Characteristics of Ordinal Data in Trend Odds Model

R N Rachmawati¹,², S Kusman¹*, K Anang¹
¹ Statistics Department, Bogor Agricultural University, Jl Raya Darmaga, Bogor 16680, Indonesia
² Statistics Department, Bina Nusantara University, Jl Kebun Jeruk Raya No. 27, Jakarta 11530, Indonesia
*email: kusmansadik@gmail.com

Abstract. Ordinal data widely appear in many scientific areas, are often assumed with proportional odds in ordinal logistic regression. When the ordinal data have monotone manner across cut-points, proportional odds assumption become less appropriate. The lack of proportionality may be possible captured using trend odds model which have constrained structural relationship between the odds and the cut-points. This paper describes briefly cumulative odds which are the foundation of trend odds model and demonstrates algebraically the trend odds model related to latent logistic and uniform distribution to obtain the explicit form of slope of the trend. A data set is used to illustrate the interpretation of trend odds model, we apply this model to cigarettes consumption which has monotone manner across cut-points that seems the proportional assumption is more likely not appropriate. However, characteristic of ordinal data that fit trend odds model properly must be specified because only monotone manner is insufficient. We perform simulation using PROC NLMIXED in SAS system and find that ordinal data must significantly monotone for unexposed group and the ratio between covariates is higher for higher level cut-points.

Keywords. logistic regression, proportional odds, cigarettes consumption, latent uniform distribution, PROC NLMIXED

1. Introduction
Generalized Linear Model (GLM) is an extension of linear model which encompasses non-normal response distributions and link functions of the mean equated to the linear predictor [1]. Example of GLM is logistic models with the logit-link function for a binomial (binary) response. Multicategory responses are generalization of binary responses that present generalizations of logistic regression for multinomial response variables that follows multinomial distribution. Generally, the multinomial distribution does not meet the exponential family, however, its relations to the Poisson distribution assures that GLM can be applied [2]. For multinomial responses, the logistic regression models are separately available for nominal and for ordinal response variables. For ordinal response, models have terms that reflect ordinal characteristics such as a monotone trend, whereby responses tend to fall in higher (or lower) categories as the value of an explanatory variable increases. Such models are more parsimonious than models for nominal responses, because potentially they have many fewer parameters [3].

In ordinal data, there is an obvious natural order among the response categories and often used in many scientific fields such as opinion polls in market research [4], psychology [5], agricultural [6] and
exposure science [7]. This data has ability to explain in great details the characteristic of observation. Proportional odds model (POM) is frequently used for analyzing ordinal data. Odds are considered proportional when all possible outcome level have the same value, or in other words, proportionality means the log odds ratio does not depend on its cut-points and proportional to distance between covariates. If the proportional assumption does not hold then an alternative is to specify a non-proportional odds model, where the regression parameters are allowed vary depending on the level of the response [8]. The lack of proportionality might be relaxed by introducing constrained structural relationship between the odds and the cut-points [9].

Capuano and Dawson [10] introduce constrained non-proportional odds model called the Trend Odds Model (TOM). TOM is very close related in health research phenomena, because there will be differences from people who have low immunity against disease and high for those who have higher immunity, or in general cases there will be an increased risk to people who are longer exposed and otherwise to fewer exposed. Another simple example is in number of cigarette consumption per day with the lowest category being zero, the move from zero to a higher category as an increase in variable such as age could have a dramatic effect on whether a person is a smoker or not. These examples show differences of log odds ratio for higher levels so that the proportional assumption is highly not appropriate and simultaneously explained that the excellence of TOM is its ability to identify high-risk groups.

TOM purpose to improve the ordinal data that have monotone manner across cut-points. Monotone means the value of ordinal data consistently increase or decrease for higher level or cut-points so that proportionality assumption seems to be violated in such cases. When value of ordinal data increase or decrease in monotone manner, it is very possible that our model influenced by a trend parameter. TOM accommodate one additional parameter that is a trend parameter to capture the lack of proportionality in ordinal logistic regression.

