

Using Ordinal Logistic Regression Analysis in Evaluating Teachers' Performance Level of High Schools (12th grades) in Kurdistan Regional Government

Luqman Sleman Omar

Statistics & Informatics Department, College of Administration and Economic, Salahaddin University –Erbil

luqmansleman@gmail.com

Delshad Shaker Ismael Botani

Statistics & Informatics Department, College of Administration and Economic, Salahaddin University –Erbil

delshd.botani@su.edu.krd

Abstract:

The Ministry of Education – Kurdistan Regional Government (MOE-KRG) currently depends only on the students' grades in the evaluation of the teachers' performance level in all subjects as low, medium, or high performance. Relying just on one variable to determine teacher's performance is not fair and this problem must be resolved statistically through finding new proposed statistical model. Therefore, this study tries to find some variables that are available in MOE-KRG for all high schools in order to use them in the proposed model to predict teachers' performance level instead of the old one. This study aims to predict teacher's performance level of high schools (12th grades) in KRG who teach Kurdish subject, and also analyze the effects of variables that have impact on the evaluation of teachers' performance level depending on the data that are available in the MOE-KRG. In this paper Ordinal Logistic Regression (OLR) method is used to find a proposed model for evaluating teachers' performance in Kurdish subject using the data of all scientific high schools in KRG (646 high schools). The teachers' status was analyzed by selecting nine variables related to the high schools: sector (governmental or non-governmental), geographic location, type of education, status of school (exemplary, non-exemplary), gender of student, year of school establishment, number of classes, number of teachers and student's average marks in Kurdish subject. It is concluded that four of these variables (student's average in Kurdish subject, number of classes, geographical location and status of school) have significant effects on teachers' evaluation in 12th class of scientific high schools, and the overall percentage of correct classification

is about 87%, it that means the Ordinal logistic regression model has an ability to predict teacher's performance level very well.

Key Words: Ordinal logistic regression. Teachers performance evaluating. Classification

1. Introduction

Kurdish subject is one of the most important curricula taught students in the twelfth grade of Scientific branch because of their importance in developing their mental abilities to analyze, develop and understand information not only in Kurdish subject but also in the other scientific subjects taught by students in the course. This is a problem that needs to be addressed by identifying the reasons for the lack of good evaluation in this article and the elimination of obstacles in order to improve the teaching level of the teachers. Several studies have been conducted in the social and medical fields, but few of them studies are about the reasons of how evaluate teachers, so we conduct this research to find out the most important problem facing the teacher in the twelfth scientific stage of the low level of rating in Kurdish subject to find some suggested to this problem, helping to raise the scientific level of the teachers.

1.2 Literature Review:

The practice of statistical approaches for ordered data can be traced back to the late 1950 (Ashford, 1959). The first efforts to evaluate asymmetrical problems, where an ordered categorical variable was measured as the dependent variable using multiple regression procedures were available in the later 1960 (Walker and Duncan, 1967). However, reviews on this matter did not appear pending the early 1980 (McCullagh, 1980, Anderson and Mohan, 1984). In a series of papers, (Goodman and Kruskal, 1979) developed log-linear models that were proposed for proportioned problems where the association between several variables, some of which were ordinal were studied. Since then regression models that fit ordinal data have been included in the larger category of Generalized Linear Models (McCullagh and Nelder, 1989) Generally, there are two main classes of models that analyze ordinal categorical data – Loglinear and Logit (binary and ordinal regression) models (Agresti, 1989) Log-linear models, which allow for the ordering of the categories, for one or more variables, are called association models and they define association patterns among the variables. With this latter method, the cell counts are modelled in a likelihood table in terms of associations among the variables, and the difference between the response and the explanatory variables is not made. In the situation of the logit models, one variable is clearly treated as the response and the others are the covariates. In medical research studies, it is more usual to observe the relationship of the outcome with respect to other measures, as opposed to analyze the relationships among several variables. As a consequence, the logit models are more fitting for our purpose. All the ordinal regression models decrease to the binary logistic model when the categories are collapsed into two or when only two categories are measured. One advantage of an ordered analysis over the equivalent nominal

analysis is that generally rarer parameters are required to describe a model for the response (Greenwood and Farewell, 1988). As a result, the ordinal regression models are more powerful. Also, modelling ordered categorical data is essentially more difficult than modelling continuous data due to the restrictions on the basic probabilities and the reduced amount of information that discrete outcomes contain. Adejumo, A.O & Adetunji, A.A (2013) analyzed the impact of some variables on academic performance of Nigeria students. Mode of entry, age at entry, department, and sex of students are examined as factors that could contribute to students' performance. Ordinal Logistic Regression is used to model the data and the results reveal that only sex of students is not a determinant factor of final grade that students may attained at graduation. It has also been established that younger students perform better than the older ones. (Adejumo and Adetunji, 2013).

1.3 Aim of Research

The aims of this research are the following:

1. Prediction of the teacher's performance levels (low, medium and high performance) who teach Kurdish subject in 12th grades of high school – MOE-KRG.
2. Depending on the availability of data in the MOE-KRG, trying to analyze the effects of variables that have impact on the evaluation of teachers' performance level using the ordinal logistic regression model.

