

FAKTOR-FAKTOR YANG MEMPENGARUHI STUNTING DI INDONESIA: ANALISIS *PRINCIPAL COMPONENT REGRESSION*

Rifdah Fadhilah¹, Dina Eka Putri^{2*}, Humami Syifa Amanda³

^{1,3}Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Indonesia

²Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mataram, Indonesia

*e-mail: dina.putri@unram.ac.id

DOI: 10.14710/j.gauss.14.1.128-138

Article Info:

Received: 2025-01-09

Accepted: 2025-06-13

Available Online: 2025-06-21

Keywords:

Principal Components; Indonesia; Stunting Prevalance; PCR; IHS

Abstract: Stunting, characterized by chronic malnutrition due to undernutrition, is a significant public health problem in Indonesia, with a prevalence rate of 21.5% in 2023—a slight decrease from 21.6% the previous year. The large impact of stunting on child health has made it a priority in global and national agendas. This study which aims to identify the causes of stunting to inform targeted policy interventions. Given the multicollinearity among the influencing factors, traditional regression models often fail. Principal Component Regression (PCR) is used to overcome multicollinearity and improve the predictive power of the model. Using secondary data from the 2023 Indonesian Health Survey (IHS), this study identified two principal components from seven initial predictors assumed to influence stunting: Exclusive breastfeeding, low birth weight, undernourished maternal food intake, undernourished child food intake, complete basic immunization, proper sanitation, and hereditary diseases. Component 1 with an eigenvalue of 3.271, explained 53.1% of the variance, while Component 2 with an eigenvalue of 1.425, explained 73.5% of the variance. The PCR method effectively addresses multicollinearity, as evidenced by Variance Inflation Factor (*VIF*) < 5. This study highlights the importance of using advanced statistical methods for robust policy development in the fight against stunting.

1. PENDAHULUAN

Stunting yang ditandai dengan malnutrisi kronis akibat kekurangan gizi (Mayasari et al., 2018) merupakan masalah kesehatan masyarakat yang signifikan di Indonesia, dengan tingkat prevalensi 21,5% pada tahun 2023—hanya sedikit menurun dari 21,6% pada tahun sebelumnya (Dinas Kesehatan Papua, 2024). Menurut Kementerian Kesehatan RI (2023), standar WHO menyebutkan prevalensi stunting idealnya berada di bawah 20%. Suatu negara dikategorikan memiliki prevalensi stunting rendah jika angkanya kurang dari 20%. Sebaliknya, jika suatu negara memiliki prevalensi stunting yang lebih dari atau sama dengan 20%, maka prevalensi stunting di negara tersebut tergolong ke dalam prevalensi stunting tinggi. Dampak stunting yang besar terhadap kesehatan anak telah menjadikannya sebagai prioritas dalam agenda global dan nasional, termasuk tujuan Pembangunan Berkelanjutan (SDG) 2.2 dan RPJMN Indonesia 2020-2024, yang bertujuan untuk mengurangi prevalensi stunting menjadi 14% pada tahun 2024.

SDG adalah kerangka kerja pembangunan yang berlaku selama 15 tahun, mulai dari 2015 hingga 2030. SDG merupakan inisiatif pembangunan berkelanjutan yang menjadi pedoman bagi pembangunan dan negosiasi internasional (Pratama et al., 2020). Kerangka ini mencakup 17 tujuan global dengan 169 indikator yang telah ditetapkan oleh PBB. Salah satu tujuan SDG, yaitu target 2.2, adalah menghapus segala bentuk malnutrisi. Pada tahun

2030, SDG bertujuan untuk menghilangkan kekurangan gizi, termasuk mencapai target internasional pada tahun 2025 terkait stunting dan wasting pada anak di bawah usia 5 tahun, serta memastikan pemenuhan kebutuhan gizi bagi remaja perempuan, ibu hamil, ibu menyusui, dan lanjut usia. Berdasarkan World Population Review (2023), penerapan SDG sudah mampu menekan angka stunting di beberapa negara. Namun, tantangannya adalah Indonesia menduduki peringkat 115 dalam prevalensi stunting secara global pada 2022.

Stunting merupakan permasalahan kompleks yang disebabkan oleh faktor multifaset. Untuk melaksanakan intervensi yang efektif, penting untuk memahami berbagai faktor yang mempengaruhi prevalensi stunting. Beberapa di antaranya adalah kondisi bayi yang lahir dengan kekurangan nutrisi dan tumbuh dalam keadaan kekurangan gizi. Faktor lain yang turut berkontribusi meliputi sanitasi rendah, berat badan bayi lahir rendah, pemberian ASI eksklusif, riwayat penyakit infeksi serta faktor genetik (Kementerian Kesehatan RI, 2022). Selain itu, indikator yang diukur pada Survei Kesehatan Indonesia (SKI) 2023 meliputi status kesehatan, kesehatan lingkungan, serta pengetahuan dan perilaku kesehatan menjadi faktor-faktor pendukung. Identifikasi faktor-faktor ini penting untuk merancang intervensi yang efektif. Mengingat banyaknya variabel yang mempengaruhi stunting, analisis regresi tradisional mungkin mengalami kesulitan dalam menangani multikolinieritas pada variabel bebas. Oleh karena itu, untuk mengatasi masalah multikolinieritas maka digunakan metode *Principal Component Regression* (PCR). Metode ini bekerja dengan mereduksi dimensi data melalui *Principal Component Analysis* (PCA) dan memanfaatkan komponen utama tersebut dalam model regresi memprediksi prevalensi stunting.

