

## ANALISIS FAKTOR-FAKTOR YANG MEMPENGARUHI KEMAMPUAN MEMBACA SISWA DENGAN METODE REGRESI PROBIT ORDINAL (STUDI KASUS SISWA SD SE-SULAWESI TENGGARA)

Waode Rasdiansastra<sup>1)</sup>, Gusti Ngurah Adhi Wibawa<sup>2)</sup> dan Bahriddin Abapihi<sup>3)</sup>  
Program Studi Matematika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Halu Oleo, Kendari, Indonesia  
E-mail: [srasdian22@gmail.com](mailto:srasdian22@gmail.com)

### ABSTRAK

Keterampilan membaca siswa memiliki peran penting dalam pengembangan diri secara berkelanjutan dan sekaligus sebagai fondasi dalam pendidikan siswa di lembaga formal. Tujuan penelitian ini adalah untuk mengetahui karakteristik pada variabel-variabel kemampuan membaca siswa dan estimasi model probit dari faktor-faktor yang mempengaruhi secara signifikan kemampuan membaca siswa dan siswi SD di Sulawesi Tenggara dengan menggunakan *software* SPSS. Penelitian ini menggunakan data kategori bertingkat atau biasa disebut data skala ordinal sebagai variabel dependen dan model yang digunakan adalah regresi probit ordinal. Hasil analisis menunjukkan bahwa variabel yang signifikan terhadap kemampuan membaca siswa dan siswi adalah jenis kelamin ( $X_1$ ), mengulang kelas ( $X_3$ ), belajar/membaca di rumah ( $X_4$ ), dan guru kelas pernah memperoleh pelatihan khusus tentang cara mengajar membaca ( $X_{11}$ ). Uji kesesuaian model menunjukkan bahwa nilai signifikansi  $> \alpha(0,05)$  artinya model regresi probit ordinal yang terbentuk telah sesuai atau tidak terdapat perbedaan antara hasil pengamatan dengan hasil prediksi model sebagai berikut :

$$\begin{aligned}P(Y = 1) &= \Phi[-0,746 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] \\P(Y = 2) &= \Phi[1,005 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] \\&\quad - \Phi[-0,746 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] \\P(Y = 3) &= 1 - \Phi[1,005 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})]\end{aligned}$$

**Kata Kunci:** Regresi Probit Ordinal, Data Kategorik, Kemampuan Membaca.

### ABSTRACT

*Student reading skills have a key role in sustainable and yet as a foundation in student education at the formal institution. The purpose of this study is to know the characteristics in the variables and the estimate of the probit model of the factors affecting students reading ability using the SPSS software. This study uses stratified category data or so-called ordinal scale as the dependent variable and the method is ordinal probit regression. Analysis results suggest that the significant variable in students reading ability are the gender of the students ( $X_1$ ), whether to display the class ( $X_3$ ), learn to read at home ( $X_4$ ), and the class teacher has received special training on how to teach reading ( $X_{11}$ ). Model compatibility test shows that significance  $> \alpha(0,05)$  it means that the models that are formed fit or there is no difference between the observations and the models predictions as follows :*

$$\begin{aligned}P(Y = 1) &= \Phi[-0,746 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] \\P(Y = 2) &= \Phi[1,005 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] \\&\quad - \Phi[-0,746 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] \\P(Y = 3) &= 1 - \Phi[1,005 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})]\end{aligned}$$

**Keywords:** Ordinal Probit Regression, Categorical Data, Reading Ability.

### 1. Pendahuluan

Pendidikan pada dasarnya adalah usaha sadar dari setiap individu untuk menumbuh-kembangkan potensi sumber daya dengan cara mendorong dan memfasilitasi kegiatan belajar-mengajar. Membaca adalah aktivitas pencarian informasi melalui lambang-lambang tertulis. Tampubolon (1987:5) berpendapat bahwa membaca merupakan suatu bagian atau komunikasi tulisan. Maksudnya, simbol-simbol tulisan atau huruf diubah menjadi simbol-simbol bunyi bahasa. Melalui pengajaran membaca, pendidik mengarahkan peserta didiknya agar mampu dalam memahami isi bacaan. Keterampilan membaca

memiliki peran penting dalam pengembangan diri secara berkelanjutan dan sekaligus sebagai fondasi dalam pendidikan siswa di lembaga formal. Anak pada usia sekolah dasar yang tidak segera memiliki kemampuan membaca yang baik akan kesulitan dalam mempelajari beberapa bidang studi pada tingkat setelahnya.

Peserta didik diharapkan memiliki dasar atau faktor-faktor yang mendukung sebagai langkah awal meningkatkan kemampuan membaca. *Early Grade Reading Assessment* (EGRA) yang dilakukan oleh ACDP (*Analytical and Capacity Development Partnership*) Indonesia yang bertujuan untuk

mengukur secara sistematis seberapa baik siswa di kelas-kelas awal sekolah dasar memiliki keterampilan membaca. Penilaian ini dilakukan pada tahun 2014 terhadap 4.812 siswa kelas 2, menunjukkan bahwa tidak sampai setengah dari jumlah tersebut mahir membaca dan paham apa yang mereka baca. Hanya 26% siswa dapat menjawab 3 dari 5 pertanyaan dengan benar dan 5,8% siswa tidak dapat membaca sama sekali. Juga terungkap bahwa anak-anak di Jawa dan Bali dapat membaca dan paham atas apa yang mereka baca secara lebih baik dibandingkan anak-anak di daerah lain di Indonesia. Anak-anak dari daerah timur Indonesia salah satunya Provinsi Sulawesi Tenggara membaca dengan tingkat pemahaman paling rendah. Melihat dari data tersebut maka perlu dilakukan analisis faktor-faktor yang mempengaruhi kemampuan baca peserta didik sebagai salah satu langkah awal untuk meningkatkan mutu pendidikan di Indonesia.

