MATHunesa
Jurnal Ilmiah Matematika **Volume 12 No 03**
e-ISSN : 2716-506X | p-ISSN : 2301-9115 **Tahun 2024**

PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS PEMBANGUNAN MANUSIA JAWA TIMUR DENGAN REGRESI LOGISTIK ORDINAL

Mohammad Dian Purnama
Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
e-mail: mohammaddian.20053@mhs.unesa.ac.id

A'yunin Sofro
Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
*e-mail: ayuninsofro@unesa.ac.id

Abstrak

Indeks Pembangunan Manusia (IPM) merupakan tolok ukur utama kemajuan dan kesejahteraan wilayah, serta indikator penting dalam menilai peningkatan mutu hidup manusia. Pada tahun 2023, IPM Jawa Timur mengalami kenaikan 3 tahun terakhir dengan nilai terakhir 73,38. Walaupun mengalami kenaikan, Jawa Timur merupakan provinsi dengan IPM terendah di Pulau Jawa dan Bali. Sehingga dari permasalahan tersebut tentunya diperlukan untuk memahami secara menyeluruh faktor-faktor yang berpengaruh terhadap IPM. Tujuan penelitian ini tentunya untuk memberikan gambaran tentang faktor-faktor yang mempengaruhi IPM dalam merancang kebijakan yang tepat di masa mendatang. Data yang digunakan merupakan tingkat IPM Jawa Timur 2023 dalam skala kategori ordinal. Untuk menginterpretasikan data ordinal, diperlukan metode yang sesuai yakni regresi logistik ordinal. Hasil penelitian menunjukkan bahwa variabel persentase penduduk miskin berpengaruh signifikan terhadap IPM.
Kata Kunci: Regresi Logistik Ordinal, IPM, Pemodelan.

Abstract

The Human Development Index (HDI) is a key measure of regional progress and prosperity, as well as an important indicator in assessing improvements in the quality of human life. In 2023, East Java's HDI has increased in the last 3 years with the latest value of 73.38. Despite the increase, East Java is the province with the lowest HDI in Java and Bali. So that from these problems it is certainly necessary to thoroughly understand the factors that affect HDI. The purpose of this study is of course to provide an overview of the factors that influence HDI in designing appropriate policies in the future. The data used is the 2023 East Java HDI level on an ordinal category scale. To interpret ordinal data, a suitable method is needed, namely ordinal logistic regression. The results show that the variable percentage of poor people has a significant effect on HDI.
Keywords: *Ordinal Logistic Regression, HDI Modeling.*

PENDAHULUAN

Tujuan utama pembangunan adalah kesejahteraan masyarakat. Manusia bukan hanya menjadi objek pembangunan, tetapi juga diharapkan dapat menjadi subjek yang aktif, sehingga dapat memberikan kontribusi yang bermanfaat bagi kemajuan suatu wilayah dan secara keseluruhan bagi kemajuan suatu negara. Keberhasilan pembangunan dinilai berdasarkan beberapa parameter, dengan Indeks Pembangunan Manusia (IPM) atau Human Development Index (HDI) yang menjadi salah satu yang paling populer saat ini (Pratowo, 2012).

Indeks Pembangunan Manusia (IPM) telah menjadi salah satu tolok ukur utama dalam mengevaluasi kemajuan dan kesejahteraan suatu wilayah. IPM digunakan sebagai indikator penting untuk menilai upaya pembangunan yang ditujukan pada peningkatan mutu hidup manusia (Manurung & Hutabarat, 2021) .Badan Pusat Statistik (BPS) (2021) juga menekankan pentingnya IPM sebagai indikator yang penting dalam menilai dimensi pembangunan yang lebih luas, terutama dalam upaya meningkatkan kualitas hidup manusia. IPM menggambarkan bagaimana masyarakat bisa memanfaatkan hasil pembangunan untuk mendapatkan penghasilan, layanan kesehatan, pendidikan, dan lain-lain.

Peningkatan IPM di Indonesia telah terjadi secara konsisten setiap tahunnya (Fitriyah, 2021). Pada tahun 2023, IPM Jawa Timur mengalami kenaikan selama 3 tahun terakhir, dengan nilai terakhir mencapai 73.38%. Meskipun mengalami peningkatan, Jawa Timur masih menjadi provinsi dengan IPM terendah di Pulau Jawa dan Bali (BPS, 2023).

Tentunya diperlukan kebijakan dalam meningkatkan IPM di Jawa Timur. Sehingga tujuan penelitian ini adalah untuk memberikan gambaran yang lebih komprehensif dalam merancang kebijakan yang tepat di masa mendatang. Dalam konteks ini, penelitian ini tetap relevan dan penting untuk mengidentifikasi faktor-faktor yang mempengaruhi IPM di Jawa Timur. Analisis faktor-faktor yang mempengaruhi akan memberikan wawasan yang lebih mendalam tentang dinamika pembangunan manusia di Jawa Timur yang lebih maksimal.