Using the foundation of Capuano and Dawson [10], stated that TOM appropriate in term of modeling ordinal data with monotone characteristic, this paper aims to: (1) briefly review the notation of cumulative odds model which is the main idea of Non-Propotional Odds Model (NPOM), POM and also TOM to construct the odds ratio between covariates, (2) show how latent logistic and uniform distributions relate to the TOM, and give the explicit form of slope of the trend, because by getting the explicit form we be sure that the TOM holds and appropriate for data with those specified underlying distributions, (3) we use hypothetical cigarette consumption data set which have monotone characteristic to modeled using TOM and (4) we present simulation using PROC NLMIXED to discover kind of ordinal data characteristic that produce significant trend parameter.

2. Cumulative odds models: NPOM, POM and TOM
Suppose there is an observed ordinal outcome $Y$ with $K+1$ categories $Y=0,1,2,K$ and for simplicity we use single covariate $X$. We define $\Psi_{kx}$ as the log cumulative odds which the odds of $Y$ having a level at least as high as cut-point $k$ given $X$. That is, in mathematical notation $\Psi_{kx}$ can be written as

$$\Psi_{kx} = \log \frac{P(Y \geq k \mid X = x)}{1 - P(Y \geq k \mid X = x)}. \tag{1}$$

The general form of cumulative odds is known as NPOM which has unconstrained cumulative odds model represented as

$$\Psi_{kx} = \alpha_k + \beta_k X, \quad k = 1, \ldots, K. \tag{2}$$

Note that non-proportional model in (2) marked in coefficient parameter $\beta$ are indexed by $k$, so that the odds ratio have different values for $k = 1, \ldots, K$. Suppose that $X$ represents an exposure status that is being investigated, with $X = 1$ for exposed and $X = 0$ for unexposed, with the result that different odds ratio for different cut-points are $\theta_1 = e^{\beta_1}, \theta_2 = e^{\beta_2}, \ldots, \theta_K = e^{\beta_K}$, wherein NPOM odds ratio for exposed
and unexposed depend on its cut-points. Using NPOM we need to estimate $K$ additional parameters to calculate the odds ratio, of course if we have limitation in number of parameters to be estimated, NPOM into an inefficient model.

Proportional odds model (POM) conceptually introduced by Aitchison and Silvey (1957) [11], and further developed by McCullagh (1980) [12]. In contrast, assumption that frequently used are POM where coefficient parameter $\beta$ have the same value for each cut-point, from (2) log odds ratio under POM stated as

$$\Psi_{kt} = \alpha_k + \beta X, \quad k = 1, \ldots, K.$$  

(3)

With POM, the odds of being at or above any cut-point are assumed to be the same for all cut-points; this is known as the ‘proportional odds’ assumption. In other words, instead of having several different odds ratios for exposed versus unexposed, a single odds ratio is calculated as $\theta = e^\beta$. POM make a strong assumption and therefore it widely used in modelling ordinal data because POM is easy to fit and interpret [13].

Having model that is easy to fit and interpret like POM is an efficient way to describe ordinal data, however, if proportionality assumption is not met we may using TOM to accomplish the lack of proportionality with one additional trend parameter $\gamma$. TOM use intermediate constrained structural relationship between the odds and cut-points [14-16]. In general, this can be accomplished by adding parameter $\gamma$ and a multiplicative scalar that varies with the cut-point, say $t_k$ [10]. Capuano and Dawson [10] define the constrained cumulative odds model with a single predictor case as

$$\Psi_{kt} = \alpha_k + (\beta + \gamma t_k) X, \quad k = 1, \ldots, K.$$  

(4)

Note that scalar $t_k$ in (4) could potentially have any set of scalars, for example, the scalar values of $t_1 = -1, t_2 = 1, t_3 = 0$ and $t_4 = 1$ would yield the odds ratios of $\theta_0 = e^\beta e^{-\gamma}, \theta_2 = e^\beta e^{\gamma}$ and $\theta_4 = e^\beta$, in other words under non-proportional odds models, the scalar set can be unstructured [Capuano, 2013]. The TOM is structured cumulative odds where scalars $t_k$ are selected to have monotone structure so that $t_1 \leq t_2 \leq \ldots \leq t_K$. Because cut-points in ordinal data are scalars that have monotone manner, Capuano and Dawson [10] select its cut-points $c_k$ as structured scalars set values, so under TOM equation (4) can be stated as

$$\Psi_{kt} = \alpha_k + (\beta + \gamma c_k) X, \quad k = 1, \ldots, K, \quad c_k < c_{k+1}.$$  