Beberapa penelitian sebelumnya yang berkaitan meliputi penelitian oleh Susilawati & Didiharyono (2023) berjudul “*Application of Principal Component Regression in Analyzing Factors Affecting Human Development Index*”. Penelitian yang dilakukan Nurhadi (2018) mengenai “Metode Statistik *Principal Component Regression* dalam Analisis Hubungan Kuantitatif Struktur dan Aktivitas Penghambat Lipoksigenase Senyawa Turunan Kurkumin”. Kemudian penelitian lainnya adalah “Faktor-Faktor yang Mempengaruhi Stunting pada Balita di Kabupaten Grobogan” oleh Yuwanti et al. (2021), di mana menggunakan uji *Chi Square* dan Regresi Logistik.

Meskipun penelitian terkait stunting telah dilakukan, penggunaan PCR untuk mengidentifikasi faktor-faktor yang paling berpengaruh terhadap prevalensi stunting di Indonesia merupakan pendekatan yang lebih modern dan tepat. Metode ini tidak hanya mengatasi masalah multikolinieritas. Nantinya juga memberikan pemahaman yang lebih jelas tentang variabel mana yang paling signifikan dalam menjelaskan variasi prevalensi stunting. Ini akan membantu mengarahkan kebijakan yang lebih terfokus dan berbasis data dalam mengatasi masalah stunting.

Penelitian ini bertujuan untuk mengidentifikasi penyebab stunting sebagai informasi untuk intervensi kebijakan yang tepat sasaran. Adanya multikolinieritas di antara berbagai faktor yang mempengaruhi, sehingga model regresi tradisional sering kali gagal. *Principal Component Regression* (PCR) digunakan untuk mengatasi multikolinieritas dan meningkatkan daya prediksi model. Hasil penelitian ini diharapkan mampu memberikan rekomendasi kebijakan kesehatan masyarakat yang lebih efektif dan tepat sasaran untuk mengurangi prevalensi stunting di masa depan.

2. TINJAUAN PUSTAKA

Regresi adalah metode analisis statistik yang digunakan untuk menjelaskan hubungan antara dua jenis variabel, yaitu variabel terikat (y) dan variabel bebas (x). Regresi terbagi menjadi dua jenis utama diantaranya regresi linear sederhana dan regresi linear berganda. Regresi linear sederhana menganalisis hubungan antara satu variabel terikat dengan satu

variabel bebas, sedangkan regresi linear berganda melibatkan lebih dari satu variabel bebas dalam modelnya (Prasetyo & Helma, 2022). Selain digunakan untuk memahami pola hubungan, analisis regresi juga berguna dalam melakukan peramalan atau prediksi. Sebagai contoh, terdapat kumpulan data (x_i, y_i) yang secara umum dapat direpresentasikan menggunakan model regresi berikut (Sembiring, 1995):

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \varepsilon_i, i = 1, 2, \dots, n \quad (1)$$

dengan y adalah variabel terikat, x adalah variabel bebas, β adalah nilai koefisien regresi, dan ε adalah residual, yang diasumsikan IIDN.

Sebelum melakukan analisis regresi, data variabel diperlukan proses standarisasi melalui pemusatan dan penskalaan. Standarisasi ini mentransformasi variabel penelitian ke dalam bentuk (Kutner et al., 2005):

$$y_i^* = \frac{y_i - \bar{y}}{s_y}, i = 1, 2, \dots, n \quad (2)$$

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, i = 1, 2, \dots, n, j = 1, 2, \dots, k \quad (3)$$

dengan

$$s_y = \sqrt{\frac{\sum_i (y_i - \bar{y})^2}{n-1}} \quad \text{dan} \quad s_j = \sqrt{\frac{\sum_i (x_{ij} - \bar{x}_j)^2}{n-1}}$$

Dengan menggunakan variabel-variabel yang telah distandarisi, diperoleh bentuk umum dari model persamaan regresi linear berganda sebagai berikut (Montgomery et al., 2012):

$$y_i^* = b_1 Z_{i1} + b_2 Z_{i2} + \dots + b_k Z_{ik} + \varepsilon_i, i = 1, 2, \dots, n. \quad (4)$$

Namun, dalam regresi linier dapat terjadi multikolinearitas, yaitu korelasi linier antara variabel bebas yang dapat mengurangi keakuratan model regresi (Gujarati & Porter, 2009). Deteksi multikolinearitas dilakukan menggunakan *Variance Inflation Factor* (VIF):

$$VIF = \frac{1}{1-R^2} \quad (5)$$

dimana R^2 adalah koefisien determinasi. Jika nilai $VIF > 5$, maka variabel tersebut mempunyai persoalan multikolinearitas dengan variabel bebas lainnya (Santoso, 2001).