Analisis regresi adalah salah satu metode analisis yang kerap digunakan dalam kasus variabel yang dipengaruhi oleh variabel lain. Menurut Hosmer (1989) analisis regresi digunakan untuk mempelajari hubungan antara dua variabel atau lebih, dengan maksud bahwa dari hubungan tersebut dapat diperkirakan besarnya dampak kuantitatif yang terjadi dari perubahan suatu kejadian terhadap kejadian lainnya. Variabel-variabel tersebut terdiri atas variabel bebas dan variabel tak bebas yang selanjutnya hubungan antara variabel ini dicirikan melalui model matematik yang disebut model regresi.

Data kemampuan membaca siswa bersifat kuantitatif dan dikategorikan dalam beberapa tingkat atau biasa disebut skala ordinal. Oleh karena itu metode yang digunakan adalah pengembangan dari metode analisis regresi itu sendiri yang dikembangkan khusus untuk menangani data kategorikal. Salah satu metode yang dapat digunakan adalah regresi probit ordinal. Regresi probit ordinal adalah suatu model regresi yang digunakan untuk menjelaskan hubungan antara variabel respon yang merupakan variabel kategorik bertingkat atau ordinal dan variabel prediktornya bisa merupakan variabel kontinu atau variabel diskrit berskala nominal atau ordinal.

Pendugaan parameter dalam persamaan regresi probit ordinal dapat dilakukan dengan menggunakan metode *Maximum Likelihood* (MLE) dan metode *Newton-Raphson*. Metode *Maximum Likelihood* mengestimasi parameter  $\beta$  dengan memaksimalkan fungsi likelihood dengan syarat data mengikuti distribusi tertentu. Karena model probit non linear sehingga tidak dapat dihitung secara analitik, untuk memperoleh penaksir  $\beta$  maka digunakan metode numerik dengan iterasi *Newton-Raphson*.

Paper ini ditulis dengan susunan sebagai berikut, pada bagian 2 membahas tentang kajian pustaka, bagian 3 membahas tentang hasil dan pembahasan dan bagian 4 penutup.

## 2. Kajian Pustaka

### 2.1 Skala Pengukuran

Skala pengukuran adalah suatu skala yang digunakan untuk mengklasifikasikan variabel yang akan diukur dalam suatu observasi agar tidak terjadi kesalahan dalam menentukan analisis data dan langkah-langkah selanjutnya (Danapriatna dan Setiawan, 2005:6). Macam-macam skala pengukuran adalah sebagai berikut :

#### a. Skala Nominal

Skala nominal merupakan skala pengukuran yang menggolongkan objek-objek atau kejadian ke dalam berbagai kelompok atau kategori, untuk menunjukkan kesamaan atau perbedaan ciri-ciri objek, tetapi tidak bisa diurutkan mana yang lebih tinggi atau yang lebih rendah. Tes statistik yang digunakan adalah statistik nonparametrik.

#### b. Skala Ordinal

Skala ordinal merupakan skala pengukuran pada objek-objek ke dalam kategori-kategori tertentu. Angka atau huruf yang diberikan mengandung tingkatan sehingga dari kelompok yang terbentuk dapat dibuat peringkat yang menyatakan hubungan. Tes statistik yang digunakan adalah statistik nonparametrik.

#### c. Skala Interval

Skala interval merupakan skala pengukuran yang memberikan ciri angka kepada kelompok objek yang mempunyai skala nominal dan ordinal, ditambah dengan jarak yang sama pada urutan objeknya. Tes statistik yang digunakan adalah statistik parametrik.

#### d. Skala Rasio

Skala rasio merupakan skala pengukuran yang mempunyai semua ciri angka sama dengan skala interval ditambah dengan satu sifat lain yaitu memberikan keterangan tentang nilai absolut dari objek yang diukur. Tes statistik yang digunakan adalah statistik parametrik.

### 2.2 Model Linear Terampat (*Generalized Linear Models/GLM*)

GLM merupakan suatu metode yang dapat digunakan untuk memecahkan masalah peubah respon tidak lagi kontinu melainkan kategorik, dengan menggunakan fungsi penghubung (*link function*) tertentu sehingga diperoleh suatu model yang mampu menganalisa hubungan antara peubah respon kategorik dengan satu atau beberapa peubah penjelas. Bila  $Y$  peubah respon tidak lagi mengikuti sebaran normal namun binomial atau Poisson (asalkan masih dalam keluarga eksponensial) dan

ragam  $Y$  merupakan fungsi dari nilai tengahnya sehingga dapat dipastikan bahwa ragam tidak homogen maka digunakanlah suatu model yang disebut model linear terampat (*Generalized Linear Models / GLM*). GLM merupakan pengembangan dari model linear klasik dimana peubah respon  $Y$  merupakan suatu komponen yang bebas dengan nilai tengah  $\mu$ . Ada tiga komponen utama dalam GLM (McCullagh dan Nelder, 1989):

1. Komponen acak, yaitu komponen dari  $Y$  yang bebas dan fungsi sebaran peluang  $Y$  termasuk dalam keluarga sebaran eksponensial dengan  $E(Y) = \mu$ .
2. Komponen sistematik, yaitu  $x_1, x_2, \dots, x_p$  yang menghasilkan penduga linear  $\eta$  dimana  $\eta = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$ .
3. Fungsi penghubung (*link function*)  $g(\cdot)$ , yang menggambarkan hubungan antara penduga linear  $\eta$  dengan nilai tengah  $\mu$ . Hubungan ini dapat ditulis dengan  $\eta = g(\mu)$ .

