.595-	-1.409-	.002
	15	.000	166.171	-1.412-	-2.073-	-.516-	-2.849-	-1.553-	-.003-
	16	.000	179.054	-1.513-	-2.216-	-.579-	-3.104-	-1.698-	-.008-
	17	.000	192.063	-1.616-	-2.359-	-.646-	-3.361-	-1.844-	-.014-
	18	.000	205.185	-1.722-	-2.502-	-.718-	-3.620-	-1.990-	-.019-
	19	.000	218.406	-1.831-	-2.646-	-.795-	-3.879-	-2.137-	-.025-
	20	.000	231.717	-1.942-	-2.790-	-.877-	-4.140-	-2.284-	-.031-

a. Method: Enter

b. Constant is included in the model.

c. Initial -2 Log Likelihood: 143.695

d. Estimation terminated at iteration number 20 because maximum iterations have been reached. Final solution cannot be found.

Table (4) shows the quality and efficiency of the model by using the likelihood ratio test ,which follows the(Chi-Square) distribution with degree of freedom equals to the number of independent variables(6). And this means:

$$\chi^2 143.695 - 0.000 = 143.695$$

Table 4. Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	143.695	6	.000
	Block	143.695	6	.000
	Model	143.695	6	.000

The results of the model summary in regression analysis is illustrated in table 5, as noted the (p- value) equals to (0.00) , which indicates the significance and the quality of the used model.

Table 5. Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	.000[a]	.710	1.000

a. Estimation terminated at iteration number 20 because maximum iterations have been reached. Final solution cannot be found.

It can be observed in table (5) that the value of (R^2Cox & Snell=0.71), which means 71% of the occurred

changes in the response variable were explained through the Logistic regression model. While the value (R^2 Nagelkerke =1) means 100% of the change in the variable of the binary response were explained through Logistic regression model.

It is clear from the results in table (6), results of (Hosmer and Lemeshow) test which follow (Chi-square) distribution with value (0.0) which is not statistically significant. Where the probable value was (1). So there is no difference between the observed value end the expected value. The results in table (7) also assure that.

Table 6. Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	.000	8	1.000

Table 7. Contingency Table for Hosmer and Lemeshow Test

		The patient is infected		The patient is not infected		
		Observed	Expected	Observed	Expected	Total
Step 1	1	**12**	12.000	0	.000	**12**
	2	**12**	12.000	0	.000	**12**
	3	**12**	12.000	0	.000	**12**
	4	**12**	12.000	0	.000	**12**
	5	**12**	12.000	0	.000	**12**
	6	**12**	12.000	0	.000	**12**
	7	**8**	8.000	4	4.000	**12**
	8	**0**	.000	12	12.000	**12**
	9	**0**	.000	13	13.000	**13**
	10	**0**	.000	7	7.000	**7**

Table (8) shows the percentage of the right classification of the injured people with Cataract disease were classified correctly 100%, Also the percentage of the uninjured people with Cataract disease was classified by 100%. While the percentage of people classified incorrectly was 0.0%. That gives us an excellent indicator to the ability of the model to classify correctly.

Table 8. Classification Table[a]

Observed		Predicted		
		type of disease		
		infected	Not infected	Percentage Correct
Step 1	infected	80	0	100.0
	Not infected	0	36	100.0
	Overall Percentage			100.0

a. The cut value is .500

Table (9) shows the estimated values for ($\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$), and the Logistic regression model as:

$$\hat{Y} = 231.717 - 1.942\ X_1 - 2.79\ X_2 - 0.877\ X_3 - 4.14\ X_4 - 2.284\ X_5 - 0.031\ X_6 + \zeta i$$

As we noted there is a statistical value represented by (WALD) for each estimated of the capabilities. While column number (6) represents the percentage of recoil for injury once there is a change in the value of the independent variable linked with ($Exp(\beta)$). Column number (7) explains the confidence limits for the estimation.

If the value of ($Exp(\beta)$) was larger than 1, the percentage of recoil for injury increased, while if the value of ($Exp(\beta)$) is less than 1, this means there is an increase in the value of independent variable which leads at the end to a decrease in the percentage of recoil for injury.

According to the variable (x_1) which represents the Gender, value equals to (0.143) indicates the probability of the person to be injured by (0.143) in males is higher than in females.

According to the variable (x_2) which represents the Age, value equals to (0.061). which means that every increase in the age of patient by one year, the percentage of injury increases by (0.061).

According to the variable(x_3) which represents the pressure in the right eye, value equals to (0.4161). This indicates every increase in the pressure of the right eye for a patient by one unit, the chance of injury increased by (0.4161).

According to the variable(x_4) which represents the pressure in the left eye, value equals to (0.016). This indicates that with every increase in the pressure of the left eye for the patient by one unit, the chance of injury increased by (0.016).

According to the variable(x_5) which represents HB-A1C (sugar cumulative) Value equals to (0.102), which indicates that with every increase in the pressure of the left eye of the patient by one unit, the chance of injury increased by (0.102).

According to the variable(x_6) which represents Anemia, value equals to (0.970). This indicates that with every increase in the Anemia by one unit, the chance of injury increased by (0.970).