1.4 Importance of research

This research is very important for MOE-KRG through finding a better evaluation model to predict teachers' performance level using ordinal logistic regression method as an important tool instead of using current MOE-KRG model which depends only on the students' grades in Kurdish subject.

1.5 Research Problem

The problem of the research is lack of an appropriate model in MOE-KRG for identifying the most important variables that have great impacts on the teachers' performance who teach Kurdish subject in 12th grade of scientific high schools.

1.6 Research Hypothesis

This Research depends on the choice of the following general hypothesis:

$$H_0: B_0 = B_1 = \dots B_9 = 0,$$

$$H_1: B_0 \neq B_1 \neq \dots B_9 .$$

Using OLR method in this research to determine which explanatory variables has a significant effect on the dependent variable. Significance of regression coefficients are zero in the population from which sample was

reserved. $H_0: B_k = 0$, that means none of independent variables have a significant effect on the dependent variable. $H_1: B_0 \neq B_1 \neq \dots B_9$, that means independent variables have a significant effect on the dependent variable (Hosmer Jr et al., 2013).

1.7 Statistical Research Tools

Given that the multiple explanatory variables and the dependent variable is a multi-dependent variable, the OLR method is used and through applying SPSS (version 25) package the researcher will get the desire results.

1.8 Limits of Research

Human limits: Teachers of 646 schools of the 12th grade of scientific branch.

Spatial limits: Most schools in the educational directorates of KRG.

Time limits: The academic year 2017.

1.9 Teachers evaluation

The public views teachers' evaluation is a major problem in the school's system (Soar et al., 1983). Where responsible bodies aware of the concern on this issue need to present and mandate more effective evaluation. Common methods for evaluating teachers, such as measurement tests of teacher characteristics, student achievement test scores, and ratings of teachers' classroom performance, have been ineffective. Some research has been done to improve the evaluation process, but teacher evaluation, in general, remains unorganized.

MOE officially depends on the student's success rate in the school compared to the general average in the governorate to which the school belongs. It means that, they depend on the success rate of the students and the grades obtained by the students only, (Vernez et al., 2016) and for this, they use the standard scores for all subjects studied in the schools. Here, a serious question arises:

Is it enough to take students' grades and success rates to evaluate teachers' performance in KRG's schools? What about the other factors and variables that should be taken into considering for evaluating teacher's performance? These issues can be solved using statistical methods by insertion of some new factors and variables.

2. Methods and Materials

2.1 Methodology

Since this study trying to understand and predict the teachers' performance level in 12th grades of high schools, and data had been collected from historical sources of MOE-KRG, the applied research (Quantitative Research)

is being taken into consideration. The OLR method has been used to analyze and predict the teachers' performance level because it is one of the issues for teachers and MOE-KRG.

2.2 Ordinal Logistic Regression (OLR)

Logistic regression may be advantageous when a model is suggested for categorical dependent variable as a function of one or more independent variables existence the dependent variable consumes two outcomes. OLR is a kind of logistic regression analysis when the dependent variable has more than two categorizes with having ordinary order or rank. In statistics, the ordered logit model is a regression model for ordinal dependent variable. It is ordinary to deliberate methods for more categorical dependents having more than two probable values. The most famous of these ordinal logistic regression methods is termed the proportional odds model. The basic idea of the proportional odds model is re-expressing the categorical variable in terms of a number of tertiary variables based on interior cut points in the ordinal scale. We can reflect the three tertiary logistic models equivalent to regressing each of the Separately against the X s. Many variables of attention are ordinal. That is, we can rank the values, but the actual distance between categories is indefinite (Hilbe, 2009).

$$\text{logit}[p(y \geq 3)] = \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1)$$

2.3 Ordinal Regression Assumptions

There are some assumptions that are given below:

- 1- The dependent variable must be measured at the ordinal level.
- 2- Ordinal independent variables should be also continuous or categorical.
- 3- Multicollinearity should not be existing.
- 4- Each independent variable has an equal effect at each collective split of the ordinal dependent variable.
- 5- Each pair of outcome groups is the same in relationship.
- 6- The effects of any descriptive variables are reliable or proportional across the different thresholds.

(Hosmer Jr et al., 2013).

2.4 The Ordinal Logistic Regression Model

The stages in model building for an ordinal logistic regression model are basically the equivalent as described in the binary logistic regression model for the multinomial logistic regression model (Kleinbaum and Klein, 2010):

$$\ln(\theta_j) = \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_9 X_9 \quad (2)$$

where, j goes from 1 to the number of categories minus 1.