Dalam model linear klasik, komponen (1) menyebar normal dan komponen (3) merupakan fungsi identitas. Sedangkan dalam GLM, komponen (1) mungkin berasal dari salah satu anggota keluarga sebaran eksponensial lainnya dan komponen (3) merupakan fungsi monoton lainnya. Dengan demikian GLM dapat dimodelkan dengan :

$$g(E(Y_i|x_i)) = g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} = \eta(x_i)$$

Fungsi kanonik *link* untuk distribusi binomial adalah  $\eta = \ln\left(\frac{p}{1-p}\right)$  atau sering disebut dengan fungsi logistik atau fungsi logit. Hubungan antara  $p$  dan  $\eta$  adalah non linear, yaitu menyerupai kurva S. Tiga fungsi *link* yang lain, yang dapat digunakan untuk menganalisis variabel respon berdistribusi binomial. Asumsi yang harus dipenuhi untuk fungsi *link* yang variabel respon berdistribusi binomial, yaitu nilai  $p$  berada ( $0 \leq p \leq 1$ ) dimana  $\eta$  bernilai ( $-\infty < \eta < \infty$ ). Beberapa fungsi *link* yang memenuhi asumsi tersebut adalah sebagai berikut:

1. Fungsi logistik atau logit,  $\eta = \ln\left(\frac{p}{1-p}\right)$  dan probabilitas dari logit,  $p = \frac{\exp\{\eta\}}{1 + \exp\{\eta\}}$ .
2. Fungsi invers kumulatif normal standar atau probit,  $\eta = \Phi^{-1}(p)$  dan probabilitas dari probit,  $p = \Phi(\eta)$ .
3. Fungsi *complementary log-log* atau fungsi *clog-log*,  $\eta = \log\left(-\log(1-p)\right)$ . Probabilitas dari *complementary log-log*,  $p = 1 - \frac{1}{\exp\{\exp(\eta)\}}$ .

### 2.3 Model Regresi Probit

Model regresi probit diperkenalkan oleh Chester Bliss pada tahun 1935. Model probit merupakan model non linear yang digunakan untuk menganalisis hubungan antara variabel dependen

(variabel respon) yang biasa disimbolkan  $Y$  dengan skala pengukuran dikotomis (biner), dan variabel independen (variabel prediktor) biasa disimbolkan  $X$  yang skala pengukuran bersifat dikotomis, polikotomis atau kontinu. Regresi probit merupakan singkatan dari "*Probability Unit*", sehingga regresi probit merupakan sebuah metode regresi yang berkaitan dengan unit-unit probabilitas. Regresi probit juga dikenal sebagai model Normit yang merupakan singkatan dari "*Normal Probability Unit*" karena berdasarkan fungsi sebaran peluang normal kumulatif baku. Estimasi parameter menggunakan metode *Maximum Likelihood Estimation* (MLE) dan menggunakan metode *Newton-Raphson*. Model regresi probit dapat ditulis seperti persamaan sebagai berikut :

$$\Phi^{-1}(\pi_i) = \mathbf{x}_i^T \boldsymbol{\beta} \text{ untuk } i=1,2,\dots,n$$

dengan  $\mathbf{x}_i^T = [1 \ x_1 \ x_2 \ \dots \ x_k]$  adalah vektor variabel independen pada pengamatan ke- $i$ ,  $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k]^T$  adalah vektor parameter,  $\Phi^{-1}(\pi_i)$  adalah probit *link function* dan  $\pi_i = P(Y_i = y)$  adalah probabilitas variabel respon pada pengamatan ke- $i$  mempunyai kategori  $y$ .

### 2.4 Regresi Probit Ordinal

Regresi probit ordinal adalah salah satu metode yang digunakan untuk menjelaskan hubungan antara variabel respon yang merupakan variabel kategorik bertingkat atau ordinal dengan variabel prediktor yang berupa variabel diskrit, kontinyu maupun campuran antara keduanya. Menurut Greene (2000), pemodelan regresi probit ordinal diawali dengan memperhatikan model sebagai berikut :

$$Y^* = \beta_0 + \boldsymbol{\beta}^T \mathbf{x} + \varepsilon$$

dimana :

$Y^*$ : variabel respon yang merupakan variabel kontinyu

$\beta_0$ : parameter intersep yang tidak diketahui

$\boldsymbol{\beta}$ : vektor parameter koefisien dengan

$$\boldsymbol{\beta} = [\beta_1 \ \beta_2 \ \dots \ \beta_p]^T$$

$\mathbf{x}$ : vektor variabel bebas dengan

$$\mathbf{x} = [X_1 \ X_2 \ \dots \ X_p]^T$$

$\varepsilon$ : error yang diasumsikan berdistribusi  $N(0, \sigma^2)$

Selanjutnya dilakukan transformasi ke dalam bentuk  $Z = \frac{Y^* - (\beta_0 + \boldsymbol{\beta}^T \mathbf{x})}{\sigma}$ , dimana  $Z \sim N(0,1)$ . Fungsi kepadatan peluang variabel  $Y^*$  adalah sebagai berikut

$$f(Y^*) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{Y^* - \boldsymbol{\beta}^T \mathbf{x}}{\sigma}\right)^2\right)$$