(5)

Note that in (5) TOM holds if and only if trend parameter is non-zero, $\gamma \neq 0$, the parameter space of $\alpha_k$ and $\beta$ similar to the POM in any real number from minus infinity to infinity and fitted using the same fitting procedure in POM. The odds ratio for TOM are decrease or increase exponentially in monotone manner, related odds ratio can be calculated from the baseline odds ratio. We called baseline odds ratio because the odds ratio is calculated at the first cut-point, so that $\theta = e^\beta$, the odds ratio for higher level cut-points are calculated as

$$\theta_k = \theta e^{\gamma (k-1)}, \quad k = 1, 2, \ldots, K.$$  

For simple example, if we have four cut-points and four odds ratios, and we have increasing trend with positive $\gamma = 2$ and baseline $\theta = 1.25$, so we will have the first odds ratio is the baseline, the second odds ratio is $\theta_2 = 1.25e^2$ and $\theta_3 = 1.25e^4$, $\theta_4 = 1.25e^6$ for the third and fourth odds ratios respectively.

3. Latent Distributions for the TOM

Suppose that the ordinal variable $Y$ based on the latent variable $Y^*$ according to specific $c_k$ and let

$$Y = k \leftrightarrow Y^* \in (c_k, c_{k+1}] \quad \text{and} \quad Y \geq k \leftrightarrow Y^* \geq c_k \quad \text{for} \quad k = 0, 1, 2, \ldots, K.$$  

Suppose that a response variable $Y^*$ follows a logistic distribution with scale parameters $\theta$ and location parameters $l$, being functions of a predictor variable $X$, so that $Y^* \mid X = x : \text{Logistic}(l_x, s_x)$.
We know that $E(Y^* \mid X = x) = l_x$ and $\text{Var}(Y^* \mid X = x) = \frac{\pi^2}{3} s_x^2$. For specific cut-point $c_k$ we have cumulative density function of latent logistic distribution $Y^*$ as

$$F(Y^* = c_k \mid X = x) = P(Y^* < c_k \mid X = x) = \frac{1}{1 + \exp\left(-\frac{c_k - l_x}{s_x}\right)}.$$  

(6)

By taking the logit function, we have the log odds as

$$\Psi_{kx} = \log \frac{P(Y^* \geq c_k \mid X = x)}{1 - P(Y^* \geq c_k \mid X = x)} = \frac{l_x - c_k}{s_x}.$$  

(7)

We are interested in determining the explicit form of slope of the trend, because it is one of the conclusive evidences to show that TOM holds for latent logistic distribution. Slope of the trend obtained from the first derivative of log odds ratio for different covariates as stated below

$$\frac{d}{dc_k}(\Psi_{kx} - \Psi_{k0}).$$  

(8)

Note that TOM is structured cumulative odds model which one of NPOM assumption, so that the log odds ratio defined in (8) are the function of specific cut-point $c_k$. If equation (8) has zero value, then TOM is not hold, in other words, the log odds ratio is constant and not the function of specific cut-point $c_k$, thus POM is hold for that case. If equation (8) have constant value, TOM is hold with linear trend.

In latent logistic distribution, we have log odds ratio written as

$$\Psi_{kx} - \Psi_{k0} = \frac{l_x - c_k}{s_x} - \frac{l_0 - c_k}{s_0}.$$  

(9)

If the scale parameters have the same value i.e. $s_x = s_0 = s$, then equation (9) become

$$\Psi_{kx} - \Psi_{k0} = \frac{l_x - c_k}{s} - \frac{l_0 - c_k}{s} = \frac{l_x - l_0}{s},$$  

(10)

so that for known location and scale parameters we have log odds ratio in (10) is constant, in other words, its first derivative of cut-point $c_k$ is zero then POM is hold for this case. The opposite happened when there are shift or different in scale parameters with or without shift in location parameters, we obtain

$$\frac{d}{dc_k}(\Psi_{kx} - \Psi_{k0}) = \frac{d}{dc_k}\left(\frac{l_x - c_k}{s} - \frac{l_0 - c_k}{s_0}\right) = \frac{1}{s} - \frac{1}{s_0},$$  

(11)

constant value for the first derivative of log odds ratio as stated in (11). So, for latent logistic distribution with shift in scale parameters TOM assumption holds with linear trend. The slope of the trend if there is a shift in both location and scale parameters are summarized in Table 1.

