

## Pemodelan Indeks Pembangunan Manusia (IPM) Menggunakan Analisis Regresi Probit (Studi Kasus: Indeks Pembangunan Manusia (IPM) di Pulau Kalimantan Tahun 2017)

### *Modelling Human Development Index (HDI) Using Probit Regression Analysis (Case Study: Human Development Index (HDI) in Kalimantan Island Year 2017)*

Santo Christyadi<sup>1</sup>, Andi M. Ade Satriya<sup>2</sup>, dan Rito Goejantoro<sup>3</sup>

<sup>1,2</sup>Laboratorium Statistika Ekonomi dan Bisnis FMIPA Universitas Mulawarman

<sup>3</sup>Laboratorium Statistika Komputasi FMIPA Universitas Mulawarman

<sup>1</sup>E-mail : santochristyadi@gmail.com

#### **Abstract**

Ordinal probit regression analysis is non-linear regression analysis that used to find affected independent variables for ordered categorical dependent variable and regression model in this analysis used Normal cumulative distribution function. Parameter estimation in this model used Maximum Likelihood Estimation (MLE) method. This model has been applied to Human Development Index (HDI) in Borneo Island in 2017 case study. HDI is the most important measurement in improving the human development quality in all cities/regencies in Indonesia. Some factors that affected to IPM, they are Life Expectancy ( $X_1$ ), School Expectancy ( $X_2$ ), Spending per Capita ( $X_3$ ), Average School Duration ( $X_4$ ), and Labour Force Participation Rate ( $X_5$ ). Based on research that was performed by researcher, resulted two factors affecting to HDI, those are Life Expetancy and Average School Duration. This model has classification accuracy of 89,29%, APER (Apparent Error Rate) value of 10,71%, and AIC (Akaike Information Criterion) value of 39,75; this model was very good because prediction value is almost approaching to observation value (actual value).

**Keywords:** AIC, APER, Human Development Index (HDI), Classification Accuracy, Ordinal Probit Regression

#### **Pendahuluan**

Sesungguhnya, statistika sangat diperlukan bukan saja dalam penelitian atau riset, tetapi juga perlu dalam bidang pengetahuan lainnya seperti: teknik, industri, ekonomi, astronomi, biologi, kedokteran, asuransi, pertanian, pemiagaan, bisnis, sosiologi, antropologi, pemerintahan, pendidikan, psikologi, meteorologi, geologi, farmasi, ekologi, pengetahuan alam, pengetahuan sosial dan lain sebagainya (Sudjana, 2005).

Analisis regresi adalah analisis data yang menggambarkan pengaruh antara variabel bebas dengan variabel tak-bebas. Kata "Regresi" pertama kali dikenalkan oleh Francis Galton. Analisis regresi adalah alat utama dalam ekonometrika. Galton menghubungkan antara tinggi badan orang tua dengan tinggi anak-anak sebagai gambaran/ilustrasi dalam analisis regresi (Gujarati, 1993).

Regresi probit adalah regresi yang menggunakan fungsi distribusi kumulatif (*Cumulative Distribution Function/CDF*) dari distribusi normal dan melibatkan variabel tak-bebas yang bersifat kualitatif, baik itu dikotomis maupun polikotomis. Regresi probit diperkenalkan oleh Chester Bliss pada Tahun 1934 untuk pertama kalinya. Untuk pengestimasi parameter-parameter yang terdapat dalam model regresi probit, digunakan metode *Maximum Likelihood Estimation* (MLE) (Agresti, 2007).

Indeks Pembangunan Manusia (IPM) atau *Human Development Index* (HDI) adalah salah

satu tolak ukur yang paling penting dalam menunjang keberhasilan pembangunan seluruh kota di Indonesia. IPM terbentuk oleh tiga dimensi dasar, yaitu umur panjang dan hidup sehat, pengetahuan, dan standar hidup layak. Ketiga dimensi tersebut akan memberi pengaruh yang sangat besar pada nilai IPM di daerah tertentu di Indonesia. Biasanya, nilai IPM akan bergantung pada berbagai permasalahan yang ada di masyarakat, seperti pengangguran, kemiskinan, buta huruf, tingkat ketahanan pangan, interaksi sosial, pemerintahan dan permasalahan lainnya (BPS, 2018).

IPM di Kalimantan Timur pada Tahun 2013 memiliki nilai IPM sebesar 73,21. Sampai pada Tahun 2017, IPM di Kalimantan Timur meningkat dengan nilai 75,12 (BPS Provinsi Kalimantan Timur, 2018). Pada tahun 2016, IPM di Kalimantan Barat memiliki nilai sebesar 65,88 dan meningkat menjadi 66,26 pada tahun 2017. Ini menempatkan Kalimantan Barat ke posisi 30 dari 34 provinsi se-Indonesia (BPS Provinsi Kalimantan Barat, 2018). IPM di Kalimantan Tengah tahun 2017 secara nasional mencapai angka 69,79, meningkat dibandingkan tahun sebelumnya, 2016 yang mencapai angka 69,13. IPM di Kalimantan Tengah tergolong IPM sedang ( $60 \leq \text{IPM} < 70$ ) (BPS Kalimantan Tengah, 2018). IPM di Kalimantan Selatan pada tahun 2017 telah meningkat menjadi 69,65 yang sebelumnya pada tahun 2016 mencapai angka 69,05 dan IPM di Kalimantan termasuk dalam kategori IPM sedang

(BPS Kalimantan Selatan, 2018). Kalimantan Utara pada tahun 2017 memiliki nilai IPM sebesar 69,84 dan ini meningkat daripada tahun sebelumnya, 2016 yang mencapai 69,20. Pada tahun 2013, provinsi ini memiliki nilai IPM sebesar 67,99. Kalimantan Utara termasuk dalam kategori IPM sedang (BPS Kalimantan Utara, 2018).