$Y^*$  berdistribusi normal dengan rata-rata  $\boldsymbol{\beta}^T \mathbf{x}$  dan varians  $\sigma^2$  atau dapat ditulis  $Y^* \sim N(\boldsymbol{\beta}^T \mathbf{x}, \sigma^2)$ . Selanjutnya dilakukan pengkategorian terhadap  $Y^*$  secara ordinal yaitu untuk  $Y^* \leq \gamma_1$  dikategorikan dengan  $Y = 1$ , untuk  $\gamma_1 < Y^* \leq \gamma_2$  dikategorikan dengan  $Y = 2$ , untuk  $\gamma_{i-1} < Y^* \leq \gamma_i$  dikategorikan dengan  $Y = i$ , untuk  $Y^* > \gamma_k$  dikategorikan dengan

$Y = k$ . Berikut ini adalah grafik fungsi kepadatan peluang dari  $Y^*$  dengan  $\gamma_1, \gamma_2, \dots, \gamma_{k-1}$  adalah batasan (*threshold*) yang membagi  $Y^*$  menjadi  $k$  kategori. Sehingga model regresi probit ordinal adalah sebagai berikut :

$$P(Y = 1) = \Phi(\gamma_1 - \beta^T X)$$

$$P(Y = 2) = \Phi(\gamma_2 - \beta^T X) - \Phi(\gamma_1 - \beta^T X)$$

$$\vdots$$

$$P(Y = k) = 1 - \Phi(\gamma_{c-1} - \beta^T X)$$

dengan  $Y = 1$  untuk kategori terendah,  $Y = k$  untuk kategori tertinggi dan  $\Phi$  yaitu fungsi distribusi kumulatif distribusi normal standar.

### 2.5 Efek Marginal

Efek marginal digunakan untuk menginterpretasikan model regresi probit ordinal dan menyatakan besarnya pengaruh setiap variabel prediktor yang signifikan terhadap peluang tiap kategori pada variabel respon (Greene, 2002). Berikut ini adalah formula dari efek marginal :

$$\frac{\partial P(Y = 1|x)}{\partial x} = (-\beta)\phi(\gamma_1 - \beta^T x)$$

$$\frac{\partial P(Y = 2|x)}{\partial x} = \beta[\phi(\gamma_1 - \beta^T x) - \phi(\gamma_2 - \beta^T x)]$$

$$\vdots$$

$$\frac{\partial P(Y = k|x)}{\partial x} = \beta[\phi(\gamma_{c-1} - \beta^T x)]$$

dengan  $\phi$  adalah fungsi kepadatan peluang dari distribusi normal standar.

### 2.6 Estimasi Parameter

Pendugaan parameter dalam persamaan regresi probit ordinal dapat dilakukan dengan menggunakan metode *Maximum Likelihood* (MLE) dan metode *Newton-Raphson*. Metode *Maximum Likelihood* mengestimasi parameter  $\beta$  dengan memaksimalkan fungsi likelihood dengan syarat data mengikuti distribusi tertentu. Berikut merupakan persamaan dari fungsi likelihood :

$$L(\beta) = \prod_{i=1}^n [p_1(x_i)]^{y_{1i}} [p_2(x_i)]^{y_{2i}} \dots [p_c(x_i)]^{y_{ci}}$$

### 2.7 Pengujian Parameter

Model yang terbentuk perlu dilakukan pengujian parameter untuk mengetahui apakah variabel-variabel prediktor yang digunakan berpengaruh signifikan terhadap variabel respon. Pengujian ini dilakukan sebanyak dua tahap, yaitu uji serentak dan uji parsial (Hosmer & Lemeshow, 2000).

#### 1) Uji Serentak

Uji serentak adalah pengujian yang dilakukan untuk mengetahui signifikansi koefisien  $\beta$  terhadap variabel respon secara serentak atau keseluruhan. Uji yang digunakan untuk pengujian parameter secara simultan dalam penelitian ini yaitu uji  $G$  atau *Likelihood Ratio Test* (LRT) dengan hipotesis sebagai berikut :

$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$  (secara bersama-sama variabel independen tidak mempengaruhi model).

$H_1 : \text{minimal ada satu } \beta_i \neq 0, \text{ dimana } i = 1, 2, \dots, p$  dimanap adalah jumlah variabel prediktor.

Statistik Uji  $G^2$  atau LRT (*Likelihood Ratio Test*) :

$$G^2 = -2 \ln \left( \frac{\binom{n_0}{n}^{n_0} \binom{n_1}{n}^{n_1} \binom{n_2}{n}^{n_2}}{\prod_{i=1}^n [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]} \right)$$

Statistik uji  $G^2$  mengikuti distribusi  $\chi^2$  maka pengujian dilakukan dengan membandingkan antara nilai statistik uji  $G^2$  dan nilai tabel  $\chi^2$  dengan derajat bebas  $v$  (banyaknya parameter) pada taraf signifikan  $\alpha$ .  $H_0$  ditolak apabila nilai  $G^2_{hit} > \chi^2_{(v,\alpha)}$  atau  $Pvalue < \alpha$ .

#### 2) Uji Parsial

Uji parsial dilakukan jika pada pengujian serentak didapatkan hasil tolak  $H_0$ . Pengujian parameter secara parsial dilakukan untuk menguji signifikansi dari masing-masing variabel prediktor dalam model. Uji yang digunakan untuk pengujian parameter secara parsial dalam penelitian ini yaitu Uji Wald dengan hipotesis sebagai berikut :

$H_0 : \beta_i = 0$  (parameter tidak signifikan atau variabel bebas tidak memiliki hubungan yang kuat dengan variabel respon).