Latent logistic distribution as the foundation of underlying distribution in POM as we can see in [12] and [10]. We also interested in determining slope of the trend using latent continuous uniform distribution. Although uniform denote the simplest continuous distribution, however, from a theoretical perspective, this distribution is a key on a risk analysis in order to generate random samples from other distribution. Suppose we have latent uniform continuous distribution $Y^* \mid X = x$: Uniform$(a_x, b_x)$ at a given $X$, so we have cumulative density function as

$$F(Y^* = c_k \mid X = x) = P(Y^* < c_k \mid X = x) = \int_{a_x}^{c_k} \frac{1}{b_x - a_x} dy = \frac{c_k - a_x}{b_x - a_x}.$$  

(12)

By taking the logit function, we have the log odds as
\[ \Phi_{kx} = \frac{1 - \left( \frac{c_k - a_x}{b_x - a_x} \right)}{\left( \frac{c_k - a_x}{b_x - a_x} \right)} \cdot \frac{b_x - c_k}{c_k - a_x}. \]

If we have the same lower bound parameters i.e. \( a_x = a_0 = a \), then we have the log odds ratio

\[ \Phi_{kx} - \Phi_{k0} = \frac{b_x - c_k}{c_k - a} - \frac{b_0 - c_k}{c_k - a} = \frac{b_x - b_0}{c_k - a}. \]  

Note that in equation (13) is the function of cut-point \( c_k \), therefore TOM holds for latent continuous uniform distribution with shift in lower bound parameters. For known upper bound parameters and specific cut-point \( c_k \) we have the slope of the log odds ratio has non-linear value as stated in equation (14) as

\[ \frac{d}{dc_k} (\Phi_{kx} - \Phi_{k0}) = -\frac{b_x}{(c_k - a)^2}. \]  

We also consider if there is a shift in lower bounds i.e. \( a_x \neq a_0 \) but constant in upper bounds i.e. \( b_x = b_0 = b \) then we have non-linear slope of the trend which is summarized in Table 1.

**Table 1.** Underlying latent Logistic and Uniform distribution.

| Parameter shift                      | Log odds ratio                                      | Assumption holds |
|--------------------------------------|-----------------------------------------------------|------------------|
| Logistic Location (\( l_x \))        | \( \frac{l_x - l_0}{s} \)                          | POM              |
| Logistic Scale (\( s_x \))           | \( \left( m - c_k \right) \left( \frac{1}{s_0} - \frac{1}{s_x} \right) \) | TOM – Linear     |
| Logistic Location and scale (\( l_x, s_x \)) | \( \frac{l_x - c_k}{s_x} - \frac{l_0 - c_k}{s_0} \) | TOM – Linear     |
| Uniform Upper bound (\( a_x \))      | \( \frac{b_x - b_0}{c_k - a} \)                    | TOM – Non-Linear |
| Uniform Lower bound (\( b_x \))      | \( \frac{b - c_k}{c_k - a} - \frac{b - c_k}{c_k - a_0} \) | TOM – Non-Linear |
| Uniform Lower and upper bound (\( a_x, b_x \)) | \( \frac{b_x - c_k}{c_k - a_x} - \frac{b_0 - c_k}{c_k - a_0} \) | TOM – Non-Linear |

**4. Application to Cigarette Consumption**

We already known along with the rough effect of cigarette consumption for our health. Not only for the active smokers, the passive smokers can also have the same health effect such as Tuberculosis, lung cancer, heart disease, complications that lead to death from long-term and high intensity of smoking. According to WHO data, Indonesia is the third country with the largest number of smokers in the world after China and India. Increased consumption of cigarettes has an impact on the higher burden of smoking diseases and increased mortality from smoking. By 2030 it is estimated that the death rate of smokers that world will reach 10 million people and 70% of them come from developing countries.
As developing country, today, more than 36% of Indonesia’s population is categorized as smokers. Among young people aged 13-15 years, there are 20% of smokers, 41% of them are teenagers and 3.5% of them are girls [17]. We have a survey of 2.330 smokers in Indonesia, smokers surveyed filled out the online questionnaire that opened from January - March 2014. The data taken from Nusaresearch Indonesia as the pioneer of Online Research Company in Indonesia. We interested to discover the information how education effect cigarette consumption in young productive aged, i.e. 17 years and above. Based on this purpose, we dichotomized the covariates into two categories i.e. higher education and lower education. Higher education means smokers who have collage education background or above, and lower education means smokers who have high school education background or below.