Metode yang digunakan untuk mengatasi multikolinieritas adalah *Principal Component Analysis* (PCA). Proses PCA melibatkan reduksi variabel-variabel bebas yang saling berkorelasi menjadi variabel baru yang disebut komponen utama, yang saling bebas. Komponen utama yang dihasilkan bersifat ortogonal dan tidak berkorelasi satu sama lain (Johnson & Wichern, 2007).

Selanjutnya, *Principal Component Regression* (PCR) menggabungkan pendekatan PCA dengan analisis regresi untuk menghilangkan multikolinieritas dan meningkatkan akurasi model. Menurut Mattjik & Sumertajaya (2011), langkah-langkah dalam PCR dimulai dengan menstandarisasi data variabel, menghitung nilai eigen (λ_j) dan vektor eigen (a_j), dan komponen utama (K_j). Komponen utama terpilih kemudian digunakan dalam analisis regresi terhadap variabel terikat, sehingga diperoleh model regresi komponen utama sebagai berikut:

$$y = w_0 + w_1 K_1 + w_2 K_2 + \dots + w_m K_m. \quad (6)$$

dimana w_0 adalah *intercept*, w_1, w_2, \dots, w_m adalah koefisien regresi, dan K_1, K_2, \dots, K_m adalah komponen utama. Komponen utama ini merupakan kombinasi linier dari variabel yang telah distandarisi seperti berikut.

$$\begin{aligned} K_1 &= a_1^t Z = a_{11} Z_1 + a_{12} Z_2 + \dots + a_{1k} Z_k \\ K_2 &= a_2^t Z = a_{21} Z_1 + a_{22} Z_2 + \dots + a_{2k} Z_k \\ &\vdots \\ K_m &= a_m^t Z = a_{m1} Z_1 + a_{m2} Z_2 + \dots + a_{mk} Z_k. \end{aligned} \quad (7)$$

Model regresi komponen utama dapat ditransformasi ke dalam model regresi berbasis variabel yang telah distandarisasi seperti berikut.

$$y = b_0 + b_1Z_1 + b_2Z_2 + \dots + b_kZ_k + \varepsilon. \quad (8)$$

Selanjutnya dapat ditransformasikan kembali menjadi model regresi berbasis variabel bebas asal seperti berikut:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k + \varepsilon. \quad (9)$$

Kemudian untuk mengetahui apakah hubungan yang ditemukan dalam model tersebut signifikan, dilakukan uji signifikansi parameter, baik secara simultan maupun parsial (Montgomery et al., 2012). Uji secara Simultan (bersama-sama) menggunakan uji F seperti berikut:

$$F_{hitung} = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE}. \quad (10)$$

Sedangkan uji secara parsial (sendiri) menggunakan uji t seperti berikut:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}. \quad (11)$$

Selain itu, model regresi juga harus memenuhi asumsi klasik, salah satunya adalah normalitas residual. Uji normalitas dilakukan untuk memastikan bahwa *error* mengikuti distribusi normal, salah satunya dengan uji Kolmogorov-Smirnov (Ghozali, 2016):

$$D = \text{Max}|F_0(\varepsilon_i) - S_N(\varepsilon_i)|. \quad (12)$$

Uji heteroskedastisitas dilakukan untuk memastikan bahwa varian *error* konstan. Salah satu metode yang digunakan adalah uji Glejser, dengan rumusan statistik uji sebagai berikut: (Ghozali, 2016):

$$t = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}. \quad (13)$$

Uji autokorelasi bertujuan untuk mengidentifikasi apakah terdapat korelasi antara *error* pada waktu yang berbeda (Kuncoro, 2007). Salah satu metode yang digunakan adalah Durbin-Watson (Draper & Smith, 1998):

$$d = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2}. \quad (14)$$

Proses akhir, untuk mengukur sejauh mana model dapat menjelaskan variabel terikat, dilakukan uji kebaikan model (*Goodness of Fit*). Nilai koefisien determinasi R^2 dihitung dengan rumus berikut (James et al., 2013):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (15)$$

dimana nilai R^2 berada dalam rentang 0 hingga 1, di mana semakin mendekati 1 menunjukkan model yang semakin baik dalam menjelaskan variabel terikat.

3. METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah faktor-faktor yang mempengaruhi prevalensi stunting di Indonesia tahun 2023. Sumber data diperoleh dari data sekunder yang dipublikasikan oleh Survei Kesehatan Indonesia (SKI). Unit penelitian terdiri dari 38 Provinsi di Indonesia. Penelitian ini menggunakan variabel terikat (y) adalah prevalensi stunting di Indonesia. Variabel bebas (x) adalah faktor-faktor yang mempengaruhi prevalensi stunting di Indonesia. Variabel-variabel yang digunakan terfokus pada indikator yang diukur pada SKI 2023. Berikut beberapa variabel tersebut disajikan dalam Tabel 1.