It is the part we certainly need to discover. We have been thinking of assumption dependents to miscarriage as a set of three unordered dependents but there is a very clear and planned ordering to these dependents like as

low performance, medium performance, and high performance. If we identify that a category is ordinal then there are superior models that express for us how independent variables relation to somebody being higher or lower on the scale. We might practice the ordered logit model so that we might practice the categorized straight as our dependent variable. Different relatives lead to proportional odds models or ordered probity models. The model cannot be dependably estimated using ordinary least squares, it is regularly estimated using maximum likelihood. The partial proportional odds model and proportional odds model are superior cases of the cumulative logit model. This is a cumulative logit model that accepts that the odds of dependent under a given dependent level are constant irrespective of which level we selected. This model allows separated intercepts for the cumulative logit but controlled the parameter sets for the predictors to be the similar across all logits. A proportional odds model that compels some predictors to have communal parameters and leaves other predictors free to have separate parameters is named a partial proportional odds model. The basic form of the generalized linear model is:

$$\text{link}(\gamma_j) = \frac{\theta_j - [\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k]}{\exp(\tau_1 Z_1 + \tau_2 Z_2 + \dots + \tau_m Z_m)} \quad (3)$$

where,

γ_j the cumulative probability for the j^{th} category.

θ_j the threshold for the j^{th} category.

β_1, \dots, β_k are the regression coefficients.

X_1, \dots, X_k are the predictor variables.

and k is the number of predictors.

The numerator on the right side specified the location of the model. The numerator of the equation determinate the scale. The τ_1, \dots, τ_m are coefficients for the scale component and Z_1, \dots, Z_m are m predictor variables for the scale component chosen from the same set of variables as the X 's. (Hosmer Jr et al., 2013)

2.5 Ordinal Logit Model Fitting

If the arrangement between levels of the ordinal dependent scale is identified, therefore numerical scores can practically be allocated to the dependent levels, then a mean dependent model can be fit. When we have more than two proceedings, we can extend the ordinal logistic regression model. Assume the original association to be categorized is: (Agresti, 2007)

$$\gamma_i = X' \beta + \varepsilon \quad (4)$$

where,

γ_i is the particular but undetected dependent variable

X is the vector of independent variables.

and β is the vector of regression coefficients which we request to estimate.

The Model fitting Information gives the $-2 \log$ likelihood ($-2LL$) values for the starting point and the last model and a chi-square to test the alteration between the $-2LL$ for the two models. The goodness of fit statistics specified that the model fits much better than the location only model. From the detected and predictable frequencies, we can compute the typical Pearson and Deviance goodness of fit measures.

The Pearson goodness of fit statistic is (Hosmer Jr et al., 2013):

$$\chi^2 = \sum \sum \left(\frac{O_{ij} - E_{ij}}{E_{ij}} \right)^2 \quad (5)$$

The deviance measure is:

$$D = 2 \sum \sum O_{ij} \ln \left(\frac{O_{ij}}{E_{ij}} \right) \quad (6)$$

Both of the goodness-of-fit statistics would be used only for models that have practically large predictable values in each cell. If we have a continuous independent variable or many categorical interpreters or some predictors with many values, we may have many cells with small expected values. There are several R^2 like statistics that can be used to measure the strong point of the association between the dependent variable and the interpreter variables. They are not as advantageous as the R^2 statistic in regression, subsequently their interpretation is not straight forward. Three commonly used statistics are (Kleinbaum and Klein, 2010):

$$1- \text{Cox and Snell } R^2 = 1 - \left(\frac{L(B^{(0)})}{L(\hat{B})} \right)^{\frac{2}{n}}$$

$$2- \text{Nagelkerke's } R^2 = \frac{\text{cox and snell } R^2}{1 - L(B^{(0)})^{\frac{2}{n}}} \quad \text{and}$$

$$3- \text{McFadden's } R^2 = 1 - \left(\frac{L(\hat{B})}{L(B^{(0)})} \right).$$

where, $L(\hat{B})$ is the log-likelihood function for the model with the estimated parameters, and $L(B^{(0)})$ is the loglikelihood with just the thresholds and n is the number of cases.

2.6 The ordinal logistic regression Parameter Estimates

The proportional odds model accepts that the true β values are the same in all three models, the adjacent-category (AC), the continuation-ratio (CR) and the proportional odds (PO) models (Hosmer Jr et al., 2013), The model is regularly estimating via maximum likelihood. The technique of maximum likelihood parallels too many renowned estimation methods in statistics. Therefore, the only modification in models is the intercept terms, α_c , $c = 1, 2, 3$. This means that the estimates from the three ordinal models could be mutual to afford just one set of β estimates. Through exponentiation the mutual estimate relation to a specified predictor, i.e. taking e^{β_j} we got an estimate of the common odds ratio (OR) that defines the relative odds for $y \geq 3$ for values of X_j different by one unit. The maximum likelihood estimate (MLE) of θ is that value of θ that maximized $lik(\theta)$:

it is the value that made the detected data the “maximum probable”. If the X_i are *iid*, then the likelihood abridged to (Harrell Jr, 2015):

$$lik(\theta) = \prod_{i=1}^n f(X_i|\theta) \quad (7)$$

Rather than maximizing this product which might be quite monotonous, we frequently used the statistic that the logarithm is an increasing function so it would be correspondent to maximized the log likelihood (Dayton, 1992):

$$l(\theta) = \sum_{i=1}^n \log(f(X_i|\theta)) \quad (8)$$

2.7 Model Adequacy Checking

2.7.1 The Deviance and Goodness of Fit

A most general method to detected nonexistence of fit searches for any way the model fails. A goodness of fit test compared the model fit with the data. This method regards the data as representing the fit of the most intricate model probable the saturated model, which has a separated parameter for each observation. In confident cases, this test statistic has a large sample chi squared null distribution. The likelihood ratio statistic for this test is the deviance and Pearson of the model. (Archer and Lemeshow, 2006).