Berdasarkan penelitian sebelumnya, yaitu mengenai jumul yang berjudul “Analisis Faktor-Faktor Yang Mempengaruhi Indeks Pembangunan Manusia (IPM) Menggunakan Metode Regresi Logistik Ordinal dan Regresi Probit Ordinal” dengan studi kasusnya adalah IPM di seluruh kota/kabupaten di Provinsi Jawa Timur Tahun 2014 yang ditulis oleh Nurmalasari dkk. (2017), diperoleh kesimpulan bahwa model regresi logistik ordinal sama baiknya dengan model regresi probit ordinal disebabkan karena nilai ketepatan klasifikasi serta nilai AIC yang sama. Oleh karena itu, penulis tertarik melakukan penelitian yang berjudul “Pemodelan Indeks Pembangunan Manusia (IPM) Menggunakan Analisis Regresi Probit (Studi Kasus: Indeks Pembangunan Manusia (IPM) di Pulau Kalimantan Tahun 2017)”. Dari penelitian ini, akan dicari faktor-faktor mana saja yang sangat berpengaruh terhadap IPM di seluruh kota/kabupaten di seluruh Provinsi di Pulau Kalimantan dan model regresi probit yang akan terbentuk serta interpretasi dari model tersebut.

**Analisis Regresi**

Istilah regresi diperkenalkan oleh Francis Galton. Dalam satu makalah yang terkenal, Galton menemukan bahwa meskipun ada kecenderungan bagi orang tua yang tinggi mempunyai anak-anak yang tinggi dan bagi orang tua yang pendek mempunyai anak-anak yang pendek, distribusi tinggi suatu populasi tidak berubah secara menyolok (besar) dari generasi ke generasi. Dengan kata-kata Galton, ini adalah “kemunduran ke arah sedang”. Tetapi penafsiran modern regresi sungguh berbeda, secara umum, analisis regresi berkenaan dengan studi ketergantungan satu variabel, variabel tak bebas, pada satu atau lebih variabel lain, variabel yang menjelaskan (*explanatory variables*), dengan maksud menaksir dan atau meramalkan nilai rata-rata hitung (*mean*) atau rata-rata (populasi) variabel tak bebas, dipandang dari segi nilai yang diketahui atau tetap (dalam pengambilan sampel berulang) variabel yang menjelaskan (yang belakangan) (Gujarati, 1993).

**Regresi Probit**

Model probit pertama kali diperkenalkan oleh Chester Bliss pada tahun 1934 dalam penerapannya di bidang teknologi. Model probit merupakan model non-linier yang menggunakan

biner (variabel boneka) sebagai variabel tak-bebasnya dan mengandalkan faktor galat  $\varepsilon_i$  yang berdistribusi normal (Agregsti, 2007).

Misalkan  $Y_i^*$  adalah variabel respon yang tak teramati. Variabel tak bebas yang dipengaruhi oleh variabel bebas memiliki persamaan yang ditulis sebagai berikut:

$$Y_i^* = \beta_0 + \beta_1 X_{li} + \dots + \beta_p X_{pi} + \varepsilon_i \quad (1)$$

Model probit dinyatakan dalam persamaan sebagai berikut:

$$P_i = \phi(Y_i^*) = \phi(\beta_0 + \beta_1 X_{li} + \dots + \beta_p X_{pi}) \quad (2)$$

Dari persamaan (2), maka fungsi  $\phi$  dapat diinverskan dalam bentuk linier. Hasil inversnya dapat ditulis sebagai berikut:

$$\phi^{-1}(P_i) = \beta_0 + \beta_1 X_{li} + \dots + \beta_p X_{pi} \quad (3)$$

dengan:

- $P_i$  : peluang “sukses” pada model ke- $i$
- $\phi(Y_i)$  : fungsi Distribusi Kumulatif (CDF) Normal
- $\beta_k$  : koefisien parameter model regresi dengan  $k = 1, 2, 3, \dots, p$
- $X_{ki}$  : variabel bebas dengan  $k = 1, 2, 3, \dots, p$  dan model ke- $i$
- $\varepsilon_i$  : galat pada model ke- $i$ . (Yuliza, 2017)

**Regresi Probit Ordinal**

Regresi probit ordinal adalah model regresi yang digunakan untuk menjelaskan hubungan antara variabel tak bebas yang merupakan variabel diskrit yang berskala ordinal dan variabel bebas yang mengandung variabel kontinu, diskrit, atau keduanya. Pemodelan regresi probit ordinal diawali dengan memperhatikan model berikut:

$$Y^* = \mathbf{X}^T \boldsymbol{\beta} + \varepsilon \quad (4)$$

Keterangan:

- $Y^*$  : variabel tak bebas yang merupakan variabel diskrit
- $\boldsymbol{\beta}$  : vektor parameter koefisien,  $\boldsymbol{\beta} = [\beta_0 \ \beta_1 \ \dots \ \beta_p]^T$ , ukuran  $p \times 1$
- $\mathbf{X}$  : vektor variabel bebas,  $\mathbf{X} = [1 \ X_1 \ \dots \ X_p]^T$ , ukuran  $1 \times p$
- $\varepsilon$  : galat yang berdistribusi  $N(0, \sigma^2)$

Berdasarkan persamaan (4), dilakukan transformasi ke nilai Z, di mana rumus nilai Z ditulis sebagai berikut:

$$Z = \frac{Y^* - \mathbf{X}^T \boldsymbol{\beta}}{\sigma} \tag{5}$$

di mana  $\sigma$  adalah nilai simpangan baku populasi. Nilai  $Z$  diasumsikan berdistribusi  $N(0, 1)$  dan dilakukan pengkategorian terhadap  $Y^*$  secara ordinal yaitu untuk  $Y^* \leq y_1$  dikategorikan sebagai  $Y = 0$ , untuk  $y_1 < Y^* \leq y_2$  dikategorikan sebagai  $Y=1$ , untuk  $y_{j-1} < Y^* \leq y_j$  dikategorikan sebagai  $Y = j - 1$ , dan untuk  $Y^* > y_j$  dikategorikan sebagai  $Y = j$ . Model regresi probit ordinal dapat ditulis sebagai berikut:

$$P(Y = 0) = \phi\left(\frac{\delta_1 - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) \tag{6}$$

$$P(Y = 1) = \phi\left(\frac{\delta_2 - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) - \phi\left(\frac{\delta_1 - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) \tag{7}$$

$$P(Y = j) = 1 - \phi\left(\frac{\delta_j - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) \tag{8}$$

dengan  $Y = 0$  untuk kategori terendah dan  $Y = j$  untuk kategori tertinggi dan  $\phi(\cdot)$  adalah fungsi distribusi kumulatif (CDF) pada distribusi normal. (Greene, 2000)

**Estimasi Parameter Regresi Probit Ordinal**

Pengestimasian parameter pada model regresi probit ordinal dapat dilakukan dengan metode *Maximum Likelihood Estimation* (MLE).

Berdasarkan hasil penaksiran untuk parameter  $\beta$  dengan metode MLE tidak dapat ditemukan bentuk yang bentuk tertutup/*closed form*. Oleh karena itu, penaksir parameter  $\beta$  tidak bisa diperoleh secara analitis. Metode numerik Newton– Raphson dengan bantuan komputasi akan sangat mungkin untuk bisa mendapatkan nilai parameter  $\beta$  dari model regresi probit ordinal. (Nurmalasari, 2017)

Tahapan-tahapan untuk mendapatkan penaksir parameter  $\beta$  dengan metode Newton-Raphson adalah sebagai berikut:

1. Menentukan nilai awal  $\boldsymbol{\beta}^{(0)}$ .
2. Menghitung nilai  $\mathbf{g}^{(0)}$  dan  $\mathbf{H}^{(0)}$  yang besarnya tergantung pada  $\boldsymbol{\beta}^{(0)}$ .
3. Memulai dari  $t = 0$ , lakukan iterasi:  $\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - [\mathbf{H}^{(t)}]^{-1} \mathbf{g}^{(t)}$ .
4. Mengentikan proses iterasi jika sudah konvergen atau  $|\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}| \leq \epsilon$ . Jika tidak,

maka teruskan ke  $t = t + 1$ , hitung  $\mathbf{g}^{(t)}$  dan  $\mathbf{H}^{(t)}$ , kemudian kembali ke langkah (3). (Wulandari, 2009)

**Efek Marjinal dari Regresi Probit Ordinal**

Dalam penginterpretasian suatu model regresi probit ordinal, digunakan rumus efek marjinal (*marginal effect*). Rumus-rumus efek marjinal ditulis sebagai berikut (Greene, 2000):

1. Rumus untuk menentukan besarnya pengaruh variabel bebas  $X_k$  untuk  $k = 1, 2, \dots, p$  terhadap  $P(Y = 0)$

$$\frac{\partial P(Y = 0)}{\partial x_k} = \phi\left(\frac{\delta_1 - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) \left(\frac{-\beta_k}{\sigma}\right) \tag{9}$$

2. Rumus untuk menentukan besarnya pengaruh variabel bebas  $X_k$  untuk  $k = 1, 2, \dots, p$  terhadap  $P(Y = 1)$

$$\frac{\partial P(Y = 1)}{\partial x_k} = \phi\left(\frac{\delta_2 - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) - \phi\left(\frac{\delta_1 - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) \left(\frac{\beta_k}{\sigma}\right) \tag{10}$$

3. Rumus untuk menentukan besarnya pengaruh variabel bebas  $X_k$  untuk  $k = 1, 2, \dots, p$  terhadap  $P(Y = j)$

$$\frac{\partial P(Y = j)}{\partial x_k} = \phi\left(\frac{\delta_j - \mathbf{X}^T \boldsymbol{\beta}}{\sigma}\right) \left(\frac{\beta_k}{\sigma}\right) \tag{11}$$

**Pengujian Signifikansi Parameter**

Dalam model regresi probit ordinal, harus melewati tahap pengujian signifikansi parameter terlebih dahulu agar model regresi probit ordinal yang diperoleh *valid*. Pengujian signifikansi parameter ada dua, yaitu secara simultan dan parsial.