Tabel 1. Variabel Penelitian

Simbol	Variabel	Satuan
y	Prevalensi Stunting	Persen (%)
x_1	Asi Eksklusif	Persen (%)
x_2	BBLR (Berat Bayi Lahir Rendah)	Persen (%)

x_3	Asupan Makanan Ibu Kurang Gizi	Persen (%)
x_4	Asupan Makanan Anak Kurang Gizi	Persen (%)
x_5	Imunisasi Dasar Lengkap	Persen (%)
x_6	Sanitasi Layak	Persen (%)
x_7	Penyakit Keturunan	Persen (%)

Tabel 2. Struktur Data Penelitian

Provinsi di Indonesia	y	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	$y_{1(1)}$	$x_{1(1)}$	$x_{2(1)}$	$x_{3(1)}$	$x_{4(1)}$	$x_{5(1)}$	$x_{6(1)}$	$x_{7(1)}$
2	$y_{1(2)}$	$x_{1(2)}$	$x_{2(2)}$	$x_{3(2)}$	$x_{4(2)}$	$x_{5(2)}$	$x_{6(2)}$	$x_{7(2)}$
3	$y_{1(3)}$	$x_{1(3)}$	$x_{2(3)}$	$x_{3(3)}$	$x_{4(3)}$	$x_{5(3)}$	$x_{6(3)}$	$x_{7(3)}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
38	$y_{1(38)}$	$x_{1(38)}$	$x_{2(38)}$	$x_{3(38)}$	$x_{4(38)}$	$x_{5(38)}$	$x_{6(38)}$	$x_{7(38)}$

Penelitian ini menggunakan metode *Principal Component Regression* (PCR). Langkah-langkah analisis yang dilakukan sebagai berikut:

- Melakukan standarisasi data penelitian menggunakan pemusatan dan penskalaan.
- Mendeteksi adanya multikolinieritas, dan jika terdeteksi maka dilakukan metode PCR.
- Menentukan nilai eigen dari matriks korelasi.
- Menentukan skor komponen utama yang terbentuk.
- Melakukan regresi linier berganda antara variabel terikat dan komponen utama.
- Menguji signifikansi parameter model PCR secara parsial dan simultan.
- Mengubah persamaan regresi linier berganda ke dalam bentuk variabel standar.
- Mengubah kembali persamaan regresi linier berganda ke dalam bentuk variabel awal.
- Melakukan uji asumsi klasik regresi linier berganda, termasuk uji normalitas, uji heteroskedastisitas, dan uji autokorelasi.
- Melakukan uji kebaikan model regresi yang dihasilkan dari metode PCR dengan memeriksa nilai koefisien determinasi yang diperoleh.

4. HASIL DAN PEMBAHASAN

Langkah pertama adalah melakukan standarisasi data, karena data yang digunakan memiliki satuan yang berbeda. Proses standarisasi dilakukan dengan mengubah variabel awal menjadi variabel standar menggunakan nilai rata-rata dan simpangan baku. Hasil dari variabel-variabel yang telah di standarisasi menjadi simbol Y^* , Z_1 , Z_2 , Z_3 , Z_4 , Z_5 , Z_6 , Z_7 .

Multikolinieritas dapat diidentifikasi melalui nilai koefisien korelasi atau VIF. Hasil perhitungan uji multikolinieritas ditunjukkan pada Tabel 2 berikut. Tabel 2 menunjukkan bahwa variabel asi eksklusif (Z_1) dan asupan makanan anak kurang gizi (Z_4) memiliki nilai $VIF > 5$, yang mengindikasikan adanya masalah multikolinieritas dalam model regresi. Mengatasi masalah ini, maka digunakan metode *Principal Component Regression* (PCR). Pendekatan metode ini diterapkan jika terdapat minimal satu variabel bebas yang mengalami multikolinieritas sehingga menghasilkan model yang stabil dan akurat.

Tabel 2. Hasil Uji Multikolinieritas

Variabel	VIF	Keputusan
Z_1	13,170	Terjadi Multikolinieritas
Z_2	1,707	Tidak Terjadi Multikolinieritas
Z_3	3,301	Tidak Terjadi Multikolinieritas
Z_4	6,200	Terjadi Multikolinieritas
Z_5	1,902	Tidak Terjadi Multikolinieritas
Z_6	2,302	Tidak Terjadi Multikolinieritas
Z_7	4,507	Tidak Terjadi Multikolinieritas

Matriks korelasi menunjukkan tingkat hubungan atau korelasi antara variabel-variabel bebas. Hasil dari matriks korelasi dapat dilihat sebagai berikut.

$$R = \begin{bmatrix} 1,000 & -0,242 & 0,754 & 0,885 & -0,417 & -0,399 & 0,857 \\ -0,242 & 1,000 & -0,350 & -0,210 & 0,204 & -0,278 & -0,305 \\ 0,754 & -0,350 & 1,000 & 0,700 & -0,088 & -0,057 & 0,576 \\ 0,885 & -0,210 & 0,700 & 1,000 & -0,281 & -0,459 & 0,707 \\ -0,417 & 0,204 & -0,088 & -0,281 & 1,000 & 0,450 & -0,448 \\ -0,399 & -0,278 & -0,057 & -0,459 & 0,450 & 1,000 & -0,377 \\ 0,857 & -0,305 & 0,576 & 0,707 & -0,448 & -0,377 & 1,000 \end{bmatrix}$$