$H_1 : \beta_i \neq 0$  dimanai = 1, 2, ... ,  $p$  (parameter signifikan atau variabel bebas memiliki hubungan yang kuat dengan variabel respon).

Statistik uji Wald mengikuti distribusi normal standar menggunakan mean sebesar 0 serta variansi sebesar 1, maka pengujian dilakukan dengan membandingkan antara statistik uji Wald dengan distribusi normal standar pada taraf signifikan  $\alpha$ .

$$W = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)}, SE(\hat{\beta}_i) = [var(\hat{\beta}_i)]^{1/2}$$

dengan :

$\hat{\beta}_i$  : penaksir dari parameter  $\beta_i$

$SE(\hat{\beta}_i)$  : standar error dari  $\hat{\beta}_i$

$H_0$  ditolak jika nilai  $|W| > Z_{\alpha/2}$  atau  $Pvalue < \alpha$ .

### 2.8 Uji Kesesuaian Model

Uji kesesuaian model dilakukan untuk menguji apakah model yang dihasilkan sudah layak atau tidak terdapat perbedaan antara hasil pengamatan dengan hasil prediksi model. Hipotesis yang digunakan adalah sebagai berikut

$H_0$ : Model sesuai (tidak ada perbedaan antara hasil observasi dengan kemungkinan hasil prediksi model)

$H_1$ : Model tidak sesuai (ada perbedaan antara hasil observasi dengan kemungkinan hasil prediksi model).

Statistik uji *Pearson* :

$$\chi^2_{Pearson} = \sum_{i=1}^g \sum_{j=1}^k \frac{(Y_{ij} - n_{ij} \hat{\pi}_j(X_i))^2}{n_{ij} \hat{\pi}_j(X_i)}$$

dengan

$\hat{\pi}_j(X_i)$  =peluang terjadinya nilai peubah respon Y ke-j pada peubah prediktor X ke-i

$Y_{ij}$  =hasil pengamatan peubah respon Y ke-j pada peubah prediktor X ke-i

$i = 1, \dots, p$

$j = 1, \dots, k$

$m$  =banyaknya parameter yang diduga dalam model

$g$  =banyaknya kombinasi peubah prediktor

$k$  =banyaknya kategori peubah respon Y

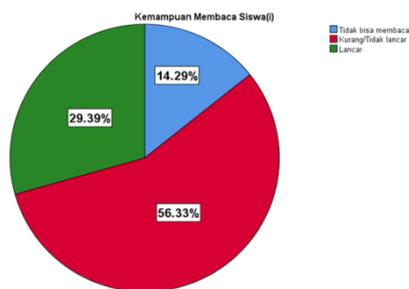
$p^{nr}$  =perkalian antara banyaknya kategori peubah respon dikurangi satu dengan banyaknya parameter yang diduga dalam model

Statistik *Pearson* mengikuti sebaran khi-kuadrat dengan derajat bebas  $g(k - 1) - p^{nr}$ . Keputusan tolak  $H_0$  apabila  $\chi^2_{pearson} > \chi^2_{tabel}$  atau  $Pvalue < \alpha$  yang menyatakan bahwa model tidak sesuai (ada perbedaan antara hasil observasi dengan kemungkinan hasil prediksi model).

### 3. Hasil dan Pembahasan

#### 3.1 Statistika Deskriptif

Statistika deskriptif adalah metode yang berkaitan dengan pengumpulan data dan penyajian suatu gugus data sehingga menaksir kualitas data berupa jenis variabel, ringkasan statistik dan representasinya dalam bentuk tabel atau diagram hingga memperoleh informasi yang lebih mudah dipahami. Analisis deskriptif dilakukan untuk melihat distribusi frekuensi dan persentase dari tiap variabel penelitian sehingga diketahui gambaran karakteristik kemampuan membaca siswa SD di Provinsi Sulawesi Tenggara. Sebagai gambaran awal berikut ditampilkan persentase kategori kemampuan membaca siswa(i) SD di Sulawesi Tenggara pada tahun 2019.



Gambar 3.1. Persentase Kategori Kemampuan Membaca

Gambar 3.1 menunjukkan bahwa sebanyak 56,33% siswa(i) di Provinsi Sulawesi Tenggara yang termasuk dalam kategori kurang/tidak lancar membaca, 29,39% siswa(i) yang termasuk dalam kategori lancar membaca, dan 14,29% siswa(i) yang termasuk dalam kategori tidak bisa membaca.

#### 3.2 Uji Independensi antar Variabel

Hubungan antar variabel bebas yaitu  $X_1$  (jenis kelamin siswa(i)),  $X_2$  (tamat PAUD/TK),  $X_3$  (mengulang kelas),  $X_4$  (belajar/membaca di rumah),  $X_5$  (Ayah membantu belajar/membaca di rumah),  $X_6$  (Ibu membantu belajar/membaca di rumah),  $X_7$  (akreditasi sekolah),  $X_8$  (lokasi sekolah),  $X_9$  (ijazah tertinggi guru kelas),  $X_{10}$  (guru kelas berlatar belakang pendidik),  $X_{11}$  (guru kelas pernah memperoleh pelatihan khusus tentang cara mengajar membaca),  $X_{12}$  (guru kelas mengetahui tentang GLS/Gerakan Literasi Sekolah) dengan variabel tak bebas yaitu Y (kemampuan membaca siswa(i)).