1. Uji Simultan (Uji Rasio *Likelihood/Likelihood Ratio*)

Secara umum, hipotesis yang akan digunakan dalam uji ini adalah:

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_p = 0$$

(Semua variabel bebas tidak mempengaruhi variabel tak bebas)

$$H_1 : \exists \beta_k \neq 0 \text{ untuk } k = 0, 1, 2, \dots, p$$

(Minimal terdapat satu variabel bebas yang mempengaruhi variabel tak bebas)

Statistik uji yang digunakan untuk uji simultan adalah:

$$G^2 = -2 \ln\left(\frac{L_1(\theta)}{L_2(\theta)}\right) \tag{12}$$

dengan:

$L_1(\theta)$  : fungsi *likelihood* tanpa variabel bebas

$L_2(\theta)$  : fungsi *likelihood* dengan variabel bebas

Kriteria uji yang digunakan adalah perbandingan nilai  $G^2$  dengan nilai  $\chi^2_{(\alpha;p)}$  dengan  $\alpha$  adalah taraf signifikansi yang digunakan dan  $p$  adalah banyaknya variabel bebas. Jika nilai  $G^2 > \chi^2_{(\alpha;p)}$ , maka  $H_0$  ditolak.

2. Uji Parsial (Uji Wald)

Secara umum, hipotesis yang akan digunakan dalam uji ini adalah:

$H_0 : \beta_k = 0, k = 0, 1, 2, \dots, p$

(Variabel bebas tidak berpengaruh terhadap variabel tak bebas secara signifikan)

$H_1 : \beta_k \neq 0, k = 0, 1, 2, \dots, p$

(Variabel bebas berpengaruh terhadap variabel tak bebas secara signifikan)

Statistik uji yang digunakan untuk uji parsial adalah:

$$W = \left( \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \right)^2 \tag{13}$$

dengan:

$\hat{\beta}_k$  : penaksir parameter  $\beta_k$

$SE(\hat{\beta}_k)$  : nilai *Standard Error* dari parameter  $\beta_k$ .

Nilai  $SE(\hat{\beta}_k)$  dapat dicari dengan menggunakan rumus

$$SE(\hat{\beta}_k) = \sqrt{Var(\hat{\beta}_k)} \tag{14}$$

Kriteria uji yang digunakan adalah perbandingan nilai  $W$  dengan nilai  $\chi^2_{(\alpha;1)}$  dengan  $\alpha$  adalah taraf signifikansi yang digunakan. Jika nilai  $W > \chi^2_{(\alpha;1)}$ , maka  $H_0$  ditolak.

(Nurmalasari, 2017)

**Pengujian Kelayakan Model**

Hipotesis yang digunakan dalam pengujian ini adalah:

$H_0$  : Tidak terdapat perbedaan antara hasil observasi dengan hasil prediksi. (Model layak digunakan)

$H_1$  : Terdapat perbedaan antara hasil observasi dengan hasil prediksi. (Model tidak layak digunakan)

Statistik uji yang digunakan dalam pengujian ini adalah:

$$D = -2 \sum_{i=1}^n \sum_{j=1}^G \left[ y_{ij} \ln \left( \frac{\hat{\pi}_{ij}}{y_{ij}} \right) \right] \tag{14}$$

dengan:

$n$  : banyaknya sampel

$G$  : banyaknya kategori dalam variabel tak bebas

$y_{ij}$  : nilai observasi ke- $i$  pada kategori ke- $j$  dalam variabel tak bebas

$\hat{\pi}_{ij}$  : nilai peluang observasi ke- $i$  pada kategori ke- $j$ .

Kriteria uji yang digunakan adalah perbandingan nilai  $D$  dengan nilai  $\chi^2_{(\alpha;df)}$  dengan  $\alpha$  adalah taraf signifikansi yang digunakan. Nilai  $df$  dapat dicari dengan rumus:

$$df = J - (p + 1) \tag{15}$$

di mana  $J$  adalah jumlah kovariatnya, dan  $p$  adalah banyaknya parameter dalam model. Jika nilai  $D > \chi^2_{(\alpha;df)}$ , maka  $H_0$  ditolak.

(Nurmalasari, 2017)

**Akaike's Information Criterion (AIC)**

Salah satu kriteria yang terbaik di samping pengujian signifikansi parameter dalam model tersebut adalah *Akaike Information Criterion* (AIC) atau Kriteria Informasi Akaike. Model yang optimal adalah model yang memiliki nilai pengepasan yang hampir mendekati dengan nilai aktual. Rumus yang digunakan dalam menentukan nilai AIC adalah sebagai berikut (Agregti, 2007):

$$AIC = -2 (\log \text{likelihood} - \text{banyaknya parameter}) \tag{16}$$

Penentuan nilai AIC juga dapat dihitung dengan menggunakan rumus sebagai berikut (Nurmalasari, 2017):

$$AIC = \frac{-2 \ln(P^*)}{n} + \frac{p^*}{n} \tag{17}$$

di mana  $\ln(P^*)$  adalah nilai *Maximum Likelihood* pada model yang mengandung  $p^*$  variabel bebas,  $p^*$  adalah banyaknya parameter dalam model (konstanta termasuk) dan  $n$  adalah ukuran sampel.

