

Analisis Jumlah Anak yang Pernah Dilahirkan dan Faktor yang Memengaruhinya pada Wanita Menikah di Indonesia Menggunakan Regresi *Generalized Poisson*

Wulan Bella Juliantika *, Abdul Kudus

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

wulanbellaj@gmail.com, abdul.kudus@unisba.ac.id

Abstract. Population growth is a significant global issue, including in Indonesia, and one of its main indicators is the fertility rates, which can be measured by the number of children ever born. As a count variable, it is commonly analyzed using the *Poisson* regression model. However, *Poisson* regression assumes equidispersion (mean equals variance), which is often violated due to overdispersion or underdispersion. In this study, the data showed underdispersion, so an alternative method, *Generalized Poisson* Regression (GPR), is used for its flexibility in handling dispersion mismatch. This study analyzes factors influencing the number of children ever born to married women in Indonesia using *Generalized Poisson* Regression. The data used are secondary data from the 2017 Indonesian Demographic and Health Survey (IDHS) with a total of 34,086 married women. The variables used include the respondent's age, province, type of residence, education level, gender of the head of household, household wealth index, knowledge of contraception, and use of contraception. Based on the estimation results, it was concluded that the age of women, most provinces, type of residence, educational level, gender of the household head, household wealth index, knowledge of contraception, and use of contraception significantly influence the number of children ever born.

Keywords: *Fertility, Children Ever Born, Generalized Poisson Regression.*

Abstrak. Pertumbuhan penduduk merupakan isu global yang signifikan, termasuk di Indonesia, dan salah satu indikator utamanya adalah tingkat kelahiran atau fertilitas. Fertilitas dapat diukur melalui jumlah anak yang pernah dilahirkan (*children ever born*). Data jumlah anak yang pernah dilahirkan berupa diskrit sehingga umumnya dianalisis menggunakan model regresi *Poisson*. Namun, model *Poisson* memiliki asumsi dasar equidispersi (nilai *mean* sama dengan varians), yang dalam praktiknya sering tidak terpenuhi karena adanya overdispersi atau underdispersi. Pada penelitian ini, data menunjukkan adanya underdispersi, sehingga digunakan *Generalized Poisson* Regression (GPR) sebagai metode alternatif yang lebih fleksibel dalam menangani ketidaksesuaian dispersi. Penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi jumlah anak yang pernah dilahirkan oleh wanita menikah di Indonesia dengan menggunakan metode regresi *Generalized Poisson*. Data yang digunakan merupakan data sekunder dari Indonesian Demographic and Health Survey (IDHS) tahun 2017 dengan total 34.086 wanita menikah. Variabel yang digunakan meliputi usia responden, provinsi, jenis tempat tinggal, tingkat pendidikan, jenis kelamin kepala rumah tangga, indeks kekayaan rumah tangga, pengetahuan kontrasepsi, dan penggunaan kontrasepsi. Berdasarkan hasil estimasi, disimpulkan bahwa usia wanita, sebagian besar provinsi, jenis tempat tinggal, tingkat pendidikan, jenis kelamin kepala rumah tangga, indeks kekayaan rumah tangga, serta pengetahuan dan penggunaan kontrasepsi secara signifikan memengaruhi jumlah anak yang pernah dilahirkan.

Kata Kunci: *Fertilitas, Anak yang Pernah Dilahirkan, Regresi Generalized Poisson.*

A. Pendahuluan

Distribusi *Poisson* digunakan ketika suatu peristiwa bersifat jarang terjadi dengan setiap kejadian di suatu peristiwa diasumsikan bersifat saling bebas (Walpole dkk., 2012). Distribusi *Poisson* merupakan standar untuk data berupa *count* dan semua model *count* lainnya merupakan penyesuaian atau variasi dari model *Poisson* dasar. Beberapa asumsi dalam *Poisson*, diantaranya variabel dependen berupa bilangan cacah (*count data*) yang nilainya non-negatif, nilai nol pada variabel respons memungkinkan untuk terjadi, observasi bersifat independen satu sama lain, serta nilai *mean* dan varians pada variabel respons bernilai identik (Hilbe, 2014). Karakteristik penting dari distribusi *Poisson* adalah nilai *mean* dan variansnya sama, yaitu equidisersi.

$$E(Y) = \text{Var}(Y) = \lambda \quad (1)$$

Namun, sering ditemukan kondisi yang menunjukkan bahwa asumsi equidisersi ini tidak terpenuhi, seperti varians data lebih kecil dari rata-rata datanya yang disebut underdispersi atau bahkan varians data lebih besar dari rata-rata datanya yang disebut overdispersi (Hilbe, 2014). Oleh karena itu, uji equidisersi penting untuk mengidentifikasi apakah data memenuhi asumsi equidisersi atau tidak karena ketidaksesuaian antara varians dan nilai harapan dapat memengaruhi akurasi estimasi parameter dan uji signifikansi dalam regresi *Poisson*. Jika asumsi equidisersi tidak terpenuhi maka digunakan metode lain untuk mengatasi masalah tersebut, seperti binomial negatif, *Generalized Poisson*, dan metode lain yang sesuai dengan masalah dispersi yang terjadi.