3. Ordinal Logistic Regression Analysis

The research was based on data taken from the Ministry of Education – General Directorate of Examinations and contains a number of importance variables that may lead to the evaluation of teachers. The study community is the most scientific schools in the Kurdistan region, which is 646 schools of the scientific branch. In this chapter, data will be presented and analyzed by the ordinal logistic regression using SPSS package version 25. Also, the logistic regression will be used to illustrate the importance of each variable of independent variables and their contribution to teachers' evaluation and the significance of each variable and its difference from the remnant of the variables.

3.1 Codes and Signs used for Research Variables

The dependent variable was evaluation of teachers which was categorized as low performance, medium performance and high performance.

Table 3.1: Nature of the dependent variable with the corresponding code

	Y = 0	Low performance	If standard degree < -1
--	-------	-----------------	-------------------------

Dependent variable of Kurdish subject teacher's evaluation	Y = 1	Medium performance	If $1 \leq \text{standard degree} \leq -1$
	Y = 2	High performance	If standard degree > -1

The independent variables that used in this research are: sector, location, type of education, status of school, gender of student, year of school establishment, number of classes in school, number of teachers in school and average of Kurdish subject in school.

Table 3.2: List of the independent variables with their respective names and category code.

Independent variables	X ₁	Average of Kurdish subject in school
	Sector X ₂	Governmental = 1
		Non-governmental = 2
	Location X ₃	Central = 0
		District = 1
		Sub-district = 2
		Village = 3
	Type of education X ₄	Scientific = 0
		Both = 2 (Scientific and literary branch)
	Status of school X ₅	Non exemplary = 1
		Exemplary = 2
	Gender of Student X ₆	Boys school = 0
		Girl school = 1
Coeducational = 2		
X ₇	Year of school establishment	
X ₈	Number of classes in school	
X ₉	Number of teachers in school	

3.2 Descriptive Statistics

Table 3.3 shows the descriptive statistics of the independent variables with respect to the dependent variable (Kurdish teacher's Assessment) in order to describe the characteristics of these variables.

Table 3.3 descriptive statistics of dependent and independent variables

Independent variables	Kurdish teacher's evaluating (Scientific branch)								
	Low performance			Medium performance			High performance		
	N	%	Mean	N	%	Mean	N	%	Mean

Sector	Governmental	84	92.3		486	98.2		51	85	
	Non-Governmental	7	7.7		9	1.8		9	15.0	
Location	Central	30	33		147	29.7		17	28.3	
	District	21	23.1		107	21.6		10	16.7	
	Sub-District	20	22		127	25.7		16	26.7	
	Village	20	22		114	23		17	28.3	
Type of education	Scientific	14	15.4		87	17.6		24	40	
	Both (scientific and literary)	77	84.6		408	82.4		36	60	
Status of school	Non-Exemplary	84	92.3		477	96.4		39	65.0	
	Exemplary	7	7.7		18	3.6		21	35.0	
Gender of student	For Boys	31	34.1		114	23.0		6	10.0	
	For Girls	17	18.7		157	31.7		10	16.7	
	Coeducational	43	47.3		224	45.3		44	73.3	
	Total	91	100		495	100		60	100	
Year of establishment (Binned)	<=1970	3	3.3		28	5.7		3	5.0	
	1971-1980	5	5.5		32	6.5		3	5.0	
	1981-1990	3	3.3		43	8.7		4	6.7	
	1991-2000	10	11		75	15.2		11	18.3	
	2001-2010	40	44		194	39.2		20	33.3	
	2011+	30	33		123	24.8		19	31.7	
	Total	91	100		495	100		60	100	
Number of classes		91		9	495		13	60		12
Number of teachers		91		20	495		28	60		25
Average of Kurdish subject in school (Scientific branch)		91		57	495		71.6	60		83.8

The first variable is sector of schools, the count and percentage of schools those are in the governmental sector which are in the medium performance (486) schools and their percentage is (98.2%) are greater than other levels. The second variable is location which consist of 4 dummy variables, most schools are located in the central location (center of governorate). The minimum count of schools with their percentages are located in the district location under the evaluating performance levels. As for type of education variable, (77) count of both (scientific and literary) schools with (84.6%) percentage has the low performance, (408) schools with (82.4%) percentage has medium performance and (36) schools with (60%) percentage has high performance. The percentage of schools those are in the non-exemplary which are in the low performance is (92.3%), in the medium performance is (96.4%) and the schools has the high performance is (65.0%). It is described that most of schools in the KRG are non-exemplary schools. The count and percentage of gender of student shows that most of schools are coeducational schools in KRG, which they has maximum count and percentage than the schools for boys and girls at the all performance level and most of them are established in (2001-2010) period

with the large count and percentage than the others period at all performance level. Number of classes and teachers in the all schools are equal with (91) count. Nevertheless, they have different mean which the greater mean of number of classes is (13) classes for each school in medium performance level. The same expression is true for number of teacher's variable which the medium performance level has the greater mean which is (28) teacher for each school. Finally, average of Kurdish subject in scientific school's variable, the mean of (91)