##### 3.2.1 Uji *Chi-Square*

Dalam penelitian ini menggunakan 245 sampel dengan skala data nominal dan ordinal yang dikategorikan dan memiliki tabel kontingensi 2x2 dan 2x3 sehingga cocok untuk menggunakan *Pearson Chi square*. Uji ini bertujuan untuk mengetahui variabel-variabel penjelas yang signifikan berhubungan dengan variabel kemampuan membaca siswa(i). Adapun hipotesisnya sebagai berikut :

$H_0$  : Tidak ada hubungan antara variabel bebas dan variabel tak bebas

$H_1$  : Ada hubungan antara variabel bebas dan variabel tak bebas

Tabel 3.1. Hasil uji independensi antar variabel

Keterangan	Variabel	Sig. ( <i>Chi-square</i> )	Keputusan
Jenis kelamin siswa(i)	$X_1$	0,011	$H_0$ ditolak
Tamat PAUD/TK	$X_2$	0,156	$H_0$ diterima
Mengulang kelas	$X_3$	0,004	$H_0$ ditolak
Belajar/membaca di rumah	$X_4$	0,029	$H_0$ ditolak
Ayah membantu belajar/membaca di rumah	$X_5$	0,988	$H_0$ diterima
Ibu membantu belajar/membaca di rumah	$X_6$	0,913	$H_0$ diterima
Lokasi sekolah	$X_8$	0,310	$H_0$ diterima
Guru kelas berlatar belakang pendidik	$X_{10}$	0,243	$H_0$ diterima
Guru kelas pernah memperoleh pelatihan khusus tentang cara mengajar membaca	$X_{11}$	0,009	$H_0$ ditolak
Guru kelas mengetahui tentang GLS (Gerakan Literasi Sekolah)	$X_{12}$	0,392	$H_0$ diterima

Tabel 3.1 memperlihatkan hasil analisis independensi *Chi-square* dengan kriteria pengujian

menolak  $H_0$  apabila  $\chi^2_{hitung} > \chi^2_{tabel}$  atau signifikansi P-value  $< 0,05$ . Karena variabel bebas yaitu  $X_1, X_3, X_4, X_{11}$  memiliki nilai P-value  $< 0,05$  maka keputusan yang diambil pada taraf nyata 5% bahwa  $H_0$  ditolak. Sehingga kesimpulannya terdapat hubungan yang signifikan dengan variabel tak bebas yaitu kemampuan membaca siswa(i).

### 3.2.2 Uji Korelasi Gamma

Uji korelasi gamma digunakan untuk mengetahui hubungan antara dua variabel dimana kedua variabel tersebut mempunyai skala data ordinal. Adapun hipotesisnya sebagai berikut :

$H_0$  : Tidak ada hubungan antara variabel independen dengan kemampuan membaca.

$H_1$  : Ada hubungan antara variabel independen dengan kemampuan membaca.

**Tabel 3.2.** Hasil uji korelasi gamma

Keterangan	Variabel	Sig. (Gamma)	Keputusan
Akreditasi sekolah	$X_7$	0,075	$H_0$ diterima
Ijazah tertinggi guru kelas	$X_9$	0,417	$H_0$ diterima

Tabel 3.2 memperlihatkan hasil analisis korelasi Gamma dengan kriteria pengujian menolak  $H_0$  apabila signifikansi P-value  $< 0,05$ . Karena variabel bebas yaitu  $X_7$  dan  $X_9$  memiliki nilai P-value  $> 0,05$  maka keputusan yang diambil pada taraf nyata 5% bahwa  $H_0$  diterima. Sehingga kesimpulannya kedua variabel tersebut tidak terdapat hubungan yang signifikan dengan variabel tak bebas yaitu kemampuan membaca siswa(i).

### 3.3 Analisis Regresi Probit Ordinal

#### 3.3.1 Estimasi Parameter

Model yang terbentuk dari variabel-variabel yang signifikan selanjutnya dilakukan estimasi parameter dengan metode maksimum *Likelihood* yang disajikan pada tabel di berikut:

**Tabel 3.3.** Estimasi regresi probit ordinal

Variabel	Estimasi	
Konstanta	Y = 1	-0,746
	Y = 2	1,005
Prediktor	$X_1$	0,405
	$X_3$	-1,084
	$X_4$	0,333
	$X_{11}$	0,407

Tabel 3.3 memperlihatkan estimasi parameter dari regresi probit ordinal dengan nilai dugaan parameter konstan  $\hat{\beta}_{01}$  (tidak bisa membaca) adalah -0,746 dan  $\hat{\beta}_{02}$  (kurang/tidak lancar) adalah 1,005.

Sedangkan untuk  $\hat{\beta}_1$  (jenis kelamin siswa(i)) adalah 0,405,  $\hat{\beta}_3$  (mengulang kelas) adalah -1,084,  $\hat{\beta}_4$  (belajar/membaca di rumah) adalah 0,333, dan  $\hat{\beta}_{11}$  (guru kelas pernah memperoleh pelatihan khusus tentang cara mengajar membaca) adalah 0,407.

### 3.3.2 Uji Signifikansi Parameter

#### a) Uji Serentak

Pengujian parameter secara serentak dilakukan untuk mengetahui apakah secara bersama-sama (simultan) parameter yang digunakan berpengaruh signifikan terhadap model.

**Tabel 3.4.** Hasil uji serentak regresi probit ordinal

Model	-2 Log Likelihood	Chi-Square	Sig.
Intercept Only	105,310	30,938	0,000
Final	74,372		

Pada Tabel 3.4 didapatkan nilai statistik uji *Likelihood Ratio Test* sebesar 30,938 dengan kriteria pengujian mengambil taraf nyata 5% dari tabel distribusi chi-square  $\chi^2_{(0,05;4)} = 9,488$  karena nilai statistik  $G > \chi^2_{(0,05;4)}$  maka kesimpulannya  $H_0$  ditolak, artinya minimal terdapat satu parameter yang signifikan terhadap model. Untuk mengetahui parameter manakah yang signifikan terhadap model maka dilanjutkan dengan uji parsial.