**Apparent Error Rate (APER)**

Rumus yang digunakan dalam penentuan nilai APER adalah sebagai berikut:

$$APER = \frac{\text{Banyak Sampel Salah Klasifikasi}}{\text{Banyak Sampel Keseluruhan}} \tag{18}$$

× 100%

Satuan nilai APER adalah dalam persen. Ketepatan klasifikasi ditentukan dengan mengurangkan nilai 100 dengan nilai APER yang diketahui. Semakin kecil nilai APER, maka semakin besar nilai ketepatan klasifikasinya.

(Nurmalasari, 2017)

**Indeks Pembangunan Manusia (IPM)**

Indeks pembangunan manusia (IPM) adalah suatu ukuran yang digunakan untuk mengukur pencapaian pembangunan manusia di suatu wilayah. IPM menjelaskan bagaimana penduduk dapat mengakses hasil pembangunan dalam perolehan pendapatan, kesehatan, pendidikan, dan sebagainya. Berbagai manfaat dari Indeks Pembangunan Manusia adalah sebagai berikut (BPS, 2018):

1. Sebagai indikator yang penting untuk mengukur keberhasilan dalam upaya membangun kualitas hidup manusia.
2. Sebagai penentu peringkat atau level pembangunan suatu wilayah atau negara.
3. Bagi Indonesia, sebagai data strategis karena sebagai ukuran kinerja Pemerintah, IPM juga digunakan sebagai salah satu alokator penentuan Dana Alokasi Umum (DAU).

IPM diperkenalkan oleh *United Nations Development Programme* (UNDP) yang dilaporkan melalui *Human Development Report* pada tahun 1996 dan berlanjut secara terus menerus tiap tahunnya. Pembangunan manusia didefinisikan sebagai “*a process of enlarging people’s choices*” atau proses yang meningkatkan aspek kehidupan masyarakat. Aspek kehidupan masyarakat dilihat dari umur seseorang yang panjang dan gaya hidup yang sehat, jenjang pendidikan yang memadai, dan standar hidup yang layak. Secara spesifik UNDP menetapkan elemen utama dalam pembangunan manusia, yaitu produktivitas (*productivity*), pemerataan (*equity*), keberlanjutan (*sustainability*), dan pemberdayaan (*empowerment*) (Setiawan, 2013).

Ada beberapa komponen yang harus dihitung sebelum memperoleh nilai IPM di suatu wilayah dan setiap komponen IPM harus distandarisasi dengan nilai minimum dan maksimum. Komponen-komponen yang dimaksud adalah kesehatan, pendidikan, dan pengeluaran. Rumus-rumus yang digunakan adalah sebagai berikut:

1. Dimensi Kesehatan

$$I_{\text{kesehatan}} = \frac{AHH - AHH_{\min}}{AHH_{\max} - AHH_{\min}} \quad (19)$$

2. Dimensi Pendidikan

$$I_{\text{HLS}} = \frac{HLS - HLS_{\min}}{HLS_{\max} - HLS_{\min}} \quad (20)$$

$$I_{\text{RLS}} = \frac{RLS - RLS_{\min}}{RLS_{\max} - RLS_{\min}} \quad (21)$$

$$I_{\text{Pendidikan}} = \frac{I_{\text{HLS}} + I_{\text{RLS}}}{2} \quad (22)$$

3. Dimensi Pengeluaran

$$I_{\text{Pengeluaran}} = \frac{\ln(Pglrn) - \ln(Pglrn_{\min})}{\ln(Pglrn_{\max}) - \ln(Pglrn_{\min})} \quad (23)$$

dengan:

- AHH : Angka Harapan Hidup
- HLS : Harapan Lama Sekolah
- RLS : Rata-rata Lama Sekolah
- Pglrn : Pengeluaran.

(BPS, 2018)

IPM dibentuk melalui tiga dimensi dasar, yaitu dimensi umur panjang, dimensi pendidikan, dan dimensi hidup yang layak. Ketiga dimensi memiliki pengertian yang sangat luas karena terkait banyak faktor. Dirumuskan sebagai berikut (Pratowo, 2013):

$$IPM = \sqrt[3]{I_{\text{kes.}} \times I_{\text{pend.}} \times I_{\text{pglrm}}} \times 100 \quad (24)$$

**Hasil dan Pembahasan**

**Statistika Deskriptif**

Jumlah sampel yang diambil pada variabel Harapan Lama Sekolah (HLS), Angka Harapan Hidup (AHH), Pengeluaran per Kapita (PPK), Rata-rata Lama Sekolah (RLS), dan Tingkat Partisipasi Angkatan Kerja (TPAK) adalah sebanyak 56 data. Statistika deskriptif disajikan dalam Tabel 1 dan 2.