Korelasi yang terbentuk memiliki tingkat yang lemah hingga sangat kuat, dimana terdapat juga hubungan positif ataupun negatif. Kemudian untuk menentukan nilai eigen dapat menggunakan matriks korelasi. Hasil perhitungan determinan matriks akan memperoleh nilai eigen, dimana penjabarannya sebagai berikut:

$$|R - \lambda I| = 0 \tag{16}$$

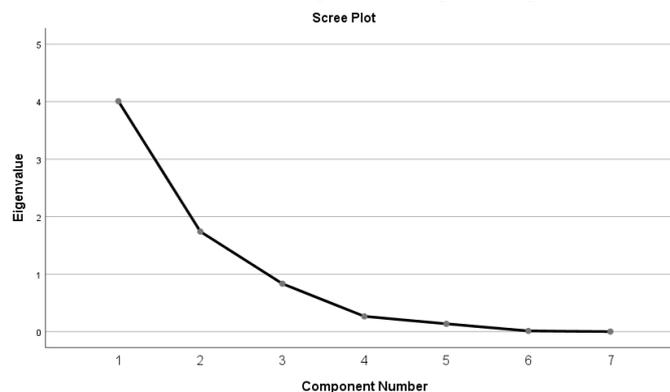
$$\begin{bmatrix} 1 - \lambda & -0,242 & 0,754 & 0,885 & -0,417 & -0,399 & 0,857 \\ -0,242 & 1 - \lambda & -0,350 & -0,210 & 0,204 & -0,278 & -0,305 \\ 0,754 & -0,350 & 1 - \lambda & 0,700 & -0,088 & -0,057 & 0,576 \\ 0,885 & -0,210 & 0,700 & 1 - \lambda & -0,281 & -0,459 & 0,707 \\ -0,417 & 0,204 & -0,088 & -0,281 & 1 - \lambda & 0,450 & -0,448 \\ -0,399 & -0,278 & -0,057 & -0,459 & 0,450 & 1 - \lambda & -0,377 \\ 0,857 & -0,305 & 0,576 & 0,707 & -0,448 & -0,377 & 1 - \lambda \end{bmatrix} = 0$$

Komponen utama yang terbentuk dilihat melalui *nilai eigen* > 1. Berdasarkan nilai eigen, diperoleh juga nilai keragaman total setiap komponen. Adapun nilai eigen dan keragaman total ditunjukkan pada Tabel 3 berikut.

Tabel 3. Nilai Eigen dan Keragaman Total

Komponen	Nilai Eigen	Keragaman Total
K_1	3,721	0,531
K_2	1,425	0,735
K_3	0,960	0,872
K_4	0,353	0,922
K_5	0,305	0,966
K_6	0,185	0,992
K_7	0,051	1,000

Tabel 3 menunjukkan bahwa 2 komponen yang dapat dibentuk dari tujuh variabel yang telah dianalisis yaitu 3,271 dan 1,425. Nilai eigen pertama sebesar 3,271 > 1 maka menjadi komponen 1 dan mampu menjelaskan 53,1% variasi. Sedangkan nilai eigen kedua sebesar 1,425 > 1 maka menjadi komponen 2 dan mampu menjelaskan 73,5% variasi. Selanjutnya, visualisasi dari komponen utama terbentuk dapat ditampilkan pada Gambar 1 berikut.



Gambar 1. Scree Plot Nilai Eigen

Skor komponen utama dimanfaatkan untuk menyusun persamaan *principal component regression* melalui perhitungan vektor eigen. Vektor eigen diperoleh dengan mensubstitusi nilai eigen dari kedua komponen utama.

$$(R - \lambda I)v_j = 0 \quad (17)$$

Untuk $\lambda = 3,271$

$$\begin{bmatrix} -2,271 & -0,242 & 0,754 & 0,885 & -0,417 & -0,399 & 0,857 \\ -0,242 & -2,271 & -0,350 & -0,210 & 0,204 & -0,278 & -0,305 \\ 0,754 & -0,350 & -2,271 & 0,700 & -0,088 & -0,057 & 0,576 \\ 0,885 & -0,210 & 0,700 & -2,271 & -0,281 & -0,459 & 0,707 \\ -0,417 & 0,204 & -0,088 & -0,281 & -2,271 & 0,450 & -0,448 \\ -0,399 & -0,278 & -0,057 & -0,459 & 0,450 & -2,271 & -0,377 \\ 0,857 & -0,305 & 0,576 & 0,707 & -0,448 & -0,377 & -2,271 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = 0$$

Untuk $\lambda = 1,425$

$$\begin{bmatrix} -0,425 & -0,242 & 0,754 & 0,885 & -0,417 & -0,399 & 0,857 \\ -0,242 & -0,425 & -0,350 & -0,210 & 0,204 & -0,278 & -0,305 \\ 0,754 & -0,350 & -0,425 & 0,700 & -0,088 & -0,057 & 0,576 \\ 0,885 & -0,210 & 0,700 & -0,425 & -0,281 & -0,459 & 0,707 \\ -0,417 & 0,204 & -0,088 & -0,281 & -0,425 & 0,450 & -0,448 \\ -0,399 & -0,278 & -0,057 & -0,459 & 0,450 & -0,425 & -0,377 \\ 0,857 & -0,305 & 0,576 & 0,707 & -0,448 & -0,377 & -0,425 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \end{bmatrix} = 0$$

Kemudian hasil perhitungan vektor eigen ditunjukkan pada Tabel 4 berikut.