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	905.463			
Final	387.703	517.760	12	.000

schools at the low performance level is (57.0). The mean of (495) schools at the medium performance level is (71.6) and the mean of (60) schools at the high-performance level is (83.8) which is the greater mean than the other level.

3.3 Results of Ordinal Logistic Regression

The model fitting information described in below table gives the -2log likelihood for intercept only and final models.

Table 3.4: Model Fitting Information

In Table 3.4, the initial log likelihood value (905.463) is a measure of constant or intercept only, but no independent variables. The final log likelihood value (387.703) is a measure of independent variables and it is computed after all the independent variables have been entered into the logistic regression. The difference between these two measures is the model chi-square value i.e. $(905.463 - 387.703 = 517.760)$ that is tested for statistical significance. The test statistic value, 517.760, has a significant level less than 0.05 i.e. $p = 0.00 < 0.05$. Thus, the null hypothesis that there is no difference between the null and final models is rejected and is concluded that there is evidence to support that there is a relationship between the dependent and explanatory variables.

From the practical data we can compute the normal person and deviance goodness of fit measure. Both of the goodness of fit statistic should be only for models that have reasonably large expected values in each cell.

Table 3.5: Goodness of Fit

	Chi-Square	df	Sig.
--	------------	----	------

Pearson	538.171	1278	1.000
Deviance	387.703	1278	1.000

From the table 3.5, the outcomes of our analysis recommend that the model is fit very well ($p > 0.05$) that is mean fail to reject the null hypothesis depending on the detected data. Correspondingly, the model fits effectively.

The correlation measure provided by ordinal logistic regression analysis that showed in table 3.6 is *pseudo R²*.

Table 3.6: Pseudo R-Square

Cox and Snell	.551
Nagelkerke	.731
McFadden	.572

The pseudo-Square as appearing in Table 3.7, accounts for the amount of variance explained in the dependent variable by the independent variables. The Cox and Snell, and Nagelkerke pseudo-*R²* in Table above recommend that the variation in the level of evaluation outcomes explained by the evaluation factors ranges between 73.1% and 57.2%. Thus, a relatively good level of variation is explained by the model. When we fitting a model with the nominal, ordinal and scale parameters likewise location parameters, parameter estimates for all of them are presented.

Table 3.7 presents the parameter estimates (also known as the coefficients of the model). In Parameter Estimates table we realize the coefficients, their standard errors, the Wald test and related p-values (Sig.), the 95% confidence interval of the coefficients and odds ratios. Then p-values less than α level they are statistically significant otherwise not.

Table 3.7: Parameter Estimation

Parameter Estimates								
		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Y1 = 0]	18.588	19.893	.873	1	.350	-20.402	57.577
	[Y1 = 1]	28.227	19.953	2.001	1	.157	-10.881	67.334
Location	X7	-.002	.010	.040	1	.842	-.021	.017
	X8	.122	.042	8.522	1	.004	.040	.205
	X9	-.038	.021	3.180	1	.075	-.079	.004
	X1	.408	.034	144.895	1	.000	.341	.474
	[X2=1]	.748	.880	.723	1	.395	-.977	2.474
	[X2=2]	0 ^a	.	.	0	.	.	.
	[X3=0]	-1.376	.522	6.950	1	.008	-2.399	-.353

[X3=1]	-1.152	.466	6.097	1	.014	-2.066	-.237
[X3=2]	-1.166	.423	7.592	1	.006	-1.996	-.337
[X3=3]	0 ^a	.	.	0	.	.	.
[X4=0]	.386	.361	1.145	1	.285	-.321	1.094
[X4=2]	0 ^a	.	.	0	.	.	.
[X5=1]	-2.436	.676	12.997	1	.000	-3.761	-1.112
[X5=2]	0 ^a	.	.	0	.	.	.
[X6=0]	-.289	.384	.567	1	.451	-1.042	.464
[X6=1]	.171	.376	.205	1	.650	-.567	.908
[X6=2]	0 ^a	.	.	0	.	.	.