#### b) Uji Parsial

Pengujian parameter secara parsial atau individu dilakukan untuk mengetahui variabel mana saja yang berpengaruh secara signifikan terhadap kemampuan membaca siswa(i) SD di Sulawesi Tenggara. Statistik uji yang digunakan adalah statistik uji *Wald*.

**Tabel 3.5** Hasil uji parsial regresi probit ordinal

Variabel	Estimasi	SE	Wald	Sig.
Y (kemampuan membaca)				
1 (tidak bisa membaca)	-0,746	0,139	28,976	0,000
2 (kurang/tidak lancar)	1,005	0,143	49,113	0,000
$X_1$ (jenis kelamin)				
1 (perempuan)	0,405	0,149	7,397	0,007
$X_3$ (mengulang kelas)				
1 (Ya)	-1,084	0,370	8,591	0,003
$X_4$ (belajar/membaca di rumah)				
1 (Ya)	0,333	0,149	4,994	0,025
$X_{11}$ (guru kelas memperoleh pelatihan khusus cara mengajar membaca)				
1 (Ya)	0,407	0,166	5,981	0,014

Tabel 3.5 menunjukkan bahwa nilai mutlak statistik uji *Wald* pada variabel independen  $X_1$  (jenis kelamin) kategori 1 (perempuan),  $X_3$  (mengulang kelas) kategori 1 (Ya),  $X_4$  (belajar/membaca di rumah) kategori 1 (tidak ada), dan  $X_{11}$  (guru kelas pernah memperoleh pelatihan khusus tentang cara mengajar membaca) kategori 1 (Ya) lebih besar dari nilai tabel  $Z_{0,05/2} = 1,96$  atau dapat dilihat dari nilai sig. pada masing-masing variabel independen yang nilainya kurang dari  $\alpha$  (0,05) sehingga keputusan yang diambil adalah tolak  $H_0$  artinya variabel-variabel independen tersebut memiliki pengaruh signifikan terhadap variabel dependen secara individu/parsial atau dengan kata lain variabel-variabel independen tersebut berpengaruh signifikan terhadap model. Diperoleh nilai  $\beta_0$  yang selanjutnya digunakan untuk membangun persamaan model peluang sebagai berikut:

$$P(Y = 1) = \Phi[-0,746 - (C)]$$

$$P(Y = 2) = \Phi[1,005 - (C)] - \Phi[-0,746 - (C)]$$

$$P(Y = 3) = 1 - \Phi[1,005 - (C)]$$

dimana  $C$  merupakan fungsi probit dengan persamaan sebagai berikut:

$$C = 0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11}$$

### 3.3.3 Uji Kesesuaian Model

Uji kesesuaian model dilakukan untuk mengetahui apakah terdapat perbedaan yang signifikan antara hasil pengamatan dengan hasil prediksi model. Berikut adalah hipotesis untuk uji kesesuaian model.

$H_0$  : Tidak ada perbedaan antara hasil pengamatan dengan hasil prediksi model (model sesuai).

$H_1$  : Ada perbedaan antara hasil pengamatan dengan hasil prediksi model (model tidak sesuai).

**Tabel 3.6.** Hasil uji kesesuaian model

	Chi-Square	df	Sig.
Pearson	20,556	22	0,548
Deviance	22,434	22	0,434

Berdasarkan Tabel 3.6, hasil analisis yang dilakukan untuk uji kesesuaian model menunjukkan bahwa nilai Sig.  $> \alpha$  (0,05) dan nilai statistik uji *Goodness of Fit Test*  $= 20,556 < \chi^2_{(0,05;22)} = 33,924$ . Hal tersebut menunjukkan bahwa keputusan yang diambil adalah  $H_0$  diterima sehingga dapat disimpulkan bahwa model regresi probit ordinal yang terbentuk telah sesuai atau tidak terdapat perbedaan antara hasil pengamatan dengan hasil prediksi model (model sesuai).

### 3.3.4 Interpretasi Model Regresi Probit Ordinal

Interpretasi model dapat dilakukan berdasarkan efek marginal (*marginal effect*) yang nilainya menunjukkan seberapa besar pengaruh untuk masing-masing penambahan satu satuan/ prediksi probabilitas variabel independen terhadap probabilitas suatu faktor diklasifikasikan ke dalam masing-masing tingkat kemampuan membaca siswa(i) di Provinsi Sulawesi Tenggara. Berikut merupakan contoh persamaan untuk interpretasi efek marginal faktor jenis kelamin ( $X_1$ ) terhadap masing-masing kategori kemampuan membaca siswa(i) di Provinsi Sulawesi Tenggara.

$$\frac{\partial P(Y = 1)}{\partial X_1} = (0,405)\{\phi[-0,746 - (C)]$$

$$= (0,405)\phi[-0,746$$

$$- (0,405X_1 - 1,084X_3 + 0,333X_4$$

$$+ 0,407X_{11})]$$

$$= (0,405)\phi[-0,746$$

$$- (0,405(1) - 1,084(1)$$

$$+ 0,333(1) + 0,407(1))]$$

$$= (0,451)\phi[-0,807]$$

$$= (0,451)[0,2881398028]$$

$$= 0,1299510511 \approx 0,1299$$

Nilai *marginal effect* variabel  $X_1$  (jenis kelamin) adalah 0,1299. Hal ini menunjukkan bahwa variabel  $X_1$  (jenis kelamin) berkategori 1 atau jenis kelamin perempuan akan meningkatkan probabilitas responden untuk masuk kategori  $y = 1$  atau kategori tidak bisa membaca sebesar 0,1299 atau 12,99%.