Penelitian sebelumnya mengenai IPM telah dilakukan oleh Hoyyi dkk. (2018) yang melakukan pemodelan IPM dan persentase penduduk miskin di Jawa Tengah menggunakan pendekatan regresi multivariat. Salsabila & Husaini (2023) juga pernah melakukan penelitian terkait faktor yang mempengaruhi IPM di Provinsi Sumatera Bagian Selatan yang meliputi Sumatera Selatan, Kepulauan Bangka Belitung, Jambi, Bengkulu, dan Lampung. Kedua peenelitian terkait pemodelan faktor-faktor itu memberikan gambaran tentang hubungan antara variabel-variabel yang kompleks dan seringkali saling terkait.

Selain dalam bentuk nominal, IPM sendiri juga memiliki pengelompokan kategori seperti rendah, sedang, tinggi, hingga sangat tinggi (Prasetyoningrum & Sukmawati, 2018). Pengelompokan kategori tersebut yang memiliki suatu tingkatan biasa dinamakan skala ordinal. Dalam analisis data, pengelompokan data ke dalam kategori-kategori tertentu, seperti data ordinal, sering digunakan untuk mempermudah pemahaman dan analisis. Data ordinal sendiri mempunyai karakteristik kategori seperti data nominal, namun memiliki perbedaan dalam derajat, urutan, atau peringkat dalam objek tersebut (Sartika, 2010).

Untuk mengetahui hubungan antar variabel prediktor dan respon, terutama variabel respon dengan data ordinal, diperlukan metode yang sesuai. Dalam menginterpretasikan bentuk data ordinal itu, penggunaan regresi logistik ordinal dianggap sebagai pendekatan yang tepat untuk menangani kompleksitas tersebut .

KAJIAN TEORI

KELUARGA EKSPONENSIAL

Sebuah konsep penting yang mendasari *Generalized Linear Model* (GLM) adalah distribusi keluarga eksponensial. Anggota keluarga distribusi eksponensial semuanya memiliki fungsi kepadatan probabilitas untuk respon yang teramati y yang dapat ditulis dalam bentuk berikut

$$f(y;\theta,\phi) = exp\left(\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} - c(y_i;\phi)\right) \qquad (1)$$

Dimana $a(\phi); b(\theta_i); c(y_i;\phi)$ adalah fungsi spesifik, θ_i adalah parameter natural, sedangkan ϕ adalah parameter dispersi dan fungsi $a(\phi)$ umumnya berupa $a(\phi) = \phi.\omega$, yang mana ω adalah konstanta yang diketahui (Myers dkk., 2012). Distribusi Binomial, Multinomial, Poisson, Gamma, Beta, dan Normal adalah anggota keluarga eksponensial. Dalam penelitian ini hanya akan membahas tentang distribusi multinomial. Karena distribusi keluarga eksponensial merupakan konsep yang mendasari GLM, maka *subsection* berikutnya akan menjelaskan tentang GLM.

GENERALIZED LINEAR MODEL

Generalized Linear Model (GLM) diperkenalkan pertama kali oleh Nelder dan Wedderburn pada tahun 1972, yang merupakan salah model regresi untuk menganalisis variabel respons, baik bersifat diskrit maupun kontinu. GLM juga tidak mengharuskan data berdistribusi normal, tetapi termasuk dengan distribusi keluarga eksponensial. Dalam GLM, diasumsikan bahwa pengamatan adalah independen dan tidak mempertimbangkan korelasi di antara hasil dari berbagai pengamatan. Komponen acak dari GLM melibatkan variabel respons, yang diasumsikan sebagai Y_i dengan $i = 1,2,3,\ldots,n$ (Agresti, 2015). Bentuk GLM sendiri sebagai berikut:

$$g(\mu_i) = g[E(Y_i)] = \sum_{i=1}^{n} \boldsymbol{\beta}_i X_i \qquad (2)$$

dimana: X_i merupakan vektor variabel regresi untuk pengamatan ke-*i*, $\boldsymbol{\beta}_i$ adalah koefisien regresi atau vektor parameter ke-$i, i = 0,1,2,\ldots,p$, dan g = *link function.*

Menurut Agresti (2015), Struktur GLM memiliki tiga komponen, yaitu komponen acak, komponen sistematis, dan *link function*.

- Komponen acak merupakan komponen dari variabel respons Y_i dengan *mean* $E(Y_i) = \mu_i$.

Pengamatan Y_i memiliki distribusi yang termasuk dalam keluarga eksponensial.

- Komponen sistematis merupakan komponen yang meliputi variabel prediktor (x_i) dan memiliki bentuk umum dari $x_i'\boldsymbol{\beta}$. Komponen ini juga dapat dinyatakan dengan $\eta_i = \sum_{i=1}^{n} \boldsymbol{\beta_i x_i}$ dengan $i = 1,2,\ldots,p$.
- *Link Function* atau fungsi penghubung merupakan fungsi dari nilai yang diharapkan (*mean*) dari Y. *Link function* (g) juga menghubungkan komponen random dan komponen sistematis. Diketahui $\mu_i = E(Y_i)$ dimana $i = 1,2,\ldots,n$. Model untuk menghubungkan μ_i dengan η_i adalah g sehingga $g(\mu_i) = \eta_i$. Fungsi g menghubungkan $E(Y_i)$ dengan variabel prediktor yaitu:

$$g(\mu_i) = \eta_i = \sum_{i=1}^{n} \boldsymbol{\beta_i x_i} \quad (3)$$

Diantara fungsi- fungsi link yang dapat digunakan, ada yang disebut fungsi link kanonik yaitu fungsi hubungan yang terjadi pada saat $\eta_i = \beta_0 + \sum_{i=1}^{p} \boldsymbol{\beta_i X_i}$. Ada berbagai kemungkinan pilihan link kanonik pada formulasi ini yang akan ditunjukkan pada tabel berikut.