The threshold coefficients are representing the intercepts, specifically the point (in terms of a logit) where Kurdish teachers evaluation might be predicted into the three categories. Estimates for the parameters obtained through the maximum likelihood estimation method for the final model are shown in table 3.7 and our equation can be written as:

$$\ln(\theta_1) = 18.588 + 0.408x_1 - (1.376 + 1.152 + 1.166)x_{3(0)(1)(2)} - 2.436x_5 + 0.122x_8 \quad (1)$$

$$\ln(\theta_2) = 28.227 + 0.408x_1 - (1.376 + 1.152 + 1.166)x_{3(0)(1)(2)} - 2.436x_5 + 0.122x_8 \quad (2)$$

In the above equations, the negative sign indicates that those variables have negative effects on Kurdish teacher evaluation. From table 3.7 location, status of school, gender of student, year of establishment and number of teachers are the factors that have negative effects on Kurdish teacher evaluation. The estimates labeled location are the coefficients for the predictor variables. From the observed significance levels in above table, we can see that some explanatory variables are factors that affect Kurdish teacher evaluation. Based on the small observed significance level, we can reject the null hypothesis. The estimates labeled location are the ones we are interested in. They are the coefficients for the predictor variables. The Wald statistic is the square of the ratio of the coefficient to its standard error. The significance of the Wald statistic in the column with heading sig (< 0.05) indicates the importance of the predictor variables in the model (we reject the Null hypothesis $H_0: \beta_i = 0$) and high values of the Wald statistic shows that the corresponding predictor variable is significant. The findings indicate that Kurdish teacher evaluation is associated with sector, location, type of education, status of school, gender of student, year of school establishment, number of classes, number of teachers and average of Kurdish subject (scientific branch). From these the independent variables that have no significance association with Kurdish teacher evaluation are year of school establishment, number of teachers, sector, type of education and gender of student. And consequently, number of classes, average of Kurdish subject (scientific branch), location and status of school are found to be more significantly associated with Kurdish teacher evaluation. By taking the proponent of the mutual estimate relative to a given predictor, i.e. taking e^{B_j} , we get an estimate of the common odds ratio that describes the comparative odds for X_j different by one unit. Values greater than one indicate that the variable in request increases the odds of being low performance evaluation and values between

0 and 1 indicate a decrease in the odds of being low performance evaluation. Thus, the above table proves the location, status of school, number of classes and average of Kurdish subject increases the odds of an individual teachers being low performance evaluation is just the complement of the odds of being level high performance evaluation. $\exp(\hat{\beta})$, it is a factor by which the odds of being a low performance evaluation of individual change when with independent variable increases by one unit (i.e. from 0 to 2). For example from Table 3.7 the value of the odd ratio $\exp(0.122) = 1.129$, for number of classes variable (X_8) indicates that if (X_8) increase one unite of odd ratio teachers have 1.129 times chances to move from low performance level to medium performance level. Similarly, the values odds ratio $\exp(0.408) = 1.503$, for change of average of Kurdish, $\exp(-1.376) = 0.252$, for change of location and $\exp(-2.436) = 0.087$ for change of status of school indicates that if those variables increase one unite of odd ratio teachers have 1.503, 0.252 and 0.087 consequently time chances to move from low performance level to medium performance level.

When we fit an ordinal regression, we assume that the relationship between the independent variables and the logits are the same for all the logits. That means the results are a set of parallel lines.

Table 3.8 test of parallel lines

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	387.703			
General	192.384 (a)	195.319 (b)	12	.000

The null hypothesis conditions that the location parameters are the equivalent across dependent categories, a. The log-likelihood value cannot be more increased after maximum number of step halving, b. The Chi-Square statistic is calculated based on the log-likelihood value of the last iteration of the general model. Rationality of the test is uncertain. C Link function: Logit. One of the assumptions underlying ordinal logistic regression is that the association between each pair of outcome groups is the equivalent. This is commonly referred to as the test of parallel lines because the null hypothesis conditions that the slope coefficients in the model are the equivalent across dependent categories and lines of the equivalent slope are parallel. If we fail to reject the null hypothesis, we conclude that the assumption holds. From the table 3.6 displays parallel line test for general model with chi square value 195.319 and p value = 0.000 which is a significance, we reject the null hypothesis. Consequently, there is enough evidence to reject the null hypothesis for general model. Thus, the proportional odds assumption appears to have not held for general model.

Table 3.9 shows the percentage of correct classification for each of low performance, medium performance and high performance and total percentage of classification.

Table 3.9 Classification table for last OLR model

Observed		Predicted				
		Low performance	Medium performance	High performance	Total	
		N	N	N	N	%
Kurdish teacher's Evaluating (Scientific branch)	Low performance	54	37	0	91	59.3%
	Medium performance	24	468	3	495	94.5%
	High performance	0	20	40	60	66.7%
	Overall Percentage				646	87.0%

From the above classification table, it appears that the percentage of classification of medium performance is higher than low performance and high performance with (94.5%) total percentage. The overall percentage of correct classification is about 87%, meaning that the Ordinal logistic regression model has an ability to predict teacher's performance level.