## 4. Penutup

### 4.1 Kesimpulan

Berdasarkan hasil penelitian yang diuraikan, maka diperoleh kesimpulan sebagai berikut.

1. Karakteristik kemampuan membaca siswa(i) SD di Provinsi Sulawesi Tenggara termasuk dalam kategori kurang/tidak lancar membaca yaitu sebesar 56,33%, sebesar 14,29% termasuk dalam kategori tidak bisa membaca, dan sisanya sebesar 29,39% termasuk dalam kategori lancar membaca.
2. Pada penerapan data Berdasarkan analisis regresi probit ordinal diperoleh faktor-faktor yang berpengaruh terhadap kemampuan membaca siswa(i) SD di Sulawesi Tenggara adalah  $X_1$  (jenis kelamin),  $X_3$  (mengulang kelas),  $X_4$  (belajar/membaca di rumah), dan  $X_{11}$  (guru kelas pernah memperoleh pelatihan khusus tentang cara mengajar membaca). Sehingga didapatkan model probit ordinal sebagai berikut:

$$P(Y = 1) = \Phi[-0,746$$

$$- (0,405X_1 - 1,084X_3$$

$$+ 0,333X_4 + 0,407X_{11})]$$

$$P(Y = 2) = \Phi[1,005 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})] - \Phi[-0,746 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})]$$
$$P(Y = 3) = 1 - \Phi[1,005 - (0,405X_1 - 1,084X_3 + 0,333X_4 + 0,407X_{11})]$$

#### 4.2 Saran

Untuk penelitian selanjutnya diharapkan dikaji lagi dengan menggunakan metode lainnya yang sesuai dalam melakukan pemodelan dan dengan meninjau kembali faktor-faktor lain yang mungkin juga berpengaruh terhadap kemampuan membaca siswa(i) SD di Sulawesi Tenggara.

**Ucapan Terimakasih.** Penulis menyampaikan terimakasih yang setulus-tulusnya kepada dosen pembimbing atas segala perhatiannya sehingga tulisan ini dapat diselesaikan dengan baik. Ucapan terimakasih juga penulis sampaikan kepada seluruh pihak yang turut andil dalam penelitian ini baik langsung maupun tidak langsung

#### Daftar Pustaka

- Cahya, R.P. dan Vita, R., 2015, Pemodelan Logit, Probit dan Complementary Log-Log pada Studi Kasus Partisipasi Perempuan dalam Pembangunan Ekonomi di Kalimantan Selatan, *Jurnal Sains dan Seni ITS*, **4**(2), 2-5.
- Danapriatna, N., dan Setiawan, R. 2005. *Pengantar Statistika*. Yogyakarta: Graha Ilmu.
- Darnius, O. dan Tarigan, G. 2018. *Simulation method of model selection based on Mallows' Cp Criteria in linear regression*. International Conference on Science and Technology 2018: Journal of Physics: Conf. Series 1116 (2018) 022008.
- Goodman, L., & Kruskal, W. 1954. *Measures of association for cross classifications*. Journal American Statistical Association. 49. 732-764.
- Greene, W.H., 2000. *Econometrics Analysis*, 4th edition. Prentice Hall, New Jersey.
- Gujarati, D. (2004). *Basic Econometrics (4th Edition)*. New York: The McGraw-Hill.
- Hakim, A. 2000. *Statistika Induktif*. Yogyakarta: Ekonisia.
- Hocking, R. R. (1996). *Methods and Applications of Linear Models: Regression and Analysis of Variance*. New York: John Wiley & Sons, Inc.
- Hosmer, D. W. dan Lemeshow, S. (2000). *Applied Logistic Regression*. USA: John Wiley & Sons.
- Kismiantini. 2007. *Analisis Peubah Respon Biner*. *Jurnal Pythagoras*. **3**(1), 64-73.
- Lucia, R. M. (2018). *Perbandingan Uji Regresi Logistik Ordinal Model Logit dan Model Probit Terhadap Estimasi Prediksi Probabilitas Kejadian Kelahiran Prematur*(Skripsi). Fakultas Kesehatan Masyarakat, Universitas Sumatera Utara, Sumatera Utara.
- McCullagh, P. dan Nelder, J. A. 1989. *Generalized Linear Models 2nd Edition*. London: Chapman & Hall.
- Mendenhall, W. and Reimnuth, J. E. 1974. *Statistics for Management and Economics*. Duxbury Press, Massachusetts.
- Montgomery, D., Peck, E. Dan Vining, G. G. 1992. *Introduction to Linear Regression Analysis*. New York: Joh Willey & Sons, Inc.
- Nawari. 2010. *Analisis Regresi dengan Ms Excel 2007 dan SPSS 17*. PT.Elex Media Komputindo: Jakarta.
- Mohammad, Q. 2013. *Analisis Regresi Terapan: Teori, Contoh Kasus, dan Aplikasi dengan SPSS*. Yogyakarta: Andi.
- Sudjana. 1997. *Metoda Statistika*. Bandung: PT. Tarsito Bandung.
- Supranto. 2008. *Statistik Teori dan Aplikasi*. Jakarta: Erlangga.
- Walgito, B. (2010). *Pengantar psikologi umum*. Yogyakarta: Andi.

Diterima pada tanggal 26 Mei 2022.  
Terbit online pada tanggal 28 Juli 2022