Tabel 4. Skor Komponen Utama

Variabel	K_1	K_2
Z_1	0,963	0,025
Z_2	-0,342	-0,691
Z_3	0,758	0,424
Z_4	0,901	-0,006
Z_5	-0,517	0,387
Z_6	-0,483	0,785
Z_7	0,889	0,005

Tabel 4 menjelaskan bahwa skor komponen utama yang diperoleh merupakan hubungan variabel standar Z_j dengan komponen utama. Hubungan tersebut dapat dituliskan ke dalam persamaan (18) dan (19) berikut.

$$K_1 = 0,963Z_1 - 0,342Z_2 + 0,758Z_3 + 0,901Z_4 - 0,517Z_5 - 0,483Z_6 + 0,889Z_7 \quad (18)$$

$$K_2 = 0,025Z_1 - 0,691Z_2 + 0,424Z_3 - 0,006Z_4 + 0,387Z_5 + 0,785Z_6 + 0,005Z_7 \quad (19)$$

Berikutnya variabel baru K_1 dan K_2 akan digunakan untuk analisis regresi sehingga diperoleh estimasi parameter untuk konstanta sebesar $6,214e^{-16}$, K_1 adalah -0,244 dan K_2 adalah 0,770. Dengan demikian, model *principal component regression* yang diperoleh dapat dilihat pada persamaan (20) berikut:

$$y = 6,214e^{-16} - 0,244K_1 + 0,770K_2 \quad (20)$$

Berikut adalah hasil signifikansi parameter berdasarkan uji simultan dan parsial.

a. Uji Simultan

Uji simultan melalui uji F dengan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \exists \beta_i \neq 0, (i = 1, 2, 3, \dots, k)$$

Hasil perhitungan uji simultan ditunjukkan pada Tabel 5 berikut.

Tabel 5. Hasil Uji Simultan

Model	df	F_{hitung}	p value	Keputusan
Regression	2	32,879	0,000	Tolak H_0
Residual	35			
Total	37			

Tabel 5 diperoleh nilai $F_{0,05;2;35}$ sebesar 3,267 dengan $\alpha = 5\%$. Perbandingan nilai $F_{hitung} > F_{0,05;2;27}$ yaitu $32,879 > 3,267$ maka keputusannya tolak H_0 . Dengan demikian, dapat disimpulkan bahwa variabel bebas secara bersama-sama memiliki pengaruh signifikan terhadap variabel terikat pada tingkat signifikansi 5%.

b. Uji Parsial

Uji parsial melalui uji t dengan hipotesis sebagai berikut:

$H_0 : \beta_j = 0, j = 1, 2, 3, \dots, k$ (parameter ke- j tidak signifikan)

$H_1 : \beta_j \neq 0, j = 1, 2, 3, \dots, k$ (parameter ke- j signifikan)

Hasil perhitungan uji parsial ditunjukkan pada Tabel 6 berikut.

Tabel 6. Hasil Uji Parsial

Variabel	t_{hitung}	$ t_{hitung} $	p value	Keputusan	VIF
Konstanta	0,000	0,000	1,000	1,000	
K_1	-2,452	2,452	0,019	Signifikan	1,000
K_2	7,730	7,730	0,000	Signifikan	1,000

Tabel 6 diperoleh nilai $t_{0,05;35}$ sebesar 2,030 dengan $\alpha = 5\%$. Perbandingan nilai $|t_{hitung}| > t_{0,05;27}$ yaitu $2,452 > 2,030$ dan $7,730 > 2,030$ maka keputusannya tolak H_0 . Dengan demikian, dapat disimpulkan bahwa variabel bebas secara sendiri memiliki pengaruh signifikan terhadap variabel terikat pada tingkat signifikansi 5%. Selain itu, ditampilkan juga nilai VIF yang tidak terdapat multikolinieritas sehingga metode *principal component regression* dapat menghilangkan multikolinieritas antar variabel bebas dengan menghasilkan komponen utama yang saling bebas.

Mengembalikan model *principal component regression* ke persamaan yang memuat variabel standar seperti persamaan (21) berikut:

$$\begin{aligned}
 y &= 6,214e^{-16} - 0,244 (0,963Z_1 - 0,342Z_2 + 0,758Z_3 + 0,901Z_4 - 0,517Z_5 \\
 &\quad - 0,483Z_6 + 0,889Z_7) + 0,770 (0,025Z_1 - 0,691Z_2 + 0,424Z_3 - 0,006Z_4 \\
 &\quad + 0,387Z_5 + 0,785Z_6 + 0,005Z_7) \\
 y &= 6,214e^{-16} - 0,2157Z_1 - 0,4487Z_2 + 0,1415Z_3 - 0,2244Z_4 + 0,1307Z_5 + \\
 &\quad 0,7224Z_6 - 0,2130Z_7
 \end{aligned} \tag{21}$$