4. Conclusions and Recommendations

4.1 Conclusions

According to the study results using OLR analyses, the main conclusions are given below:

- 1- The overall classification of the predictive response category shows that our obtained predictive model could predict the teacher's performance level very well.
- 2- The evaluating of teacher's performance level is affected by some factors through their experiment in all schools. The average of Kurdish subject and status of school would greatly increase the chance of teachers to shift from low performance level to medium performance or high-performance level when compared to other variables depending on the practical data.
- 3- The independent variables that have no effects on the teacher's evaluation are sector, type of education, gender of student, year of establishment of school and number of teachers. Whereas, teacher's evaluation is strongly affected by location, status of school, number of classes, average of Kurdish subject sector.

4.2 Recommendations

Based on the consequences of this study, we recommend a number of interferences to be considering in the sixth grade scientific in Kurdish subject in the schools of the educational directorates of the KRG to evaluate teachers:

1. The educational directorates of the KRG should give importance on factors that affect teachers' evaluation.
2. The use of certain methods of Ordinal or multinomial logistic regression to evaluate teacher's performance level by educational directorates of MOE-KRG.
3. The directors of schools should give attention for teachers' evaluation and report of supervisors to solve such reproaches exist in the process of teachers evaluating.
4. The MOE-KRG should work on expanding the teachers evaluating field in the educational directorates.

References

- ADEJUMO, A. & ADETUNJI, A. 2013. Application of ordinal logistic regression in the study of students' performance. *Mathematical Theory and Modeling*, 3, 10-19.
- AGRESTI, A. 1989. Tutorial on modeling ordered categorical response data. *Psychological bulletin*, 105, 290.
- AGRESTI, A. 2007. An introduction to categorical data analysis. Wiley-Interscience.
- ANDERSON, J. & MOHAN, S. 1984. Sequential coding algorithms: A survey and cost analysis. *IEEE Transactions on communications*, 32, 169-176.
- ARCHER, K. J. & LEMESHOW, S. 2006. Goodness-of-fit test for a logistic regression model fitted using survey sample data. *The Stata Journal*, 6, 97-105.
- ASHFORD, J. 1959. An approach to the analysis of data for semi-quantal responses in biological assay. *Biometrics*, 15, 573-581.
- DAYTON, C. M. 1992. Logistic regression analysis. *Stat*, 474-574.
- GOODMAN, L. A. & KRUSKAL, W. H. 1979. Measures of association for cross classifications. *Measures of association for cross classifications*. Springer.
- GREENWOOD, C. & FAREWELL, V. 1988. A comparison of regression models for ordinal data in an analysis of transplanted-kidney function. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 325-335.
- HARRELL JR, F. E. 2015. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, Springer.
- HILBE, J. M. 2009. *Logistic regression models*, Chapman and hall/CRC.
- HOSMER JR, D. W., LEMESHOW, S. & STURDIVANT, R. X. 2013. *Applied logistic regression*, John Wiley & Sons.
- KLEINBAUM, D. G. & KLEIN, M. 2010. *Logistic Regression: A Self-Learning Text*, Springer Science & Business Media.
- MCCULLAGH, P. 1980. Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 42, 109-127.
- MCCULLAGH, P. & NELDER, J. 1989. *Generalized Linear Models* Second edition Chapman & Hall. London.
- SOAR, R. S., MEDLEY, D. M. & COKER, H. 1983. Teacher evaluation: A critique of currently used methods. *The Phi Delta Kappan*, 65, 239-246.
- VERNEZ, G., CULBERTSON, S., CONSTANT, L. & KARAM, R. 2016. *Initiatives to Improve Quality of Education in the Kurdistan Region-Iraq*, RAND Education.
- WALKER, S. H. & DUNCAN, D. B. 1967. Estimation of the probability of an event as a function of several independent variables. *Biometrika*, 54, 167-179.

الملخص

تعتمد وزارة التربية - حكومة إقليم كردستان / العراق حالياً فقط على درجات الطلبة في تقييم مستوى أداء المدرسين في جميع المواد الدراسية. الاعتماد على متغير واحد فقط لتحديد أداء المعلم ليس عادلاً ويجب حل هذه المشكلة إحصائياً من خلال إيجاد نموذج إحصائي جديد مقترح. لذلك ، حاول الباحث إيجاد بعض المتغيرات المتوفرة في وزارة التربية لجميع المدارس الثانوية من أجل استخدامها في النموذج المقترح للتنبؤ بمستوى أداء المدرسين بدلاً من النموذج القديم. تهدف هذه الدراسة إلى التنبؤ بمستوى أداء المدرسين في المدارس الثانوية (الصف الثاني عشر) في حكومة إقليم كردستان الذين يقومون بتدريس مادة اللغة الكردية، وكذلك تحليل آثار المتغيرات التي لها تأثير على تقييم مستوى أداء المدرسين اعتماداً على البيانات المتوفرة لدينا. استخدم الباحث طريقة الانحدار اللوجستي الترتيبي لإيجاد نموذج مقترح لتقييم أداء المدرسين في مادة اللغة الكردية باستخدام بيانات جميع المدارس الثانوية العلمية في حكومة إقليم كردستان (646 مدرسة ثانوية). تم تحليل حالة المدرسين من خلال اختيار تسعة متغيرات تتعلق بالمدارس الثانوية وهم القطاع ، الموقع الجغرافي ، نوع التعليم، حالة المدرسة، جنس المدرسة ، سنة تأسيس المدرسة، عدد الصفوف في المدرسة، عدد المدرسين في المدرسة، ومتوسط درجات الطالب في المادة اللغة الكردية. تم التوصل الى أن أربعة من هذه المتغيرات (معدل الطالب ، وعدد الصفوف ، والموقع الجغرافي وحالة المدرسة) لها تأثير كبير على تقييم المدرسين في الصف الثاني عشر في المدارس الثانوية العلمية. كما تم التوصل الى أن النسبة الاجمالية للتصنيف الصحيح هي حوالي 87٪، وهذا يعني أن نموذج الانحدار اللوجستي الترتيبي لديه القدرة على التنبؤ بمستوى أداء المدرسين بصورة ملائمة جداً.