Regresi *Generalized Poisson* (*Generalized Poisson Regression/GPR*) merupakan salah satu perluasan dari model regresi *Poisson* yang dirancang untuk mengatasi permasalahan overdispersi dan underdispersi dalam data *count* (cacahan). Regresi *Generalized Poisson* memiliki perbedaan dengan regresi *Poisson*, yaitu pada parameternya dimana tidak hanya memiliki parameter λ , tetapi juga memiliki parameter dispersi yang dinotasikan sebagai θ . Jika parameter $\theta > 0$ maka menunjukkan overdispersi dan jika parameter $\theta < 0$ maka menunjukkan underdispersi (Hilbe, 2014). Regresi *Generalized Poisson* memiliki rata-rata (*mean*), yaitu

$$E(y_i) = \lambda_i = e^{x_i^T \beta} \quad (2)$$

dan varians yang berbeda dari *Poisson* biasa, yaitu

$$\text{Var}(y_i) = \lambda_i(1 + \theta\lambda_i)^2 \quad (3)$$

Pertumbuhan penduduk merupakan salah satu isu global yang menjadi pusat perhatian di dunia karena pertumbuhannya yang diproyeksikan akan terus meningkat. Berdasarkan United Nations (2024), populasi dunia di tahun 2022 mencapai lebih dari 8,2 miliar jiwa dan diperkirakan akan terus tumbuh selama 50 hingga 60 tahun ke depan hingga mencapai 10,3 miliar jiwa di pertengahan tahun 2080. Indonesia, sebagai negara dengan populasi terbanyak keempat di dunia, mencatat jumlah penduduk sekitar 281,6 juta jiwa pada tahun 2024. Jumlah ini diprediksi terus meningkat, mencapai sekitar 298 juta jiwa pada tahun 2030 dan sekitar 337 juta jiwa pada tahun 2050. Pertumbuhan penduduk ini dipengaruhi oleh berbagai faktor, salah satunya adalah tingkat kelahiran (fertilitas). Fertilitas adalah kemampuan wanita melahirkan yang diukur dari jumlah anak hidup yang dilahirkan selama masa subur (usia 15–49 tahun). Jumlah ini dipengaruhi oleh berbagai faktor, seperti usia menikah, pendidikan, ekonomi, akses kesehatan, serta norma dan nilai sosial (Badan Pusat Statistik, 2011). Perubahan demografis ini menjadi tantangan bagi negara Indonesia untuk merancang perencanaan pembangunan nasional, terutama dalam pengendalian kelahiran anak. Berdasarkan Badan Perencanaan dan Pembangunan Nasional (2023) rata-rata angka fertilitas nasional pada tahun 2020 sebesar 2,18 meskipun begitu fenomena ketimpangan fertilitas antar provinsi terlihat jelas, seperti di Jakarta dengan angka fertilitas sebesar 1,75 dengan Papua dengan angka fertilitas sebesar 2,76.

Di tingkat global, permasalahan terkait demografi pun masih mejadi isu penting di negaranya masing-masing, terutama negara berkembang sehingga ada beberapa penelitian-penelitian serupa yang mengkaji terkait faktor-faktor yang memengaruhi jumlah anak yang pernah dilahirkan oleh wanita menikah. Berdasarkan penelitian yang dilakukan oleh Gebre (2024) dan Cherie dkk. (2023) dengan studi kasus negara Ethiopia menunjukkan bahwa usia seorang wanita saat melahirkan anak pertama, agama seorang wanita, tingkat pendidikan seorang wanita, indeks kekayaan rumah tangga, penggunaan kontrasepsi, daerah tempat tinggal, dan status pekerjaan wanita secara independen memengaruhi jumlah anak yang pernah lahir di antara wanita Ethiopia yang saat ini menikah. Penelitian serupa lainnya di negara Oman yang menunjukkan bahwa usia ibu, tingkat pendidikan, tempat tinggal, dan penggunaan kontrasepsi pada ibu menunjukkan bahwa faktor-faktor tersebut berpengaruh secara signifikan (Al-Balushi dkk., 2020). Penelitian-penelitian yang dikaji tersebut menggunakan regresi, seperti regresi *Poisson*, regresi binomial negatif, regresi logistik, dan sebagainya.