**Tabel 1** Analisis Statistika Deskriptif (Data Numerik)

Variabel Penelitian	Banyaknya Data	Rata-rata (Mean)
AHH (X <sub>1</sub> )	56	70,14
HLS (X <sub>2</sub> )	56	12,54
PPK (X <sub>3</sub> )	56	10.218
RLS (X <sub>4</sub> )	56	8,04
TPAK (X <sub>5</sub> )	56	72,22

**Tabel 2** Analisis Statistika Deskriptif (Data Numerik) (Sambungan dari Kolom Sebelumnya)

Variabel Penelitian	Nilai Maksimum	Nilai Minimum
AHH (X <sub>1</sub> )	73,97	62,94
HLS (X <sub>2</sub> )	14,92	11,12

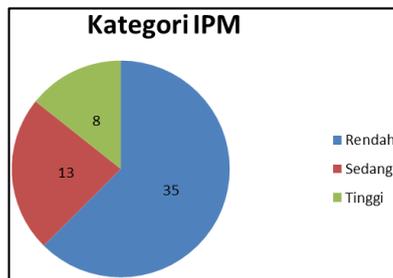
**Tabel 2** Analisis Statistika Deskriptif (Data Numerik) (Sambungan dari Kolom Sebelumnya) (Lanjutan)

Variabel Penelitian	Nilai Maksimum	Nilai Minimum
PPK (X <sub>3</sub> )	16.271	6.680
RLS (X <sub>4</sub> )	11,03	5,85
TPAK (X <sub>5</sub> )	89,42	59,69

Banyaknya kota/kabupaten yang termasuk kategori Tinggi adalah sebanyak 8 data. Sedangkan, banyaknya kota/kabupaten yang termasuk kategori Sedang adalah sebanyak 13 data dan banyaknya kota/kabupaten yang termasuk kategori Rendah adalah sebanyak 35 data. Statistika deskriptif untuk data kategorik disajikan dalam Tabel 3.

**Tabel 3** Analisis Statistika Deskriptif (Data Kategorik)

Kategori (Y)	Banyaknya Data
1 = Rendah	35
2 = Sedang	13
3 = Tinggi	8



**Gambar 1** Grafik Kategori IPM

**Persamaan Awal Regresi Probit Ordinal Serta Estimasi Parameternya**

Berdasarkan pada persamaan (4), maka bentuk model awal regresi probit ordinal sebelum dilakukan suatu estimasi parameter adalah sebagai berikut:

$$Y^* = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 \quad (25)$$

Di mana  $X_1$  adalah variabel Angka Harapan Hidup (AHH) yang dinyatakan dalam persen,  $X_2$  adalah variabel Angka Harapan Lama Sekolah (HLS) yang dinyatakan dalam tahun,  $X_3$  adalah variabel Pengeluaran Per Kapita (PPK) yang dinyatakan dalam ribu rupiah per orang per tahun,  $X_4$  adalah variabel Rata-rata Lama Sekolah (RLS) yang dinyatakan dalam tahun, dan  $X_5$  adalah variabel Tingkat Partisipasi Angkatan Kerja (TPAK) yang dinyatakan dalam persen.

Hasil estimasi parameter disajikan dalam model awal regresi probit ordinal sebagai berikut:

$$Y_1^* = 479,36 + 3,12X_1 + 11,61X_2 + 0,01X_3 + 6,03X_4 - 0,06X_5 \quad (26)$$

$$Y_2^* = 512,71 + 3,12X_1 + 11,61X_2 + 0,01X_3 + 6,03X_4 - 0,06X_5 \quad (27)$$

**Uji Simultan dan Uji Parsial**

Kedua pengujian dilakukan secara bertahap untuk menghasilkan model regresi probit ordinal yang terbaik. Pencarian model terbaik dilakukan dengan metode eliminasi mundur. Uji simultan menggunakan uji Likelihood Ratio dan uji parsial menggunakan uji Wald. Berikut adalah masing-masing pengujian untuk model tersebut:

1. Uji Simultan Hipotesis

$H_0 : \beta_0 = \beta_1 = \dots = \beta_p = 0$   
(Semua variabel bebas tidak mempengaruhi variabel tak bebas)

$H_1 : \exists \beta_k \neq 0$  untuk  $k = 0, 1, 2, \dots, p$   
(Minimal terdapat satu variabel bebas yang mempengaruhi variabel tak bebas)

Jika nilai  $G^2 > \chi^2_{(\alpha;p)}$  atau  $p\text{-value} < 0,05$  maka  $H_0$  ditolak.

**Tabel 4** Uji Simultan Bertahap

Tahap	Nilai Uji LR	p-value
Tahap 1 (Tahap Awal)	102,00	0,00
Tahap 2	102,00	0,00
Tahap 3	90,17	0,00
Tahap 4	70,25	0,00

2. Uji Parsial

$H_0 : \beta_k = 0, k = 0, 1, 2, \dots, p$   
(Variabel bebas tidak berpengaruh terhadap variabel tak bebas secara signifikan)

$H_1 : \beta_k \neq 0, k = 0, 1, 2, \dots, p$   
(Variabel bebas berpengaruh terhadap variabel tak bebas secara signifikan)

Jika nilai  $W > \chi^2_{(\alpha;1)}$ , maka  $H_0$  ditolak.

Nilai  $\chi^2_{(0,05;1)} = 3,84$ .