Table 2. Cigarette consumption/day and odds ratio

| Education   | 1-9   | 10-14 | 15-19 | >20 |
|-------------|-------|-------|-------|-----|
| Higher ed.  | 845   | 287   | 142   | 115 |
| Lower ed.   | 504   | 239   | 120   | 78  |

Odds Ratio

|       | TOM  | POM  |
|-------|------|------|
| Odds  | Ratio| Ratio|
| TOM   | 0.74 | 0.78 |
| POM   | 0.64 | 0.78 |

As we can see in Table 2, we divide our level response into four responses, in other words, we will have three cut-points. Our level responses state cigarette consumption per day which the first level response is the least amount of cigarette consumption only 1 until 9 cigarettes per day, and this amount getting higher for the higher response levels. For lower education as the reference, under TOM assumption we have significantly negative trend odds parameter $\gamma = -0.1452$ (SE(\(\gamma\)) = 0.0703, p-value = 0.0728) and coefficient parameter $\beta = -0.299$, which yields $e^\beta = 0.74$ as our baseline in the first cut-point, $e^\beta e^\gamma = 0.64$ and $e^\beta e^{2\gamma} = 0.55$ are odds ratios for the second and third cut-points respectively. As the representation from our data that have decreasingly monotone manner, we also have decreasingly monotone trend odds ratio. Under POM assumption we have the same odds ratio for all cut-points i.e. $e^\beta = 0.78$. POM representation is very different with TOM, under POM we have constant odds ratio even though our data shows the decreasing trend effect for both education in each cigarette consumption level. However, it is quite surprising that, smokers who have lower education consume cigarette more than higher educated smokers in all response levels.

From our data that have decreasing monotone manner which according to the theory, this kind of data is very applicable modelled using TOM, we expect that we have bigger trend parameter value than the currently obtained i.e. $\gamma = -0.1452$. Although this value has significant effect under $\alpha = 0.1$ significance level, trend parameter that closes to zero will reduce the meaning of the effect of the trend itself. Furthermore, adding one additional parameter under TOM should be able to increase or improve the fit model accuracy than POM assumption, however with small trend parameter values, indication to get better fit still needs to be discussed. For these reasons, in the next section, we are interested to determine the characteristic of ordinal data that really appropriate to be modelled using TOM, because from our cigarette consumption data for example, it seems that data that have only monotone manner is insufficient. Appropriate for this context we mean the trend parameter has significantly effect with non-zero (and not close to zero) and have better fit than POM.

5. Ordinal data characteristics in simulation study

We perform simulation using NLMIXED procedure in SAS system to obtain the characteristics of ordinal data that appropriate to be modelled using TOM. SAS have several procedures which offer programming capabilities to analyse ordinal data including the MODEL, NLIN and NLMIXED.
procedures. Based on the complexity of the models and the number of statement the NLMIXED procedure efficiently estimates these ordinal data models with maximum likelihood and if present can also model the random effects [18].

Simulated scenarios included different sample sizes, consists of 50, 100 and 400 per group, different reference of multinomial distribution, and different numbers of data sets \((n_1 = 1000 \text{ and } n_2 = 600)\) were randomly generated from a multinomial distribution using different seeds. The data were modelled using TOM and POM and we evaluated the percentage of \(\beta_{\text{TOM}}, \beta_{\text{POM}}\) and \(\gamma\) parameters that have significantly effect under \(\alpha = 0.1\) significance level. Table 3 shows generated multinomial distribution and its specification for each model. Note that from Table 3 we have generated Model 1 until Model 5, using two covariates, with \(X = 1\) as the exposed and \(X = 0\) as the unexposed status, with four level responses that we set from non-significant to significant decreasing for unexposed group probability and from smaller ratio to higher ratio between covariates. We have generated decreasingly monotone multinomial distribution in order to be modelled using TOM.