Berdasarkan uraian fenomena dan kajian pustaka, penelitian ini bertujuan untuk menganalisis faktor-faktor yang memengaruhi jumlah anak yang pernah dilahirkan oleh wanita menikah di Indonesia dengan menggunakan metode Regresi *Generalized Poisson* (GPR). Tujuan ini dilandasi oleh kebutuhan akan model statistik yang sesuai untuk data cacah yang menunjukkan gejala underdispersi, di mana model regresi *Poisson* biasa tidak lagi layak digunakan. Selain itu, penelitian ini juga dapat dimanfaatkan bagi pihak-pihak terkait, khususnya pemerintah dan lembaga perencanaan kebijakan kependudukan. Hasil penelitian ini diharapkan dapat menjadi bahan pertimbangan dalam merumuskan kebijakan pengendalian kependudukan yang lebih efektif dan tepat sasaran, karena didasarkan pada faktor-faktor yang terbukti secara signifikan berpengaruh terhadap tingkat kelahiran.

B. Metode

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari situs resmi *The Demographic and Health Surveys (DHS)*. Survei ini merupakan hasil kerja sama antara Badan Pusat Statistik (BPS), Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN), dan Kementerian Kesehatan Indonesia, dengan bantuan pendanaan dari Pemerintah Indonesia. Pengumpulan data dilakukan pada periode 24 Juli hingga 30 September 2017, mencakup 1.970 blok sensus di wilayah perkotaan dan perdesaan, dengan target 49.250 rumah tangga dan estimasi 59.100 perempuan usia 15–49 tahun sebagai responden. Hasil dari survei tersebut menghasilkan unit penelitian sebesar 49.627 wanita Indonesia yang berusia 15–49 tahun dan sesuai dengan tujuan penelitian ini maka hanya digunakan sebanyak 34.086 data responden dengan kriteria wanita menikah. Variable yang digunakan pada penelitian ini terdiri dari satu variable respons (Y) dan delapan variable predictor (X) yang disajikan pada table berikut.

Tabel 1. Variabel Penelitian

Variabel	Nama Variabel	Jenis Data
Y	Jumlah Anak yang Pernah Dilahirkan	Numerik
X_1	Usia Responden	Numerik
X_2	Provinsi	Kategorik dengan 34 provinsi
X_3	Jenis Tempat Tinggal	Kategorik dengan 1: Perkotaan dan 2: Pedesaan
X_4	Tingkat Pendidikan	Kategorik dengan 1: Tidak Berpendidikan, 2: Pendidikan Dasar, 3: Pendidikan Menengah, dan 4: Pendidikan Tinggi
X_5	Jenis Kelamin Kepala Rumah Tangga	Kategorik dengan 1: Laki-laki dan 2: Perempuan
X_6	Indeks Kekayaan Rumah Tangga	Kategorik dengan 1: Sangat Miskin, 2: Miskin, 3: Menengah, 4: Kaya, dan 5: Sangat Kaya
X_7	Pengetahuan terkait Kontrasepsi	Kategorik dengan 1: Tidak Tahu Metode Kontrasepsi, 2: Hanya Tahu Metode Kontrasepsi <i>Folklore</i> , 3: Hanya Tahu Metode Kontrasepsi Tradisional, dan 4: Tahu Metode Kontrasepsi Modern
X_8	Penggunaan Kontrasepsi	Kategorik dengan 1: Tidak Menggunakan Kontrasepsi, 2: Menggunakan Kontrasepsi <i>Folklore</i> , 3: Menggunakan Kontrasepsi Tradisional, dan 4: Menggunakan Kontrasepsi Modern

Regresi *Poisson*

Poisson merupakan standar untuk data berupa *count* dan semua model *count* lainnya merupakan penyesuaian atau variasi dari model *Poisson* dasar. Beberapa asumsi dalam *Poisson*, diantaranya variabel dependen berupa bilangan cacah (*count data*) yang nilainya non-negatif, nilai nol pada variabel respons memungkinkan untuk terjadi, observasi bersifat independen satu sama lain, serta nilai *mean* dan varians pada variabel respons bernilai identik (Hilbe, 2014). Jadi, regresi *Poisson* merupakan salah satu bentuk model yang digunakan untuk menganalisis pengaruh antara variabel respons berupa *count data* dengan satu atau lebih variabel prediktor. Dalam hal ini, nilai rata-rata dari variabel respons $\bar{y}_i$ dinyatakan sebagai λ_i , yaitu rata-rata jumlah kejadian yang diprediksi oleh model untuk pengamatan ke- i . Model regresi *Poisson* menggunakan fungsi *link* log yang dinyatakan sebagai berikut.

$$\log(\lambda_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} = x_i^T \beta \quad (4)$$

atau dalam bentuk eksponensialnya sebagai berikut.

$$\lambda_i = e^{\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki}} = e^{x_i^T \beta} \quad (5)$$

Pendugaan parameter pada regresi *Poisson* dilakukan dengan menggunakan metode MLE dan iterasi Newton Raphson.