Tabel 1. Link Kanonik

Distribusi	Link kanonik
Normal	$\eta_i = \mu_i$
Binomial	$\eta_i = \log\left(\frac{\mu_i}{1-\mu_i}\right)$
Multinomial	$\eta_i = \log\left(\frac{\mu_i}{\mu_k}\right), i < k$
Gamma	$\eta_i = \frac{1}{\mu_i}$
Poisson	$\eta_i = \log(\mu_i)$

Kemudian ada berbagai fungsi *link function* yang akan ditunjukkan pada tabel berikut

Tabel 2. *Link Function*

Fungsi	*Link Function*
Logit	$\eta_i = \log\left(\frac{\mu_i}{1-\mu_i}\right)$
Probit	$\eta_i = \Phi^{-1}(\mu_i)$
log-log	$\eta_i = \log[-\log(1-\mu_i)]$

Model regresi logistik ordinal pada penelitian ini menggunakan fungsi logit dengan data variabel responnya berdistribusi multinomial. Pada subsection akan dijelaskan tentang distribusi multinomial.

DISTRIBUSI MULTINOMIAL

Distribusi multinomial merupakan salah satu distribusi yang paling umum digunakan pada *Generalized Linear Model* (GLM). Apabila n percobaan berulang dapat menghasilkan lebih dari 2 *outcome* yang mungkin dengan probabilitas masing-masing konstan pada setiap percobaan, akan dihasilkan distribusi multinomial. Distribusi peluang multinomial digunakan sebagai penentuan peluang hasil yang di kategorikan ke dalam lebih dari dua kelompok. Distribusi multinomial merupakan keluarga eksponensial, yang didefinisikan sebagai berikut.

$$\pi_i = P(Y_i = r_k) \quad (4)$$

dengan π_i merupakan probabilitas respon populasi ke-i dan respon kategori r_k. Untuk r pengamatan independen dengan r_1 termasuk dalam kategori 1, r_2 termasuk dalam kategori 2, ..., r_k termasuk dalam kategori k, fungsi probabilitas multinomial dapat dirumuskan sebagai berikut

$$P(Y_i = r_1, r_2, \ldots, r_k) = \frac{r!}{r_1! r_2! \ldots r_k!} \pi_1 \pi_2 \ldots \pi_k \quad (5)$$

dengan $\sum_{i=1}^{k} r_k = n$ dan $\sum_{i=1}^{k} \pi_i = 1$. Sehingga untuk nilai mean dan juga varians dari distribusi multinomial berturut-turut adalah $E(r_i) = n\pi_i$ dan $var\ (r_i) = r\pi_i(1-\pi_i)$. Metode yang akan digunakan dalam penelitian ini adalah metode regresi logistik ordinal yang juga akan dijelaskan pada *subsection* selanjutnya.

REGRESI LOGISTIK ORDINAL

Regresi logistik ordinal merupakan suatu metode statistik yang digunakan untuk menentukan model dari variabel respons yang memiliki sifat ordinal yang berarti variabel tersebut mempunyai kategori yang diatur dalam urutan atau tingkatan tertentu. Regresi logistik merupakan salah satu metode regresi non-linier yang dapat digunakan untuk mencari hubungan dikotomus (nominal atau ordinal dalam dua kategori) atau polikotomus (berskala nominal atau ordinal dengan lebih dari dua kategori) dengan satu atau beberapa variabel prediktor (Ruspriyanti & Sofro, 2018).

Pada model ini terdapat sifat ordinal dari variabel respon yang dituangkan dalam peluang kumulatif dengan fungsi logit. Variabel respon estimasi ke-i (Y_i) sendiri yang memiliki nilai observasi y_i, yang mana pada kasus data ordinal akan bernilai sama dengan r_i

sebagai interprestasi data kategorik (Riadi & Kartikasari, 2020).

Misal $\pi_r(x) = P(Y_i \leq r|x)$ yaitu probabilitas bahwa respon bernilai r dengan total c, diberikan vektor x yang berisi j kovariat, dan γ menyatakan probabilitas kumulatif kategori respon, pada persamaan berikut (Riadi & Kartikasari, 2020).

$$\gamma_r(x) = P(Y_i \leq r|x) = \pi_1(x) + \cdots + \pi_r(x) \quad (6)$$
$$;\ r = 1,,2,\ldots,c$$