پووخته

وهزارهتی پهروهردی حکومهتی ههریمی کوردستان / عیراق ، له هه‌سه‌نگاندنی مامۆستایان ته‌نھا پشت به تیکرای نمره‌ی قوتابیان ده‌به‌ستیت بۆ ههر وانه‌یه‌ک. به‌لام ئایا ته‌نھا له‌به‌رچاوگرتنی تیکرای نمره‌ی قوتابیان به‌سه بۆه‌سه‌نگاندنی مامۆستایان؟ بۆیه پێویسته رینگه چاره‌یه‌کی تر له رینگه‌ی به‌کارهینانی شیکردنه‌وه ئامارییه‌کان بۆ ئه‌و گرفته بدۆزینه‌وه ئه‌ویش به دۆزینه‌وه‌ی مۆدیلێکی ئاماری نوێی پێشنیارکراو به‌مه‌به‌ستی هه‌سه‌نگاندنیکی دروستر بۆ مامۆستایانی ئاماده‌یه‌یه‌کان (پۆلی دوازه‌ی زانستی).

تویژه‌ر ژماره‌یه‌ک گۆراوی تری هاوشان له‌گه‌ڵ تیکرای نمره‌ی وانه‌ی زمانی کوردی له‌و داتاو زانیاریانه‌ی له وه‌زاره‌تی پهروهرده‌ به‌رده‌ست بوو ده‌ست نیشان کردووه به‌مه‌به‌ستی به‌کارهینانیان بۆدۆزینه‌وه‌ی مۆدیلێکی نوێی ئاماری پێشنیارکراو به‌مه‌به‌ستی هه‌سه‌نگاندنی ئاستی مامۆستایان له‌ شوینی رینگای کۆنی په‌یره‌وکراو.

ئامانجی ئه‌و تویژه‌یه‌وه‌یه پێشبینی کردن و هه‌سه‌نگاندنی ئاستی مامۆستایانی زمانی کوردی یه‌ له پۆلی دوازه‌ی زانستی له قوتابخانه ئاماده‌یه‌یه‌کانی ههریمی کوردستان که ژماره‌یان (٦٤٦) قوتابخانه‌یه وه‌هه‌روه‌ها شیکردنه‌وه‌ی ئه‌و گۆراوانه‌ی کاریگه‌ری دروست ده‌که‌ن له‌سه‌ر هه‌سه‌نگاندنی مامۆستایان. بۆ ئه‌و مه‌به‌سته تویژه‌ر رینگای شیکردنه‌وه‌ی ئامار (Ordinal) logistic regression ی به‌کارهیناوه بۆ دۆزینه‌وه‌ی مۆدیله ئامارییه نوێیه‌که. شیکردنه‌وه بۆ دۆخی مامۆستایان کراوه له رینگه‌ی ده‌ست نیشان کردنی (٩) گۆراوی ترئه‌وانیش: که‌رتی قوتابخانه، پینگه‌ی جوگرافی قوتابخانه، جووری خویندن له قوتابخانه، دۆخی قوتابخانه، په‌گه‌زی قوتابخانه، سالی دامه‌زراندنی قوتابخانه، ژماره‌ی مامۆستایان له قوتابخانه، ژماره‌ی پۆله‌کان له قوتابخانه، له‌گه‌ڵ تیکرای نمره‌ی وانه‌ی زمانی کوردی قوتابیان له قوتابخانه. تویژه‌ر گه‌یشه‌ ئه‌و ئه‌نجامه‌ی که (٤) له‌و گۆراوانه (تیکرای نمره‌ی قوتابیان، ژماره‌ی پۆله‌کان، پینگه‌ی جوگرافی له‌گه‌ڵ دۆخی قوتابخانه)

کاریگه‌رییان هه‌یه له‌سه‌ر ئه‌نجامدانی هه‌سه‌نگاندنی ئاستی مامۆستایان . هه‌روه‌ها کۆی گشتی رێژه‌ی پۆلینی دروست ده‌گاته نزیکه‌ی (87٪)، ئه‌وه ده‌رده‌خات نمونه‌ی پیشنیارکراوی به‌ده‌ست هاتوو زۆر به‌باشی توانای هه‌سه‌نگاندنی مامۆستایانی هه‌یه.