Equidispersi

Regresi *Poisson* memiliki kriteria penting, yaitu nilai mean dan variansnya sama yang dinyatakan dengan $E(Y) = \text{Var}(Y) = \lambda$ atau kondisi tersebut dinamakan equidispersi. Namun, sering ditemukan kondisi yang menunjukkan bahwa asumsi ini tidak terpenuhi, seperti varians data lebih kecil dari rata-rata datanya yang disebut underdispersi atau bahkan varians data lebih besar dari rata-rata datanya yang disebut overdispersi. Statistik dispersi dari model *Poisson* didefinisikan sebagai statistik Pearson X^2 dibagi dengan derajat kebebasan residual. Dimana derajat kebebasan residual adalah jumlah pengamatan dalam model dikurangi jumlah prediktor, termasuk intersep (Hilbe, 2014). Statistik dispersi tersebut dinyatakan dengan persamaan sebagai berikut.

$$\theta = \frac{X^2}{db} \quad (6)$$

Dimana

$$X^2 = \sum_{i=1}^n \frac{(y_i - \lambda_i)^2}{\lambda_i} \quad (7)$$

Hasilnya dari statistik dispersinya seharusnya memiliki nilai 1 yang menunjukkan model equidispersi. Namun, jika nilai θ lebih besar dari 1 maka menunjukkan model yang overdispersi dan jika nilai θ kurang dari 1 menunjukkan model yang underdispersi.

Regresi *Generalized Poisson*

Regresi *Generalized Poisson* memiliki perbedaan dengan regresi *Poisson*, yaitu pada parameternya dimana tidak hanya memiliki parameter λ , tetapi juga memiliki parameter kedua sebagai parameter dispersi yang dinotasikan dengan θ . Sama seperti model yang diperkenalkan sebelumnya, *Generalized Poisson* menjadi *Poisson* ketika dispersi sama dengan nol. Namun, fitur menarik dari *Generalized Poisson* adalah bahwa parameter dispersi dapat memiliki nilai negatif, yang menunjukkan penyesuaian untuk *Poisson* underdispersi. Fungsi probabilitas massanya dinyatakan sebagai berikut (Yang dkk., 2009).

$$f(y_i; \lambda_i, \theta) = \left(\frac{\lambda_i}{1 + \theta \lambda_i} \right)^{y_i} \cdot \frac{(1 + \theta y_i)^{y_i - 1}}{y_i!} \cdot \exp \left(\frac{-\lambda_i (1 + \theta y_i)}{1 + \theta \lambda_i} \right) \quad (8)$$

Jika parameter $\theta > 0$ maka menunjukkan overdispersi dan jika parameter $\theta < 0$ maka menunjukkan underdispersi (Hilbe, 2014). Regresi *Generalized Poisson* memiliki rata-rata (*mean*), yaitu $E(y_i) = \lambda_i = e^{x_i^T \beta}$ dan varians yang berbeda dari *Poisson* biasa, yaitu $Var(y_i) = \lambda_i(1 + \theta \lambda_i)^2$. Pendugaan parameter pada regresi *generalized Poisson* dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* dan iterasi Newton Raphson.

Pengujian Signifikansi Parameter

Pada analisis regresi, terdapat dua uji hipotesis, yaitu uji simultan dan uji parsial. Uji simultan atau uji secara bersamaan dilakukan untuk menguji pengaruh signifikansi dari setiap variabel independen secara bersama-sama memiliki terhadap variabel dependen. Uji simultan yang umum digunakan, yaitu *Likelihood Ratio Test* dengan hipotesis sebagai berikut.

$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$; tidak ada variabel independen yang berpengaruh secara signifikan terhadap variabel dependen

H_1 : Paling tidak ada satu $\beta_k \neq 0$ dengan $k = 1, 2, 3, \dots, k$ yang berpengaruh secara signifikan terhadap variabel dependen.