Model logit kumulatif terdiri dari beberapa jenis model, salah satunya pada model *propotional odds*. Pada model *propotional odds*, setiap logit kumulatif memiliki koefisien intersep yang berbeda, namun mempunyai nilai efek $\boldsymbol{\beta}$ yang sama. Pada model propotional odds transformasi logit dilakukan pada probabilitas kumulatif respon γ_r (Riadi & Kartikasari, 2020). Model ini adalah model logistik ordinal yang sering digunakan, dengan melakukan eksponensiasi terhadap persamaan diatas diperoleh odds $Y \leq r$ atau odds respon berada pada kategori kurang dari atau sama dengan *r*. Sehingga didapatkan persamaannya sebagai berikut (Riadi & Kartikasari, 2020)

$$\begin{aligned} \text{Logit } \gamma_r(x) &= \log \frac{\gamma_r(x)}{1-\gamma_r(x)} \\ &= \log \frac{\pi_1(x)+\cdots+\pi_r(x)}{1-(\pi_1(x)+\cdots+\pi_r(x))} \\ &= \log \frac{\pi_1(x)+\cdots+\pi_r(x)}{\pi_{r+1}(x)+\cdots+\pi_c(x)} \\ &= \log\left(\exp\left(\alpha_r - \left(\textstyle\sum_{j=1}^{p} \boldsymbol{\beta}_j \boldsymbol{x}_{ij}\right)\right)\right) \\ &= \alpha_r - \left(\textstyle\sum_{j=1}^{p} \boldsymbol{\beta}_j \boldsymbol{x}_{ij}\right) \end{aligned} \quad (7)$$

di mana $P(Y_i \leq r|x_{ij})$ adalah fungsi probabilitas, α_r merupakan koefisien intersep pada kategori *r*, x_{ij} adalah variabel prediktor ke-*i*, dan $\boldsymbol{\beta}_j \boldsymbol{x}_{ij}$ merupakan koefisien regresi variabel prediktor ke-*i*

Perhitungan odds $(Y_i \leq r)$ atau peluang suatu peristiwa terjadi dalam suatu kategori respon tertentu (π_r) didefinisikan sebagai berikut (Riadi & Kartikasari, 2020)

$$\pi_r = \frac{\exp\left(\alpha_r - \left(\sum_{j=1}^{p} \boldsymbol{\beta}_j x_{ij}\right)\right)}{1+\exp\left(\alpha_r - \left(\sum_{j=1}^{p} \boldsymbol{\beta}_j x_{ij}\right)\right)} \quad (8)$$

Dalam regresi logistik ordinal, koefisien intersep (α_r), juga merupakan nilai yang membagi rentang variabel laten kontinu ke-*r* (variabel yang tidak dapat diukur secara langsung, tetapi diasumsikan memiliki nilai yang kontinu) menjadi kategori-kategori yang diamati. Koefisien tersebut juga menggambarkan probabilitas suatu observasi jatuh pada kategori tertentu. Dalam model regresi logistik ordinal dengan fungsi logit, juga merupakan nilai peluang yang digunakan untuk mengklasifikasikan hasil prediksi ke dalam kategori biner yaitu kategori pertama atau kedua (*r-1* atau *r*). Dalam interpretasi modelnya, apabila nilai prediksi kurang atau sama dengan dari koefisien intersep (α_r) maka akan diklasifikasikan ke dalam kategori pertama, sedangkan nilai-nilai prediksi yang lebih besar dari koefisien intersep (α_r), maka akan diklasifikasikan ke dalam kategori kedua. Jika variabel dependen memiliki tiga kategori (r=1,2,3), maka menurut Riadi dan Kartikasari (2020) model yang terbentuk adalah

$$\text{Logit } 1 = \pi_1(x) = P(Y_i \leq 1|x) = P(Y_i = 1) \quad (9)$$

$$\begin{aligned} \text{Logit } 2 = \pi_2(x) &= P(Y_i \leq 2|x) - \pi_1 \\ &= P(Y_i \leq 2|x) - P(Y_i = 1) \end{aligned} \quad (10)$$

$$\begin{aligned} \text{Logit } 3 = \pi_3(x) &= P(Y_i > 2|x) \\ &= 1 - P(Y_i \leq 2|x) = 1 - \pi_2 - \pi_1 \end{aligned} \quad (11)$$

Perhitungan odds $(Y_i \leq r)$ atau peluang suatu peristiwa terjadi dalam suatu kategori respon tertentu (π_r) dengan tiga kategori respon didefinisikan sebagai berikut

$$\pi_1(x) = P(Y_i \leq 1|x) = \frac{\exp\left(\alpha_1 - \left(\sum_{j=1}^{p} \boldsymbol{\beta}_j x_{ij}\right)\right)}{1+\exp\left(\alpha_1 - \left(\sum_{j=1}^{p} \boldsymbol{\beta}_j x_{ij}\right)\right)} \quad (12)$$

$$\begin{aligned} \pi_2(x) &= P(Y_i \leq 2|x) - \pi_1(x) \\ &= \frac{\exp\left(\alpha_2 - \left(\sum_{j=1}^{p} \boldsymbol{\beta}_j x_{ij}\right)\right)}{1+\exp\left(\alpha_2 - \left(\sum_{j=1}^{p} \boldsymbol{\beta}_j x_{ij}\right)\right)} - \pi_1(x) \end{aligned} \quad (13)$$