### Table 3. Generated Multinomial Distributions

| Specification | Exposed | Unexposed |
|---------------|---------|-----------|
| Model 1       | \(Y : \text{m}(n_1,\{(0.3,0.26,0.24,0.2)\})\) | \(Y : \text{m}(n_1,\{(0.3,0.25,0.24,0.21)\})\) |
| Model 2       | \(Y : \text{m}(n_1,\{(0.4,0.3,0.2,0.1)\})\) | \(Y : \text{m}(n_1,\{(0.4,0.28,0.2,0.12)\})\) |
| Model 3       | \(Y : \text{m}(n_1,\{(0.4,0.3,0.2,0.1)\})\) | \(Y : \text{m}(n_1,\{(0.4,0.2,0.12,0.08)\})\) |
| Model 4       | \(Y : \text{m}(n_1,\{(0.4,0.3,0.2,0.1)\})\) | \(Y : \text{m}(n_2,\{(0.4,0.2,0.12,0.08)\})\) |
| Model 5       | \(Y : \text{m}(n_1,\{(0.4,0.3,0.2,0.1)\})\) | \(Y : \text{m}(n_2,\{(0.5,0.4,0.09,0.01)\})\). |

Table 4 shows the simulation results of modelling using TOM and POM for generated multinomial distributions. In Model 1 we set the same numbers of data set for exposed and unexposed group i.e. \(n_1 = 1000\) and non-significant decreasing probability that almost have the same value for higher response levels. Then under Model 1 with TOM assumption, we have only 4% significant trend and only 10% significant \(\beta\) parameters from 1000 multinomial distribution of ordinal data and 50 repetitions. We make more repetitions in expects of obtaining more percentage of significant parameters value. Using 100 and 400 repetitions we only have 4% and 7% significant trend, 10% and 10.3% significant \(\beta\) parameters respectively, of course this is not the significant addition compared to addition of simulated repetitions. The same results occur using POM assumption, only 8% \(\beta\) parameters that have significant effect and in this case, the percentage of significant parameters do not show an increase pattern for higher repetition values.

In Model 2, we set this model decreasingly monotone manner which more significant decrease from Model 1. Under TOM assumption, Model 2 produce almost the same results in percentage of significant trend parameter with Model 1, yet the percentage of significant \(\beta\) parameters is higher than Model 1. The same with Model 1, under TOM and POM assumptions, addition of significant parameters do not show an increase pattern for higher repetition values.
Table 4. Percentage of significant parameters

| Parameter | TOM | POM |
|-----------|-----|-----|
|           | γ%  | β%  | γ%  | β%  | γ%  | β%  |
| Sample sizes per group | | | | | | |
| 50 | 100 | 400 | 50 | 100 | 400 | 50 | 100 | 400 |
| Model 1 | 4.0 | 4.0 | 7.3 | 10.0 | 10.0 | 10.3 | 8.0 | 6.0 | 8.5 |
| Model 2 | 4.0 | 7.0 | 6.0 | 40.0 | 39.0 | 36.5 | 6.0 | 17.0 | 14.3 |
| Model 3 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Model 4 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| Model 5 | 100 | 98.0 | 97.5 | 100 | 100 | 100 | 100 | 100 | 100 |

Results from Model 1 and Model 2 show that in its applications under TOM, ordinal data that only have monotone manner across cut-points (or response levels) is insufficient to obtain significant effect for the γ and β parameters. The importance of obtaining significant parameters have been explained in section 4. Then, in this simulation study, we modify the previous models that have significant decreasing monotone manner. We generate Model 3 that we set the same distribution from Model 2 for exposed group, and decrease more significant than Model 2 for unexposed group. For this model, γ and β parameters are 100% significantly effect in the logistic regression using TOM and POM assumptions.