Statistik uji yang digunakan sebagai berikut.

$$G = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = 2 \left(\ln L(\hat{\Omega}) - \ln L(\hat{\omega}) \right) \quad (9)$$

Fungsi *likelihood* $L(\hat{\Omega})$ merupakan nilai maksimum *likelihood* untuk model lengkap yang mencakup seluruh variabel prediktor sedangkan $L(\hat{\omega})$ merupakan nilai maksimum *likelihood* untuk model sederhana tanpa mencakup variabel prediktor. Kriteria pengujiannya adalah tolak H_0 jika nilai statistik *likelihood ratio test* lebih besar daripada nilai tabel distribusi *Chi-Square* pada tingkat signifikansi α dengan derajat bebas $(p + 1) - 1$ atau dapat dinyatakan sebagai $G \geq \chi_{(p+1)-1, \alpha}^2$. Selain itu, tolak H_0 juga dapat ditentukan melalui nilai *P-Value* $< \alpha$.

Di sisi lain, uji parsial pada analisis regresi dilakukan untuk menguji pengaruh signifikansi dari setiap variabel independen terhadap variabel dependen secara individu. Uji parsial yang umum digunakan, yaitu uji Wald dengan hipotesis sebagai berikut.

$H_0 : \beta_k = 0$; variabel independen ke- k tidak berpengaruh secara signifikan terhadap variabel dependen

$H_1 : \beta_k \neq 0$ dengan $k = 1, 2, 3, \dots, k$; variabel independen ke- k berpengaruh secara signifikan terhadap variabel dependen

Statistik uji yang digunakan sebagai berikut.

$$W = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (10)$$

Fungsi $\hat{\beta}_k$ menunjukkan taksiran nilai estimasi dari β_k sedangkan $SE(\hat{\beta}_k)$ menunjukkan standar error dari parameter β_k . Kriteria pengujiannya adalah tolak H_0 jika nilai statistik Wald lebih besar daripada nilai tabel distribusi *Chi-Square* pada tingkat signifikansi α dengan derajat bebas 1 atau dapat dinyatakan sebagai $W \geq \chi_1^2$. Selain itu, tolak H_0 juga dapat ditentukan melalui nilai *P-Value* $< \alpha$.

Model regresi *Poisson* dapat diinterpretasikan sebagai rasio angka kejadian (*Incidence Rate Ratio/IRR*), yaitu membandingkan laju kejadian suatu peristiwa antar dua kelompok dalam periode waktu tertentu. Interpretasi koefisien regresi *Poisson* pada dasarnya adalah selisih antara logaritma ekspektasi frekuensi kejadian, sehingga koefisien regresinya juga dapat dianggap sebagai logaritma dari rasio ekspektasi frekuensi antar kelompok (Hajarisman, 2022).

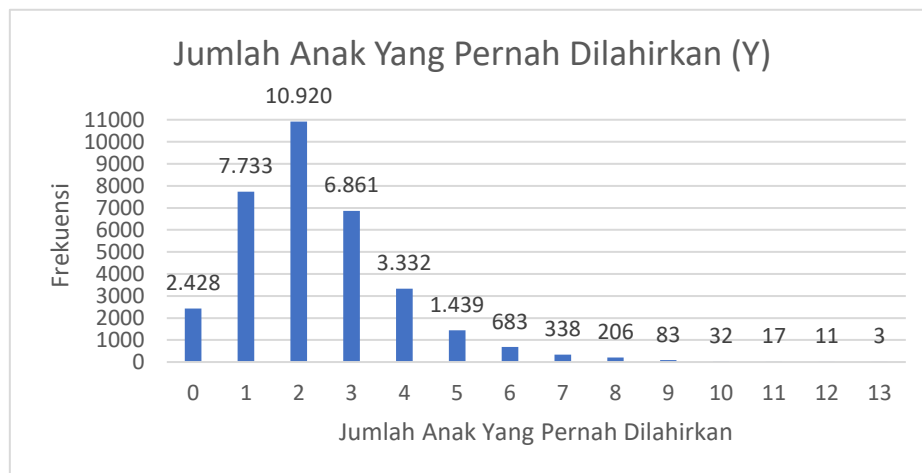
Dalam konteks regresi *Poisson* untuk rasio angka kejadian dapat dengan mengeksponensialkan koefisien regresi β yang diperoleh menggunakan dari model dengan persamaan sebagai berikut.

$$IRR = e^{\beta} \quad (11)$$

C. Hasil Penelitian dan Pembahasan

Eksplorasi Data

Sebelum dilakukan pemodelan, terlebih dahulu dilakukan eksplorasi terhadap data yang digunakan dalam penelitian ini, seperti histogram untuk variable Y sebagai berikut.