Mengembalikan persamaan ke dalam variabel awal yaitu x seperti persamaan (22) berikut:

$$\begin{aligned}
 y &= 6,214e^{-16} - 0,215 \left(\frac{x_1 - 56,2026}{13,28707} \right) - 0,4487 \left(\frac{x_2 - 5,8947}{1,35226} \right) \\
 &\quad + 0,1415 \left(\frac{x_3 - 88,5500}{3,93905} \right) - 0,2244 \left(\frac{x_4 - 52,9079}{9,39069} \right) + 0,1307 \left(\frac{x_5 - 30,6158}{15,14102} \right) \\
 &\quad + 0,7224 \left(\frac{x_6 - 82,1474}{14,00044} \right) - 0,2130 \left(\frac{x_7 - 31,4684}{8,38132} \right) \\
 y &= 6,214e^{-16} - 0,0162x_1 - 0,3318x_2 + 0,0359x_3 - 0,0239x_4 + 0,0086x_5 + \\
 &\quad 0,0516x_6 - 0,0254x_7
 \end{aligned} \tag{22}$$

Berdasarkan teori dan logika dasar pada model yang diperoleh telah memenuhi asumsi analisis regresi. Namun, interpretasi terhadap persamaan (22) tidak dapat menjelaskan dalam konteks teori kesehatan masyarakat. Hanya beberapa variabel yang dapat diinterpretasikan,

yaitu variabel asi eksklusif (x_1) dan asupan makanan ibu kurang gizi (x_4). Penambahan satu persen variabel x_1 dan x_4 maka menyebabkan penurunan prevalensi stunting di Indonesia. Dengan demikian, diperlukan upaya peningkatan terhadap variabel tersebut.

Pengujian terhadap *error* melalui beberapa uji asumsi klasik sebagai berikut:

Uji normalitas dalam penelitian ini dilakukan menggunakan metode *Kolmogorov-Smirnov*, dengan hipotesis pengujian sebagai berikut:

H_0 : *Error* berdistribusi normal ($\mu = 0$)

H_1 : *Error* tidak berdistribusi normal ($\mu \neq 0$)

Hasil perhitungan uji normalitas diperoleh nilai $D_{tabel} = 0,221$ dengan $\alpha = 0,05$. Perbandingan nilai $D < D_{tabel}$ yaitu $0,115 < 0,221$ maka keputusannya gagal tolak H_0 . Selain itu, nilai *p value* sebesar 0,200 dimana lebih besar dari 0,05 sehingga menyebabkan keputusan gagal tolak H_0 . Dengan demikian, kesimpulannya adalah *error* berdistribusi normal.

Uji heteroskedastisitas dilakukan menggunakan metode *Glejser*, dengan hipotesis pengujian sebagai berikut:

H_0 : Tidak terjadi heteroskedastisitas

H_1 : Terjadi heteroskedastisitas

Hasil perhitungan uji heteroskedastisitas diperoleh nilai $t_{0,05;35} = 2,030$ dengan $\alpha = 0,05$. Perbandingan nilai $t_{hitung} < t_{0,05;35}$ yaitu $0,272 < 2,030$ dan $0,196 < 2,030$ maka keputusannya gagal tolak H_0 . Selain itu, nilai *p value* sebesar 0,787 dan 0,846 dimana lebih besar dari 0,05 sehingga menyebabkan keputusan gagal tolak H_0 . Dengan demikian, kesimpulannya adalah tidak terjadi heteroskedastisitas atau varians *error* metode yang digunakan konsisten.

Uji autokorelasi dilakukan menggunakan metode *Durbin-Watson*, dengan hipotesis pengujian sebagai berikut:

H_0 : $\rho_1 = \rho_2 = \dots = \rho_n = 0$ (Tidak terdapat autokorelasi)

H_1 : $\rho_i \neq 0, i = 1, 2, 3, \dots, n$ (Terdapat autokorelasi)

Hasil perhitungan uji autokorelasi diperoleh nilai $dU < d < 4 - dU$ yaitu $0,913 < 1,402 < 3,087$ sehingga keputusannya gagal tolak H_0 , yang memberikan kesimpulan bahwa tidak terdapat autokorelasi.

Nilai R diperoleh adalah 0,653, yang menunjukkan bahwa tingkat hubungan (korelasi) antara variabel cukup kuat. Kebaikan model regresi dijelaskan melalui nilai *R Square* dan *Adjusted R Square*. Nilai *R Square* sebesar 0,653 atau 65,3% menunjukkan bahwa variabilitas pada variabel terikat dapat dijelaskan oleh model PCR yang digunakan. Begitu juga, *Adjusted R Square* sebesar 0,633 menunjukkan bahwa 63,3% variasi pada variabel terikat dapat dijelaskan oleh variasi variabel bebas, sementara sisanya sebesar 36,7% dipengaruhi oleh faktor-faktor lain yang tidak diteliti. Nilai *Std. Error of the Estimate* (SEE) sebesar 0,60598200, yang berarti semakin kecil nilai SEE, maka model regresi akan semakin tepat dalam memprediksi.