In Model 4 we generate the exposed group which has the same distribution from Model 2 and 3, and more significant decreasing than Model 3 for unexposed group by setting different smaller number data set i.e. n = 600. The smaller number data set in Model 4 represents higher ratio between covariates (i.e. exposed and unexposed group) than ratio between covariates in Model 3. Under TOM and POM, Model 3 and 4 consistently produce 100% significant γ and β parameters for all repetitions. We set Model 5 for exposed group which has the same distribution in Model 4, but for unexposed group we set more significant decreasing than Model 4 by setting more significant decreasing probability and also smaller number data set. This model produces 100% significant γ and β parameters using TOM and POM assumptions except for TOM in 100 and 400 repetitions. Although this model shows decreasing pattern in percentage of significant γ parameters, the resulting percentage is still proper compared to the increasing of repetitions. From Table 4, the simulation results show that more significant decreasing probability for unexposed group and higher ratio between covariates will produce higher percentage number of significant parameters.

Table 5. Models’ fit statistics

| Model | TOM-PO | 50 | 100 | 400 |
|-------|--------|----|-----|-----|
|       |        |    |     |     |
| -2RLL | AIC    | BIC | -2RLL | AIC | BIC | -2RLL | AIC | BIC |
| Model 1 | -1.32  | 0.68 | 0.76 | -1.30 | 0.70 | 0.78 | -1.33 | 0.67 | 0.75 |
| Model 2 | -3.25  | -1.25 | -1.17 | -3.09 | -1.09 | -1.01 | -3.13 | -1.13 | -1.05 |
| Model 3 | 0.00   | 2.00 | 1.39 | 0.00 | 2.00 | 1.39 | 0.00 | 2.00 | 1.39 |
| Model 4 | 0.00   | 2.00 | 1.39 | 0.00 | 2.00 | 1.39 | 0.00 | 2.00 | 1.39 |
| Model 5 | -63.57 | -61.57 | -61.50 | -59.93 | -57.93 | -57.85 | -60.27 | -58.27 | -58.19 |
To ensure that ordinal data that have significant decreasing unexposed group probability and have higher ratio between covariates give more applicable to be modelled using TOM than POM, we compare the parameter estimates using fit indices criteria. PROC NLMIXED produce three fit indices that most widely used in comparing the models’ fit statistics, i.e. -2RLL, AIC and BIC. All these criteria are in smaller is better form. Based on these three criteria we compare Model 1 until Model 5 under TOM and POM. The results are presented in Table 5.

Note that in Table 5 we compare models’ fit using the average value of -2RLL, AIC and BIC criteria under TOM and POM assumptions in each repetition. The negative value indicates the ordinal data more appropriate when modelled using TOM, the positive value indicates the opposite meaning. Model 1 until Model 4 shows that TOM do not show better fit than POM, especially in Model 3 which yields identical results with Model 4, POM has better performance than TOM. Unlike other models, Model 5 under TOM produce better performance than POM. This result increases our confidence that ordinal data that have significant decreasing probability for unexposed group and have higher ratio between covariates will yield better ordinal logistic regression model using TOM. In other words, TOM can capture the lack of proportionality in odds ratio representation for these specified ordinal data characteristics.

6. Discussions

This paper describes trend odds model for ordinal data that have monotone manner across cut-points with logistic and uniform underlying latent distributions. The trend odds model can describe non-proportional odds when latent variable distribution is not known [10]. When fitting data with statistical models, there is often a compromise between model adequate fit and simplicity, although proportional odds gives simple interpretation, it may give inadequate fit when the proportionality assumption is not met. Trend odds assumption added the complexity of only one additional parameter and works properly than proportional odds for monotone ordinal data characteristic [10].

We explore how education effect the cigarette consumption using trend odds model. We have decreasing monotone numbers of cigarette consumption per day for both higher and lower educations. The negative trend resulted from this observations and yield decreasing odds ratio for more cigarette consumption, and we find the fact that lower educated smokers tend to consume more cigarette per day than higher educated smokers.

In trend odds model application, many ordinal data that have monotone manner across cut-points could yield non-significant trend parameter and probably do not give better fit than proportional odds. We explore and examine using PROC NLMIXED in our simulation study, designed with various scenarios and result that significantly monotone unexposed group probability and higher ratio between covariates produce excellence fit statistics than proportional odds assumption and also give significant trend parameter. Odds ratio in trend odds model increase or decrease from higher to fewer levels (or vice versa) according to an exponential function, so that higher ratio between covariates is reasonable characteristic condition for ordinal data modelled using trend odds assumption.