Gambar 1. Histogram Jumlah Anak Yang Pernah Dilahirkan Pada Wanita Menikah Tahun 2017.

Berdasarkan Gambar 1, dapat dilihat bahwa distribusi jumlah anak yang pernah dilahirkan (Y) menunjukkan pola data cacah dengan nilai terbanyak berada pada 2 anak, yaitu sebanyak 10.920 responden dan jumlah responden dengan anak lebih dari 5 semakin sedikit, menunjukkan distribusi yang condong ke kanan yang sesuai dengan karakteristik data regresi *Poisson* karena variabel dependen berupa data cacah non-negatif dengan frekuensi tinggi pada angka rendah dan semakin jarang pada angka yang lebih besar. Selanjutnya dilakukan eksplorasi data secara deskriptif terhadap variabel-variabel yang digunakan dalam analisis sebagai berikut.

Tabel 2. Statistik Deskriptif

Variabel	n	Mean	Varians	Minimum	Maksimum
Y	34.086	2,353	2,409	0	13
X ₁	34.086	35,335	65,157	15	49
X ₂	34.086	15,877	83,156	1	34
X ₃	34.086	1,495	0,249	1	2
X ₄	34.086	1,795	0,498	0	3
X ₅	34.086	1,068	0,063	1	2
X ₆	34.086	2,927	2,048	1	5
X ₇	34.086	2,986	0,042	0	3
X ₈	34.086	1,760	2,045	0	3

Sumber: Hasil olah data, 2025

Secara keseluruhan, hasil statistik deskriptif menunjukkan bahwa variabel-variabel pada penelitian ini memiliki nilai yang bervariasi yang memungkinkan adanya perbedaan karakteristik antar responden. Berdasarkan tabel terlihat variabel respons Y memiliki nilai rata-rata sebesar 2,353 dan variansi sebesar 2,409 yang relatif mendekati atau identik, yang mana terindikasi tidak memenuhi asumsi equidispersi sehingga perlu dilakukan pengujian statistik dispersi, tentunya melibatkan variabel prediktor karena akan dianalisis menggunakan regresi.

Pemodelan Regresi *Poisson*

Dilakukan pemodelan regresi *Poisson* untuk mengidentifikasi faktor-faktor yang memengaruhi jumlah anak yang pernah dilahirkan dengan prediktornya. Dalam model ini, seluruh variabel kategorik dikonversi ke dalam bentuk variabel dummy. Bentuk model yang diperoleh sebagai berikut.

$$\begin{aligned}
\log(\lambda_i) = & -0,6162 + 0,042 \times \text{Usia Wanita}_i + 0,016 \times \text{Sumut}_i - 0,107 \times \text{Sumbar}_i \\
& - 0,061 \times \text{Riau}_i - 0,224 \times \text{Jambi}_i - 0,134 \times \text{Sumsel}_i - 0,174 \times \text{Bengkulu}_i \\
& - 0,249 \times \text{Lampung}_i - 0,220 \times \text{Babel}_i - 0,157 \times \text{Kepri}_i - 0,236 \times \text{Jakarta}_i \\
& - 0,241 \times \text{Jabar}_i - 0,350 \times \text{Jateng}_i - 0,449 \times \text{DIY}_i - 0,403 \times \text{Jatim}_i \\
& - 0,181 \times \text{Banten}_i - 0,290 \times \text{Bali}_i - 0,170 \times \text{NTB}_i + 0,079 \times \text{NTT}_i \\
& - 0,128 \times \text{Kalbar}_i - 0,165 \times \text{Kalteng}_i - 0,220 \times \text{Kalsel}_i - 0,096 \times \text{Kaltim}_i \\
& + 0,056 \times \text{Kalut}_i - 0,228 \times \text{Sulut}_i - 0,114 \times \text{Sulteng}_i - 0,091 \times \text{Sulsel}_i \\
& + 0,043 \times \text{Sulgar}_i - 0,154 \times \text{Gorontalo}_i + 0,020 \times \text{Sulbar}_i \\
& + 0,163 \times \text{Maluku}_i + 0,072 \times \text{Malut}_i + 0,126 \times \text{PapuaBarat}_i \\
& + 0,184 \times \text{Papua}_i - 0,008 \times \text{Desa}_i - 0,142 \times \text{PendidikanDasar}_i \\
& - 0,252 \times \text{PendidikanMenengah}_i - 0,455 \times \text{PendidikanTinggi}_i \\
& - 0,063 \times \text{KRTPerempuan}_i - 0,088 \times \text{Miskin}_i - 0,133 \times \text{Menengah}_i \\
& - 0,139 \times \text{Kaya}_i - 0,106 \times \text{SangatKaya}_i + 0,446 \times \text{TahuKF}_i \\
& + 0,056 \times \text{TahuKT}_i + 0,214 \times \text{TahuKM}_i + 0,266 \times \text{KontrasepsiTradisional}_i \\
& + 0,291 \times \text{KontrasepsiModern}_i
\end{aligned}$$