**Tabel 5** Uji Parsial Tahap 1 (Tahap Awal)

Variabel	Nilai Uji W	Keputusan
Konst. 1	$1,08 \times 10^{-2}$	Menerima $H_0$
Konst. 2	$1,08 \times 10^{-2}$	Menerima $H_0$
AHH	$4,00 \times 10^{-3}$	Menerima $H_0$
HLS	$2,03 \times 10^{-3}$	Menerima $H_0$
PPK	$3,97 \times 10^{-3}$	Menerima $H_0$
RLS	$2,02 \times 10^{-3}$	Menerima $H_0$
TPAK	$1,04 \times 10^{-4}$	Menerima $H_0$

**Tabel 6** Uji Parsial Tahap 2

Variabel	Nilai Uji W	Keputusan
Konst. 1	$1,04 \times 10^{-2}$	Menerima $H_0$
Konst. 2	$1,04 \times 10^{-2}$	Menerima $H_0$
AHH	$9,61 \times 10^{-4}$	Menerima $H_0$
HLS	$4,24 \times 10^{-4}$	Menerima $H_0$
PPK	$3,42 \times 10^{-3}$	Menerima $H_0$
RLS	$4,33 \times 10^{-4}$	Menerima $H_0$

**Tabel 7** Uji Parsial Tahap 3

Variabel	Nilai Uji W	Keputusan
Konst. 1	5,60	Menolak $H_0$
Konst. 2	5,62	Menolak $H_0$
AHH	5,71	Menolak $H_0$
PPK	3,51	Menerima $H_0$
RLS	4,32	Menolak $H_0$

Dari Tabel 5, 6, 7, 8, dijelaskan bahwa dalam mendapatkan model terbaik, dilakukan eliminasi salah satu variabel yang memiliki nilai uji statistik

W yang terkecil sampai semua variabel bebas tersebut sudah menolak  $H_0$ .

**Tabel 8 Uji Parsial Tahap 4**

Variabel	Nilai Uji W	Keputusan
Konst. 1	10,83	Menolak $H_0$
Konst. 2	11,49	Menolak $H_0$
AHH	5,96	Menolak $H_0$
RLS	18,39	Menolak $H_0$

Pada Tabel 8, didapatkan model akhir regresi probit ordinal sebagai berikut:

$$Y_1^* = 46,93 + 0,45 X_1 + 1,84 X_4 \quad (28)$$

$$Y_2^* = 50,00 + 0,45 X_1 + 1,84 X_4 \quad (29)$$

**Uji Kesesuaian Model**

Hipotesis

$H_0$  : Tidak terdapat perbedaan antara hasil observasi dengan hasil prediksi. (Model layak digunakan)

$H_1$  : Terdapat perbedaan antara hasil observasi dengan hasil prediksi. (Model tidak layak digunakan)

Jika nilai  $D > \chi^2_{(\alpha,df)}$ , maka  $H_0$  ditolak.

**Tabel 9 Uji Kesesuaian Model**

Nilai Deviance (D)	df	p-value
31,75	108	1,00

Berdasarkan pada Tabel 9, Dapat dikatakan bahwa tidak ada perbedaan antara hasil observasi dan hasil prediksi, dan model final regresi probit ordinal layak untuk digunakan.

**Nilai AIC dan Ketepatan Klasifikasi**

Penghitungan nilai AIC secara manual menggunakan rumus (16) dalam Agresti (2007):

$$AIC = -2 (\text{LogLik} - p^*) \\ = -2 (-15,88 - 4) = 39,75$$

Di mana *LogLik* adalah nilai *Log Likelihood* dan  $p^*$  adalah banyaknya parameter dalam model. Dari perhitungan tersebut, didapat nilai AIC sebesar 39,75.

Hasil tabulasi silang antara observasi dengan prediksi untuk menghitung nilai ketepatan klasifikasi disajikan dalam Tabel 10.

**Tabel 10 Hasil Tabulasi Silang Antara Observasi dengan Prediksi**

Kategori Observasi	Kategori Prediksi		
	Rendah	Sedang	Tinggi
Rendah	32	3	0
Sedang	3	10	0
Tinggi	0	0	8

$$APER = \frac{3+3+0}{32+3+0+3+10+0+0+0+8} \times 100\% \\ = 10,71\%$$

$$100\% - APER = 100\% - 10,71\% = 89,29\%$$

Berdasarkan pada tabel 10, dapat dijelaskan bahwa model final regresi probit ordinal memiliki nilai persentase kesalahan klasifikasi (APER) sebesar 10,71% dan nilai persentase ketepatan klasifikasinya sebesar 89,29%.

**Interpretasi Model**

Berikut ini adalah nilai efek marginal dari masing-masing variabel bebas dan masing-masing kategori:

**Tabel 11 Nilai Efek Marginal**

Kategori	Variabel Bebas	
	$X_1$ (AHH)	$X_4$ (RLS)
Rendah ( $Y = 0$ )	-0,162	-0,664
Sedang ( $Y = 1$ )	0,162	0,662
Tinggi ( $Y = 2$ )	$3,6 \times 10^{-4}$	$1,47 \times 10^{-3}$

Salah satu contoh interpretasi nilai efek marginal berdasarkan Tabel 11 adalah sebagai berikut:

Ketika nilai variabel AHH bertambah satu satuan, maka akan berpengaruh terhadap probabilitas suatu Kota/ Kabupaten di Pulau Kalimantan yang mendapatkan IPM kategori rendah sebesar -0,162 atau akan menurunkan peluang untuk mendapatkan IPM kategori rendah sebesar 0,162.

Berdasarkan contoh di atas, dapat disimpulkan bahwa peluang untuk mendapatkan IPM kategori rendah akan menurun ketika variabel AHH dan atau RLS bertambah satu satuan. Sedangkan, peluang untuk mendapatkan IPM kategori sedang akan meningkat ketika variabel AHH dan atau RLS bertambah satu satuan, dan peluang untuk mendapatkan IPM kategori tinggi akan meningkat sangat sedikit ketika variabel AHH dan atau RLS bertambah satu satuan.