In this paper, we explore trend odds model with dichotomized covariates and focused on models with a single predictor variable. Exploring more than a single predictor and the interaction between variables could be one of an interested further discussion because it aligns with NLMIXED procedures that allows researcher to include interaction between covariates. Continuous covariate is ubiquitous in ordinal logistic regression literature, for example, the exposed responses can be ordered using normal, mild and severe levels that in exposure science often measured in continuous variables. So that, specified characteristics for continuous covariates under trend odds model is also interested to be discussed. Although not the scope of this work, it is important to explore different link families to obtain better fit or proportionality, while in this paper we use logit function in its estimation processes. This far, we use trend odds model to covers the lack of proportionality assumption, it also be an interested research that apply proportional odds model and trend odds model simultaneously and allowing random effect to ordinal data [19].
Acknowledgments
This paper would not have been possible without data support from Nusaresearch Company, as one of Pioneer and Biggest Online Research in Indonesia. Data support from Nusaresearch is highly appreciated and hope can be used for useful activity and decision through this research. The authors also want to thank LPDP (Lembaga Pengelola Dana Pendidikan) and Kemenristek DIKTI for supporting BUDI (Beasiswa Unggulan Dosen Indonesia) Scholarship for the first author.

References
[1] A Agresti 2015 Foundation of Linear and Generalized Linear Models John Wiley & Sons, Inc., Hoboken, New Jersey.
[2] D W Hosmer and S Lemeshow 2013 Applied Logistic Regression John Wiley & Sons, Inc. New York
[3] A Agresti 2013 Cathegorical Data Analysis John Wiley & Sons, Inc. Publication
[4] F Winter 2010 Primary Scales of Measurement Wiley International Encyclopedia of Marketing
[5] Camparo J 2013 A geometrical approach to the ordinal data of likert scaling and attitude measurements: The density matrix in psychology J of Math. Psy. 57 29
[6] W M Lisa, M Jane, Y Fang, S Oyewale, J R Debra and D L Tricia 2017 A principal factor analysis to characterize agricultural exposures among Nebraska veterans J of Exp. Sci. & Envi. Epid. 27 214
[7] S Madhuri, IK Leeka, A A Onyebuchi, A D Hozefa and O Jørn 2014 Complexities of sibling analysis when exposures and outcomes change with time and birth order J of Exp. Sci. & Envi. Epid. 24 482
[8] T J McKinley, M Morters and J L N Wood 2015 Bayesian Model Choice in Cumulative Link Ordinal Regression Models Int. Soc. for Bayesian Anal. 10 1
[9] A A O’Connell 2006 Logistic Regression Models for Ordinal Response Variables. Quantitative Applications in the Social Sciences California SAGE Publications
[10] A W Capuano and J D Dawson 2013 The trend odds model for ordinal data Stat. in Med. 32 2250
[11] J Aitchison, S D Silvey 1957 The generalization of probit analysis to the case of multiple responses Biometrika 44 131
[12] P McCullagh 1980 Regression models for ordinal data. J of the Roy. Stat. Soc. 42 109
[13] A W Capuano, J D Dawson, G C Gray 2007 Maximizing power in seroepidemiological studies through the use of the proportional odds model Influenza and Other Respiratory Viruses 1 87
[14] C V Ananth, D G Kleinbaum 1997 Regression models for ordinal responses: a review of methods and applications Int. J. of Epid. 26 1323
[15] S Greenland 1994 Alternative models for ordinal logistic regression Stat. in Med. 13 1665
[16] R Lall, M J Campbell, S J Walters, K Morgan 2002 A review of ordinal regression models applied on health-related quality of life assessments Stat. Meth. Med. Res. 11 49
[17] www.depkes.go.id
[18] H Robin, 2013 Models for Ordinal Response Data NC SAS Institute Inc.
[19] A W Capuano, J D Dawson, M R Ramirez, R S Wilson, L L Barnes and R W Field 2016 Modeling likert scale outcomes with trend-proportional odds with and without cluster data Eur. J. of Res. Met. for the Behav. and Soc. Sci. 12 33