Setelah memperoleh model regresi *Poisson*, dilakukan pengujian equidisersi. Hal ini didasarkan pada hasil eksplorasi data sebelumnya yang mengindikasikan bahwa asumsi equidisersi tidak terpenuhi. Pengujian ini dilakukan untuk menilai apakah model regresi *Poisson* sesuai digunakan terhadap data yang dianalisis. Berikut pengujian statistic disperse.

$$\theta = \frac{X^2}{db} = \frac{20.449,9093}{34.037} = 0,6008$$

Berdasarkan hasil statistik dispersi, diperoleh 0,6008 yang lebih kecil dari 1 sehingga dapat dikatakan bahwa data mengalami underdispersi sehingga tidak cocok atau tidak sesuai jika menggunakan regresi *Poisson*. Dengan demikian, diperlukan metode lain untuk menangani kondisi ini, dimana pada penelitian ini digunakan metode regresi *Generalized Poisson*.

Pemodelan Regresi *Generalized Poisson*

Pengujian statistik dispersi menyatakan bahwa data mengalami underdispersi sehingga akan dilanjutkan dengan metode alternatif, yaitu regresi *Generalized Poisson*. Berikut hasil pemodelan regresi *Generalized Poisson*.

$$\begin{aligned}
\log(\lambda_i) = & -0,3837 + 0,044 \times \text{Usia Wanita}_i + 0,0159 \times \text{Sumut}_i - 0,110 \times \text{Sumbar}_i \\
& - 0,055 \times \text{Riau}_i - 0,250 \times \text{Jambi}_i - 0,142 \times \text{Sumsel}_i - 0,197 \times \text{Bengkulu}_i \\
& - 0,229 \times \text{Lampung}_i - 0,225 \times \text{Babel}_i - 0,164 \times \text{Kepri}_i - 0,248 \times \text{Jakarta}_i \\
& - 0,253 \times \text{Jabar}_i - 0,377 \times \text{Jateng}_i - 0,467 \times \text{DIY}_i - 0,405 \times \text{Jatim}_i \\
& - 0,167 \times \text{Banten}_i - 0,301 \times \text{Bali}_i - 0,199 \times \text{NTB}_i + 0,077 \times \text{NTT}_i \\
& - 0,132 \times \text{Kalbar}_i - 0,152 \times \text{Kalteng}_i - 0,234 \times \text{Kalsel}_i - 0,095 \times \text{Kaltim}_i \\
& + 0,062 \times \text{Kalut}_i - 0,238 \times \text{Sulut}_i - 0,114 \times \text{Sulteng}_i - 0,090 \times \text{Sulsel}_i \\
& + 0,038 \times \text{Sulgar}_i - 0,155 \times \text{Gorontalo}_i + 0,025 \times \text{Sulbar}_i \\
& + 0,156 \times \text{Maluku}_i + 0,063 \times \text{Malut}_i + 0,132 \times \text{PapuaBarat}_i \\
& + 0,197 \times \text{Papua}_i - 0,022 \times \text{Desa}_i - 0,138 \times \text{PendidikanDasar}_i \\
& - 0,251 \times \text{PendidikanMenengah}_i - 0,448 \times \text{PendidikanTinggi}_i \\
& - 0,059 \times \text{KRTPerempuan}_i - 0,098 \times \text{Miskin}_i - 0,153 \times \text{Menengah}_i \\
& - 0,168 \times \text{Kaya}_i - 0,140 \times \text{SangatKaya}_i + 0,406 \times \text{TahuKF}_i \\
& + 0,077 \times \text{TahuKT}_i + 0,179 \times \text{TahuKM}_i + 0,216 \times \text{KontrasepsiTradisional}_i \\
& + 0,234 \times \text{KontrasepsiModern}_i
\end{aligned}$$

Dengan parameter dispersi θ sebesar -0,2272 yang menunjukkan bahwa model mengalami underdispersi, di mana varians data lebih kecil dari rata-ratanya. Oleh karena itu, penggunaan model *Generalized Poisson Regression* lebih sesuai dibandingkan model *Poisson* biasa yang mengasumsikan kondisi equidisersi.