Table 9. shows the Variables in the Equation

| | | B | S.E. | Wald | df | Sig. | Exp(B) | 95% C.I.for EXP(B) | |
								Lower	Upper
Step 1[a]	X1	-1.942-	8528.201	.000	1	1.000	.143	.000	.
	X2	-2.790-	841.683	.000	1	.997	.061	.000	.
	X3	-.877-	8263.650	.000	1	1.000	.416	.000	.
	X4	-4.140-	6092.227	.000	1	.999	.016	.000	.
	X5	-2.284-	7041.774	.000	1	1.000	.102	.000	.
	X6	-.031-	996.373	.000	1	1.000	.970	.000	.
	Constant	231.717	143354.956	.000	1	.999	4.300E+100		

a. Variable(s) entered on step 1: X1, X2, X3, X4, X5, X6.

4. Results

1-The study shows that there is a relation between the independent variables and the response variable which represented the value of (NagelKerke R Square) that equals (1).

2-The study proved that the used logistic regression model is efficient and representative of the data shown through (Likelihood Ratio Test) and (Hosmer and Lemeshow test).

3-The study explained that the percentage of the true classification of the people who had Cataract disease was classified truly by 100%. Also, the percentage of the people who didn't have Cataract was classified truly by 100%. While the percentage of the people who were wrongly classified was (0.0). And this gives an excellent indicator of the capability of the model to classify the patients correctly.

4-The study proved that the value of (R Square Nagelkerke=1). This means that 100% of the change in the occurred changes in the response variable were explained through the Logistic regression model.

5-The study explained that the independent variables represented by the gender, the age, the pressure in the right eye, the pressure in the left eye, HbA1C, and the anemia are representative variables for the study of Cataract disease which affects the eyes.

5. Recommendations

The researcher recommends the following:

1. Adopting the logistic regression method as an efficient statistical method for the classification of patients in the medical field.

2. Conducting extensive studies on ophthalmology and expanding the sample size further to cover more and more aspects of the human eye diseases.

References

Abdul, M., & Hassan, A. I. (2009). The Use of Multi-Logistical Template to Determine Factors Affecting Disease Master, Faculty of Science, Sudan University of Science and Technology.

Al, B., & Ibrahim, H. (2005). Analysis of the Path in the Logistic Regression Model, With a Practical Application, Faculty of Administration and Economics, al-Mustansiriyah University

Ali, A., & Mohammad, A. (2010). The Logistic Regression Model to Breast Cancer Among Sudanese Women, unpublished Phd Thesis, Faculty of Science, Sudan University of Science and Technology

El-Habi, A., & El-Jazzar (2014). A Comparative Study Between Linear Discriminant Analysis and Multinomial Logistic Regression. *An-Najah University Journal Research (Humanities)*, *28*(6), 1526-1548.

Ghanem, A., & Fareed, A. (2011). The Use of Binary Logistic Regression in Response to the Study of Social and Economic Determinants of the Adequacy of Family Income (A Practical Study on a Random Sample of Households in the Province of Damascus, *Damascus University Journal of Economic and Legal Sciences*, *27*(1), 113 – 132.

Hassan, & Ahmad, F. (2014). The Use of the Logistic Model to Determine the Factors Affecting the Incidence of Anemia (Anemia) In Children. Master Thesis, Faculty of Science, Sudan University of Science and Technology.

Hosmer, D. W., & Lemeshow, S. (2013). *Applied Logistic Regression* (3rd ed.), NewYork: Wiley Retrieved from http://ihmsi.org/

Khodair, A. (2012). The Use of Logistic Regression in the Predictability of the Computable Functions of Economic Variables Belong Quality. *Kirkuk University Journal for Economic and Administrative Sciences*, *2*(2), 234-253.

Marlene, M. (2004). *Generalized Linear Models. Fraunhofer Institute for Industrial Mathematics (ITWM), (Germany).* Retrieved from http://www.marlenemueller.ed/publications/HandbookCS.pdf

Mustafa, & Khader, O. (2011). The Factors Leading to Traffic Accidents, Injuries Resulting from Accidents In the Island. Phd Thesis, Faculty of Science, Sudan University of Science and Technology.

Qasim & Abdel Razaq, B. (2011). The Analysis of Some of the Impact of Changes in the Incidence of Gum, Using the Logistic Regression Model. *Journal of Economic Sciences, 7*(27), 137-164.

Robert, B., & Richard, B. (2008). Business Research Methods and Statistics using SPSS. Five Extra Advanced Chapters. pp 568 -575. Retrieved from http://www.uk.sagepub.com/burns/website

Wuensch, Karl L. (2014). *Binary Logistic Regression with SPSS.* Retrieved from http://www.care.ecu.edu/psyc/wunesch/MV/Multerg/logistic_SPSS

Yassin, & Mohamed, H. (2008). The Use of the Logistic Regression Model to Determine the Causes and Incidence of Breast Cancer in Females Case Study of the National Center for Treatment, The Prophet's Medicine in Khartoum Master, Faculty of Science, Sudan University of Science & Technology. Retrieved from http://www.sustech.ed/sudannewar/thesis.php?collno=2&serial=6059

Yusuff, N. M., Ngah, U. K., & Yahaya, A. S. (2012). Breast Cancer Analysis Using Logistic Regression. *IJRRAS*, *10*(1), 14-22. Retrieved from http://www.arpapress.com/Volumes/Vol10Issue1/IJRRAS_10_1_02.pdf

Copyrights