$$\pi_3 = 1 - \pi_2(x) - \pi_1(x) \quad (14)$$

ESTIMASI PARAMETER REGRESI LOGISTIK ORDINAL

Metode MLE dipilih karena mempunyai beberapa kelebihan dibandingkan dengan metode lain, diantaranya dapat digunakan untuk model tidak linear seperti regresi logistik, serta hasil penaksirannya mendekati parameternya (Hosmer dan Lemeshow, 2000). Metode ini mengestimasi parameter $\boldsymbol{\beta}$ dengan memaksimumkan fungsi likelihood dengan syarat data harus mengikuti distribusi tertentu. Jika variabel dependen (*Y*) mengikuti distribusi multinomial dengan peluang hasil yang di kategorikan ke dalam lebih dari dua kelompok, maka fungsi *likelihood* untuk sampel *n* independen observasi adalah sebagai berikut:

$$l(\boldsymbol{\beta}) = \textstyle\sum_{i=1}^{n} [\pi_1(x)^{y_{1i}} \pi_2(x)^{y_{2i}} \pi_3(x)^{y_{3i}}] \quad (15)$$

Dengan nilai $i = 1,2,\ldots,n$, Sehingga didapatkan fungsi *log-likelihood* sebagai berikut

$$L(\boldsymbol{\beta}) = \sum_{i=1}^{n} [y_{1i} \log[\pi_1(x)] + y_{2i} \log[\pi_2(x)] \,| + y_{3i} \log[\pi_3(x)]] \tag{16}$$

Memaksimumkan *log-likelihood* dapat diperoleh dengan cara mendiferensialkan $L(\boldsymbol{\beta})$ terhadap $\boldsymbol{\beta}$ dan menyamakan dengan nol akan diperoleh persamaan. Penyelesaian turunan pertama dari fungsi *log-likelihood* tidak linear, sehingga digunakan metode numerik yaitu iterasi *Newton-Raphson* untuk mendapatkan estimasi parameternya.

$$\boldsymbol{\beta}^{(t+1)} = \boldsymbol{\beta}^{(t)} - \left(\boldsymbol{H}^{(t)}\right)^{-1} \boldsymbol{q}^{(t)} \tag{17}$$

dimana, $\boldsymbol{\beta} = \left[\boldsymbol{\beta_1 \beta_2} \ldots \boldsymbol{\beta_p}\right]'$ adalah parameter regresi,

$$\boldsymbol{q}^{(t)} = \left(\frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1} \frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2} \frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_3} \frac{\partial L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right)$$

$$\boldsymbol{H}^{(t)} = \begin{pmatrix} \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1^2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}_2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}_2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2^2} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2 \partial \boldsymbol{\beta}} \\ \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_1 \partial \boldsymbol{\beta}} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}_2 \partial \boldsymbol{\beta}} & \frac{\partial^2 L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}} \end{pmatrix}$$

Iterasi berhenti apabila terpenuhi kondisi konvergen yakni selisih $|\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}| \leq \varepsilon$ dimana ε merupakan bilangan yang sangat kecil.

UJI SIGNIFIKANSI PARAMETER SECARA SIMULTAN

Pengujian signifikansi secara simultan dalam dilaksanakan dengan tujuan untuk menentukan apakah semua estimasi parameter dalam model signifikan secara keseluruhan. Uji serentak digunakan dengan tujuan untuk mengetahui apakah model telah tepat (signifikan) dan memeriksa pengaruh variabel prediktor di dalam model secara bersama-sama menggunakan uji *chi-square* (Hosmer dkk., 2013). Hipotesis pengujian signifikansi koefisien parameter secara serentak adalah sebagai berikut.

$H_0: \beta_1 = \beta_2 = \cdots = \beta_j$

H_1: Paling tidak terdapat satu $\beta_j \neq 0, j = 1,2,\ldots,p$

Statistik uji:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_2}{n}\right)^{n_2} \left(\frac{n_3}{n}\right)^{n_3}}{\prod_{i=1}^{n} \pi^{Y_i} (1-\pi)^{1-Y_i}} \right] \tag{18}$$

dengan:

n_1 = Banyaknya observasi yang bernilai 1

n_2 = Banyaknya observasi yang bernilai 2

n_3 = Banyaknya observasi yang bernilai 3

n = Banyaknya observasi

Statistik uji G mengikuti Distribusi *Chi-square*, dimana db merupakan banyaknya parameter dalam model. Sehingga diperoleh keputusan tolak H_0 jika nilai statistik uji G lebih dari $\chi^2_{(\alpha,db)}$ atau *p-value* kurang dari α dengan α merupakan derajat kebebasan sebesar 0.05 atau 5%.