5. KESIMPULAN

Berdasarkan analisis data diperoleh beberapa kesimpulan, yaitu komponen 1 dengan nilai eigen 3,271, menjelaskan 53,1% dari varians, sedangkan Komponen 2 dengan nilai eigen 1,425, menjelaskan 73,5% dari varians. Kedua komponen tersebut secara signifikan mempengaruhi prevalensi stunting, seperti yang ditunjukkan oleh uji simultan dan parsial. Model regresi menggunakan *Principal Component Regression* (PCR) pada data prevalensi stunting di Indonesia seperti Persamaan (22). Model memiliki nilai kebaikannya sebesar 0,633, yang mengindikasikan bahwa 63,3% variasi stunting dapat dijelaskan oleh model, menyisakan 36,7% yang disebabkan oleh faktor eksternal. Metode PCR secara efektif

mengatasi multikolinieritas, yang dibuktikan dengan nilai *Variance Inflation Factor* (*VIF*) < 5. Studi ini menyoroti pentingnya penggunaan metode statistik canggih untuk pengembangan kebijakan yang kuat dalam memerangi stunting.

DAFTAR PUSTAKA

- Dinas Kesehatan Papua. (2024). *Menkes Budi Soroti Lambatnya Penurunan Angka Stunting di Indonesia*. <https://dinkes.papua.go.id/menkes-budi-soroti-lambatnya-penurunan-angka-stunting-di-indonesia>
- Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis*. In *Wiley Series in Probability and Statistics* (1st ed.). Wiley. <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118625590>
- Ghozali, I. (2016). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 23* (8th ed.). Univ. Diponegoro Press.
- Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics*. In *The McGraw-Hill series Economics* (5. ed). McGraw-Hill Irwin.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. In *Springer texts in statistics*. Springer.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6. ed). Pearson/Prentice Hall.
- Kementerian Kesehatan RI. (2022). *Faktor-faktor Penyebab Kejadian Stunting pada Balita*. https://yankes.kemkes.go.id/view_artikel/1529/faktor-faktorpenyebab-kejadian-stunting-pada-balita
- Kementerian Kesehatan RI. (2023). *Prevalensi Stunting di Indonesia Turun ke 21,6% dari 24,4%*. <https://sehatnegeriku.kemkes.go.id/baca/rilis-media/20230125/3142280/prevalensi-stunting-di-indonesia-turun-ke-216-dari-244>
- Kuncoro, M. (2007). *Metode Kuantitatif: Teori dan Aplikasi untuk Bisnis dan Ekonomi*. UPP STIM YKPN.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., Li, W., & Nachtsheim, C. (2005). *Applied Linear Statistical Models*. In *McGraw-Hill/Irwin series Operations and decision sciences* (5. ed., internat. ed). McGraw-Hill.
- Mattjik, A. A., & Sumertajaya, I. M. (2011). *Sidik Peubah Ganda*. IPB PRESS.
- Mayasari, D., Indriyani, R., & Sutarto. (2018). Stunting, Faktor Resiko dan Pencegahannya. *Jurnal Agromedicine*, 5(1), 540–545.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. In *Wiley series in probability and statistics* (5th ed). Wiley.
- Nurhadi, D. (2018). Metode Statistik Principal Component Regression dalam Analisis Hubungan Kuantitatif Struktur dan Aktivitas Penghambat Lipoksigenase Senyawa Turunan Kurkumin. *Teknologi: Jurnal Ilmiah Dan Teknologi*, 1(1), 60–71. <https://doi.org/10.32493/teknologi.v1i1.1418>
- Prasetyo, R. A., & Helma. (2022). Analisis Regresi Linear Berganda Untuk Melihat Faktor Yang Berpengaruh Terhadap Kemiskinan Di Provinsi Sumatera Barat. *Journal of Mathematics UNP*, 7(2), 62–68. <https://doi.org/10.24036/unpjomath.v7i2.12777>
- Pratama, N. B., Purnomo, E. P., & Agustiyara, A. (2020). Sustainable Development Goals (SDGs) dan Pengentasan Kemiskinan di Daerah Istimewa Yogyakarta. *SOSIOHUMANIORA: Jurnal Ilmiah Ilmu Sosial Dan Humaniora*, 6(2), 64–74. <https://doi.org/10.30738/sosio.v6i2.8045>
- Santoso, S. (2001). *Mengolah Data Statistik Secara Profesional*. PT. Elex Media Komputindo. Jakarta.

- Sembiring, R. K. (1995). *Analisis Regresi*. ITB Press.
- Susilawati, S., & Didiharyono, D. (2023). Application of Principal Component Regression in Analyzing Factors Affecting Human Development Index. *Jurnal Varian*, 6(2), 199–208. <https://doi.org/10.30812/varian.v6i2.2366>
- World Population Review. (2023). *Indonesia Population 2022*. <https://worldpopulationreview.com>
- Yuwanti, Y., Mulyaningrum, F. M., & Susanti, M. M. (2021). Faktor-Faktor yang Mempengaruhi Stunting pada Balita di Kabupaten Grobogan. *Jurnal Keperawatan Dan Kesehatan Masyarakat Cendekia Utama*, 10(1), 74–84. <https://doi.org/10.31596/jcu.v10i1.704>