**Kesimpulan**

Berdasarkan pengujian yang telah dilakukan sebelumnya, maka diperoleh kesimpulan sebagai berikut:

1. Model regresi probit ordinal yang telah terbentuk secara final untuk IPM di seluruh kota/kabupaten di Pulau Kalimantan adalah:

$$Y^* = \beta_0 + \beta_1 X_1 + \beta_4 X_4$$

Hasil estimasi parameter dari model di atas adalah:

$$Y_1^* = 46,93 + 0,45 X_1 + 1,84 X_4$$

$$Y_2^* = 50,00 + 0,45 X_1 + 1,84 X_4$$

2. Faktor-faktor yang mempengaruhi IPM di seluruh kota/kabupaten di Pulau Kalimantan berdasarkan model akhir regresi probit ordinal adalah faktor Angka Harapan Hidup (AHH) dan Rata-rata Lama Sekolah (RLS). Itu diperoleh melalui tahap eliminasi variabel yang memiliki nilai statistik hitung *Wald* yang paling kecil dan tidak signifikan.
3. Berdasarkan hasil perhitungan sebelumnya, pada model akhir regresi probit ordinal diperoleh nilai AIC sebesar 39,75 dan nilai ketepatan klasifikasi sebesar 89,29%.

### Daftar Pustaka

- Agresti, A. 2007. *Categorical Data Analysis*. New York: John Wiley and Sons.
- BPS. 2018. *Indeks Pembangunan Manusia*. Dapat diakses pada <https://bps.go.id> pada tanggal 5 Maret 2019.
- BPS Kalimantan Barat. 2018. *Provinsi Kalimantan Barat Dalam Angka 2018*. Dapat diakses pada <https://kalbar.bps.go.id> pada tanggal 2 Agustus 2019.
- BPS Kalimantan Tengah. 2018. *Provinsi Kalimantan Timur Dalam Angka 2018*. Dapat diakses pada <https://kalteng.bps.go.id> pada tanggal 2 Agustus 2019.
- BPS Kalimantan Timur. 2018. *Provinsi Kalimantan Timur Dalam Angka 2018*. Dapat diakses pada <https://kaltim.bps.go.id> pada tanggal 6 Maret 2019.
- BPS Kalimantan Selatan. 2018. *Provinsi Kalimantan Selatan Dalam Angka 2018*. Dapat diakses pada <https://kalsel.bps.go.id> pada tanggal 2 Agustus 2019.
- BPS Kalimantan Utara. 2018. *Provinsi Kalimantan Utara Dalam Angka 2018*. Dapat diakses pada <https://kaltara.bps.go.id> pada tanggal 2 Agustus 2019.
- Dajan, A. 1983. *Pengantar Metode Statistik Jilid I*. Jakarta: LP3ES.
- Greene, W. H. 2000. *Econometrics Analysis Sixth Edition*. New Jersey: Prentice Hall.
- Gujarati, D. 1993. *Ekonometrika Dasar*. Jakarta: Penerbit Erlangga.
- Nurmalasari, R., Ispriyanti, D., dan Sudamo. 2017. Analisis Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia (IPM) Menggunakan Metode Regresi Logistik Ordinal dan Regresi Probit Ordinal (Studi Kasus Kabupaten/ Kota di Jawa Tengah Tahun 2014). *Jurnal GAUSSIAN*, Vol. 6, Nomor 1, Hal. 113- 115.
- Riduwan. 2014. *Pengantar Statistika Sosial*. Bandung: Alfabeta.
- Setiawan, M. B. dan Hakim, A. 2013. Indeks Pembangunan Manusia Indonesia. *Jurnal Economia*, Vol. 9, No. 1, Hal.18-19.
- Soepono, B. 2002. *Statistik dalam Penelitian Ilmu-Ilmu Sosial dan Pendidikan*. Jakarta: Rineka Cipta.
- Sudjana. 2005. *Metoda Statistika*. Bandung: Tarsito.
- Sugiyono. 2005. *Metode Penelitian Bisnis*. Bandung: Alfabeta.
- Suratini. 2018. Model Probability Unit (Probit) Pinjaman Mikro Terhadap Kemiskinan Rumah Tangga di Indonesia. *Jurnal Pro Bisnis*, Vol. 11, No. 1, Hal. 1.
- Wulandari, E. dan Sutanto, H. T. 2009. *Model Regresi Probit untuk Mengetahui Faktor-Faktor yang Mempengaruhi Jumlah Penderita Diare di Jawa Timur*. *Jurnal MATHunesa*, Vol. 2, No. 1, Hal. 2.
- Yuliza, M., Nugroho, S., dan Rini, D. S. 2017. *Model Logit dan Probit Faktor-Faktor Keberhasilan Belajar Mahasiswa pada Sistem Pembelajaran Jarak Jauh (SPJJ) (Studi Kasus: Mahasiswa Jurusan Manajemen Fakultas Ekonomi Universitas Terbuka (JME-UT))*. Bengkulu: Universitas Terbuka. Dapat diakses pada <http://sigitnugroho.id/eSkripsi/2017/08/Model%20Logit%20dan%20Probit%20Faktor%20Keberhasilan%20Belajar.pdf> tanggal 28 Januari 2019.