UJI SIGNIFIKANSI PARAMETER SECARA PARSIAL

Uji parsial adalah sebuah metode pengujian yang dilakukan pada masing-masing parameter secara terpisah dalam suatu model regresi. Pengujian ini bertujuan untuk menilai signifikansi dari masing-masing parameter dalam model tersebut. Dalam analisis regresi logistik ordinal, uji Z merupakan alat penting untuk menilai signifikansi individu dari koefisien regresi. Setiap koefisien dalam model diuji untuk menentukan apakah ada bukti yang cukup untuk menyatakan bahwa koefisien tersebut berbeda dari nol. Ketika standar deviasi (σ) dari populasi diketahui, maka secara otomatis digunakan pendekatan uji Z dan uji ini sangat cocok pada penelitian ini karena sampel lebih dari 30 (Sofro dkk., 2019). Hipotesis uji Z didefinisikan sebagai berikut:

$H_0 = \beta_j = 0$

$H_1 = \beta_j \neq 0, j = 1,2,\ldots,p$

Dengan statistik uji:

$$Z_{hitung} = \frac{\text{Estimasi Koefisien}}{\text{Standar Eror}} = \frac{\hat{\beta}}{SE(\hat{\beta})} \tag{19}$$

Daerah penolakan H_0 saat $Z_{Hitung} > Z_\alpha$ atau *p-value* $< \alpha$ dengan α merupakan derajat kebebasan sebesar 0.05 atau 5%.

METODE PENELITIAN

DATA PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder yang merupakan data dari publikasi laman Badan Pusat Statistik (BPS) Provinsi Jawa Timur. Data pengamatan yang dianalisis pada penelitian ini adalah seluruh 38 kabupaten/kota di Provinsi Jawa Timur tahun 2023. Pemilihan tahun 2023 didasarkan pada data terbaru pada variabel prediktor yang diterbitkan oleh BPS Jawa Timur pada tahun 2024, yang mencakup informasi dari tahun 2023. Pada penelitian ini variabel respon (Y) yang digunakan adalah Indeks Pembangunan Manusia (IPM) yang telah dikategorikan menjadi data ordinal. Berikut adalah pengkategorian IPM menurut Badan Pusat Statistik.

Tabel 3. Kategori IPM

Skala	Kategori
Sedang	60 ≤ IPM < 70
Tinggi	70 ≤ IPM < 80
Sangat Tinggi	IPM ≥ 80

Sedangkan pada variabel prediktor (x) adalah pertumbuhan ekonomi (x_1), persentase penduduk miskin (x_2), tingkat partisipasi angkatan kerja (x_3), dan pendapatan domestik regional bruto (PDRB) per kapita (x_4).

RANCANGAN PENELITIAN

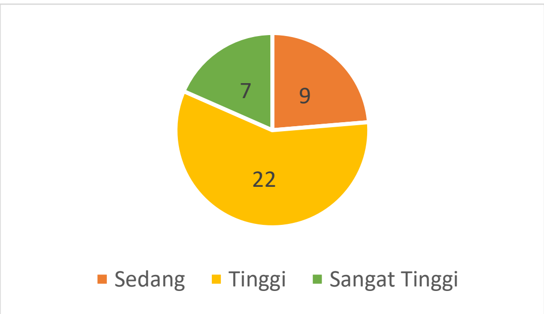
Rancangan penelitian digunakan untuk mencapai tujuan dari penelitian yang dilakukan. Analisis statistik deskriptif untuk mendeskripsikan karakteristik kategori Indeks Pembangunan Manusia (IPM) dan faktor-faktor yang mempengaruhi untuk mendapatkan gambaran dari data yang dikumpulkan. Selanjutnya dilakukan analisis regresi logistik ordinal yang merupakan Analisis untuk mengetahui hubungan variabel prediktor dan variabel respon pada data kategori ordinal. Untuk mendapatkan estimasi parameter pada model regresi logistik ordinal, maka dilakukan penaksiran parameter menggunakan *Maximum Likelihood Estimation* (MLE). Selanjutnya dilakukan pengujian signifikasi parameter model yang dilakukan secara simultan menggunakan uji G dan secara parsial menggunakan uji wald.

HASIL DAN PEMBAHASAN

STATISTIK DESKRIPTIF

Berikut adalah hasil analisis statistik deskriptif dari Indeks Pembangunan Manusia (IPM) berdasarkan kategorinya.

Gambar 1. Karakteristik IPM



Berikut hasil analisis statistik deskriptif faktor-faktor yang mempengaruhi yakni pertumbuhan ekonomi (x_1), persentase penduduk miskin (x_2), tingkat partisipasi angkatan kerja (x_3), dan pendapatan domestik regional bruto (PDRB) per kapita (x_4).

Tabel 4. Kategori IPM

Variabel	Rata-rata	Minimum	Maximum
x_1	4.71	1.20	6.19
x_2	10.29	3.31	21.76
x_3	73.16	66.89	81.64
x_4	70645.06	23842.1	541112.5

MODEL REGRESI LOGISTIK ORDINAL

Hasil estimasi parameter pada regresi logistik ordinal menggunakan *Maximum Likelihood Estimation* (MLE) ditunjukkan pada tabel sebagai berikut:

Tabel 5. Estimasi Parameter

Parameter	Est	Std Error	Z-Value	P-Value
α_1	-14.892	9.843	2.289	0.130
α_2	-8.882	9.540	0.867	0.352
β_1	-0.802	0.570	1.978	0.160
β_2	-0.617	0.196	9.894	0.002
β_3	-0.056	0.114	0.238	0.625
β_4	0.00003	0.00002	2.940	0.086

Berdasarkan tabel di atas dapat dibentuk model regresi logistik ordinal sebagai berikut:

Logit 1

$$Logit\ P(Y_i \leq 1|x) = -14.892 - (-0.802\ x_1 - 0.617\ x_2 - 0.056\ x_3 + 0.00003\ x_4)$$

Logit 2

$$Logit\ P(Y_i \leq 2|x) = -8.882 - (-0.802\ x_1 - 0.617\ x_2 - 0.056\ x_3 + 0.00003\ x_4)$$

Berdasarkan model di atas, maka nilai odds atau model peluang tingkat indeks pembangunan manusia (IPM) masing-masing kategori adalah

IPM dengan kategori sedang

$$\pi_1(x) = \frac{\exp(-14.892 - (-0.802\ x_1 - 0.617\ x_2 - 0.056\ x_3 + 0.00003\ x_4))}{1 + \exp(-14.892 - (-0.802\ x_1 - 0.617\ x_2 - 0.056\ x_3 + 0.00003\ x_4))}$$

IPM dengan kategori sedang

$$\pi_2(x) = \left(\frac{\exp(-8.882 - (-0.802\ x_1 - 0.617\ x_2 - 0.056\ x_3 + 0.00003\ x_4))}{1 + \exp(-8.882 - (-0.802\ x_1 - 0.617\ x_2 - 0.056\ x_3 + 0.00003\ x_4))}\right) - \pi_1(x)$$

IPM dengan kategori sedang

$$\pi_3(x) = 1 - \pi_2(x) - \pi_1(x)$$

Dari persamaan diatas, misal untuk menghitung nilai peluang pada Kabupaten Pacitan, maka nilai variabel pertumbuhan ekonomi (x_1) sebesar 4.76, persentase penduduk miskin (x_2) sebesar 13.65, tingkat partisipasi angkatan kerja (x_3) sebesar 81.64, dan pendapatan domestik regional bruto (PDRB) per kapita, (x_4) sebesar 33148.6 disubstitusi ke dalam model ke dalam persamaan probabilitas kumulatif (π) sebagai berikut.

$\pi_1 = 0.232266758$
$\pi_3 = 0.759686783$
$\pi_3 = 1 - \pi_2 - \pi_1 = 0.008046459$

Berdasarkan nilai peluang dari persamaan di atas, maka probabilitas IPM di Kabupaten Pacitan kategori sedang sebesar 0.2323 atau sekitar 23.23%, kategori tinggi sebesar 0.7597 atau sekitar 75.97%, dan peluang kategori sangat tinggi sebesar 0.0080 atau 0.08%. Dapat disimpulkan Kabupaten Pacitan dapat diprediksi IPM akan masuk pada kategori tinggi.

UJI SIGNIFIKANSI PARAMETER SECARA SIMULTAN

Tahap selanjutnya dalam model regresi logistik ordinal adalah melakukan uji simultan yang digunakan untuk menunjukkan hubungan antara variabel prediktor terhadap variabel respon secara simultan. Uji ini dilakukan dengan menggunakan statistik uji G dengan hipotesis sebagai berikut :

$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0$ atau tidak terdapat variabel prediktor yang berpengaruh signifikan terhadap model
H_1 : Paling sedikit ada satu $\beta_j \neq 0$ dengan *j=1,2,...,p* atau terdapat minimal satu variabel prediktor yang berpengaruh signifikan

Tabel 6. Uji Signifikansi Parameter Secara Simultan

Model	Loglik	df	Chi-Sq	P-Value
Final	36.532	4	37.126	0.000

Berdasarkan hasil uji likelihood ratio test padatabel di atas, diketahui bahwa nilai χ^2 yaitu 37.126 lebih besar dari nilai $\chi^2_{(3;0.05)}$ yaitu 9.488 atau jika berdasar *p-value* diketahui bahwa *p-value* yaitu 0.000 yang mana lebih kecil dari α (0.05). Hal itu berarti diperoleh keputusan H_0 ditolak, yang artinya minimal ada satu variabel independen yang berpengaruh signifikan terhadap IPM di Jawa Timur.

UJI SIGNIFIKANSI PARAMETER SECARA PARSIAL

Pada uji signifikansi parameter secara parsial, nilai uji Wald diperoleh dengan membagi nilai estimasi dengan standar eror atau dapat dilihat pada statistik uji Wald bahwa nilainya setara dengan nilai Z–hitung. Kriteria pengujiannya yaitu tolak H_0 jika $Z_{Hitung} > Z_\alpha$ atau $p - value < \alpha$ dengan $\alpha = 0.05$ atau 5%. Berdasarkan nilai *p-value* pada tabel 5 diketahui bahwa nilai hanya variabel persentase penduduk miskin (x_2) yang meiliki *p-value* < 0.05. Sehingga diambil kesimpulan bahwa hanya variabel persentase penduduk miskin (x_2) yang berpengaruh secara signifikan terhadap IPM di Jawa Timur tahun 2023.

INTERPRETASI HASIL

Nilai koefisien pertumbuhan ekonomi (x_1) terhadap IPM adalah -0.802, sehingga dapat diinterpretasikan bahwa apabila terjadi penurunan pertumbuhan ekonomi sebesar 1% dapat menyebabkan peningkatan IPM sebesar exp (-0.802) = 0.45 kali. Pada persentase penduduk miskin (x_2) nilai koefisien menunjukkan -0.617 yang berarti apabila terjadi penurunan persentase penduduk miskin sebesar 1% maka akan meningkatkan IPM sebesar $\exp(-0.617)$ = 0.54 kali. Sedangkan pada tingkat partisipasi angkatan kerja (x_3), menunjukkan nilai koefisien sebesar -0.056 yang berarti apabila terjadi penurunan TPAK sebesar 1% dapat menyebabkan peningkatan IPM sebesar $\exp(-0.056)$ = 0.94 kali, dan pada pendapatan domestik regional bruto (PDRB) per kapita (x_4), nilai kofisien yang sebesar 0.000003 menunjukkan apabila terjadi peningkatan PDRB sebesar 1% maka akan meningkatkan IPM sebesar $\exp(0.000003)$ = 1 kali.

PENUTUP

SIMPULAN

Dari uji signifikansi parameter secara simultan didapatkan bahwa nilai χ^2 yaitu 36.532 lebih besar dari nilai $\chi^2_{(3;0.05)}$ yaitu 9.488 atau jika berdasar *p-value* yaitu 0.000 yang mana lebih kecil dari α (0.05), maka ada satu variabel prediktor yang berpengaruh signifikan terhadap IPM di Jawa Timur. Dari uji signifikansi parameter secara parsial didapatkan bahwa hanya variabel hanya variabel persentase penduduk miskin (x_2) yang meiliki *p-value* < 0.05. Sehingga diambil kesimpulan bahwa hanya variabel persentase penduduk miskin (x_2) yang berpengaruh secara signifikan terhadap IPM di Jawa Timur tahun 2023.

SARAN

Berdasarkan penelitian yang telah dilakukan, disarankan agar penelitian selanjutnya memperhatikan dengan lebih teliti pemilihan variabel yang akan digunakan. Selain itu, disarankan juga untuk menambahkan variabel prediktor tambahan yang mungkin memiliki pengaruh, namun belum dimasukkan dalam penelitian ini karena keterbatasan data yang tersedia

DAFTAR PUSTAKA

Agresti, A., 2015. Foundations of linear and generalized linear models. John Wiley & Sons

Agresti, A., 2019. An Introduction to Categorical Data Analysis, vol. 792. JohnWiley & Sons

Badan Pusat Statistika Indonesia. (2021). Indeks Pembangunan Manusia Indonesia. https://www.bps.go.id/id/publication/2021/04/30/8e777ce2d7570ced44197a37/indeks-pembangunan-manusia-2020.html

Badan Pusat Statistika Indonesia. (2021). Indeks Pembangunan Manusia menurut Provinsi. https://sumsel.bps.go.id/indicator/26/593/1/-metode-baru-indeks-pembangunan-manusia-menurut-provinsi-.html

Hosmer Jr, D.W., Lemeshow, S., dan Sturdivant, R. X., 2013. Applied logistic regression, vol. 398. John Wiley & Sons.

Hoyyi, A., Safitri, D., Sugito, S., & Prahutama, A. (2018). PENDEKATAN REGRESI LINIER MULTIVARIAT UNTUK PEMODELAN INDEKS PEMBANGUNAN MANUSIA (IPM) DAN PERSENTASE PENDUDUK MISKIN DI JAWA TENGAH. Jurnal Statistika Universitas Muhammadiyah Semarang, 6(2).

Manurung, E. N., & Hutabarat, F. (2021). Pengaruh Angka Harapan Lama Sekolah, Rata-Rata Lama Sekolah, Pengeluaran per Kapita Terhadap Indeks Pembangunan Manusia. Jurnal Ilmiah Akuntansi Manajemen, 4(2), 121-129.

Myers, R. H., Montgomery, D. C., Vining, G. G., dan Robinson, T. J., 2012. Generalized linear models: with applications in engineering and the sciences. John Wiley & Sons.

Prasetyoningrum, A. K., & Sukmawati, U. S. (2018). Analisis pengaruh Indeks Pembangunan Manusia (IPM), pertumbuhan ekonomi dan pengangguran terhadap kemiskinan di Indonesia. Equilibrium: Jurnal Ekonomi Syariah, 6(2), 217-240.

Pratowo, N. I. (2012). Analisis faktor-faktor yang berpengaruh terhadap Indeks Pembangunan Manusia. Jurnal Studi Ekonomi Indonesia, 1(1), 15-31.

Riadi, R. A., dan Kartikasari, M. D., 2020. Implementasi k-means clustering dan regresi logistik ordinal terhadap kinerja cabang pt. x. PROSIDING SENDIKA, 6(1).

Ruspriyanty, D. I., & Sofro, A. (2018, November). Analysis of hypertension disease using logistic and probit regression. In Journal of Physics: Conference Series (Vol. 1108, No. 1, p. 012054). IOP Publishing.

Sartika, E., 2010. Pengolahan data berskala ordinal. Sigma-Mu, 2(1), 60–69.

Sofro, A., Oktaviarina, A., dan Maulana, D. A., 2019. Metode Statistika. Unesa Press