Selanjutnya dilakukan uji hipotesis simultan untuk mengetahui pengaruh signifikansi dari setiap variabel independen secara bersama-sama terhadap variabel dependen menggunakan uji *likelihood ratio*. Berikut pengujian hipotesis uji simultan.

$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_{50} = 0$; tidak ada variabel independen yang berpengaruh secara signifikan terhadap variabel Jumlah Anak yang Pernah Dilahirkan

H_1 : Paling tidak ada satu $\beta_k \neq 0$ dengan $k = 1, 2, 3, \dots, 49$ yang berpengaruh secara signifikan terhadap variabel Jumlah Anak yang Pernah Dilahirkan

Diperoleh statistik ujinya sebagai berikut.

$$G = 2(-52.258 - (-60.592))$$

$$G = 16.668$$

Berdasarkan uji simultan, dengan taraf signifikansi 5% diperoleh hasil G sebesar 16.668 lebih besar dari $\chi^2_{49; 0,05}$ sebesar 66,3387 sehingga tolak H_0 atau dikatakan paling tidak ada satu $\beta_k \neq 0$ dengan $k = 1, 2, 3, \dots, 49$ yang berpengaruh secara signifikan terhadap variabel Jumlah Anak yang Pernah Dilahirkan. Sebagaimana kesimpulan uji simultan yang telah dilakukan maka dilakukan juga uji hipotesis parsial menggunakan uji Wald dengan pengujian hipotesis sebagai berikut.

$H_0 : \beta_k = 0$ dengan $k = 1, 2, 3, \dots, 49$; variabel independen ke- k tidak berpengaruh secara signifikan terhadap variabel Jumlah Anak yang Pernah Dilahirkan

$H_1 : \beta_k \neq 0$ dengan $k = 1, 2, 3, \dots, 49$; variabel independen ke- k berpengaruh secara signifikan terhadap variabel Jumlah Anak yang Pernah Dilahirkan

Diperoleh hasil statistik uji parsial sebagai berikut.

Tabel 3. Hasil Uji Parsial

Variabel	Kategori	β	IRR	P-Value
Usia Wanita (X_1)	Usia Wanita Saat Ini	0,044	1,045	<,0001*
	Aceh	Ref		
	Sumatera Utara	0,016	1,016	0,3587
	Sumatera Barat	-0,110	0,896	<,0001*
	Riau	-0,054	0,947	0,0133*
	Jambi	-0,250	0,779	<,0001*
	Sumatera Selatan	-0,142	0,868	<,0001*
	Bengkulu	-0,197	0,821	<,0001*
	Lampung	-0,229	0,796	<,0001*
	Bangka Belitung	-0,225	0,799	<,0001*
	Kepulauan Riau	-0,164	0,849	<,0001*
	Jakarta	-0,248	0,780	<,0001*
	Jawa Barat	-0,253	0,777	<,0001*
	Jawa Tengah	-0,377	0,686	<,0001*
	Yogyakarta	-0,467	0,627	<,0001*
	Jawa Timur	-0,405	0,667	<,0001*
	Banten	-0,167	0,846	<,0001*
	Bali	-0,301	0,740	<,0001*
	Nusa Tenggara Barat	-0,199	0,820	<,0001*
	Nusa Tenggara Timur	0,077	1,080	<,0001*
	Kalimantan Barat	-0,132	0,876	<,0001*
	Kalimantan Tengah	-0,152	0,859	<,0001*
	Kalimantan Selatan	-0,234	0,791	<,0001*
	Kalimantan Timur	-0,095	0,909	<,0001*
	Kalimantan Utara	0,062	1,064	0,0148*
	Sulawesi Utara	-0,238	0,788	<,0001*
	Sulawesi Tengah	-0,114	0,892	<,0001*
	Sulawesi Selatan	-0,090	0,914	<,0001*
	Sulawesi Tenggara	0,039	1,039	0,0467*
	Gorontalo	-0,155	0,857	<,0001*
	Sulawesi Barat	0,025	1,025	0,1917
	Maluku	0,157	1,170	<,0001*
	Maluku Utara	0,063	1,065	0,0034*
	Papua Barat	0,132	1,141	<,0001*
Provinsi (X_2)	Perkotaan	Ref		
	Pedesaan	-0,022	0,978	0,001*
Jenis Tempat Tinggal (X_3)	Perkotaan	Ref		
	Pedesaan	-0,022	0,978	0,001*

Variabel	Kategori	β	IRR	P-Value
Tingkat Pendidikan (X_4)	Tidak Berpendidikan	Ref		
	Pendidikan Dasar	-0,138	0,871	<,0001*
	Pendidikan Menengah	-0,251	0,778	<,0001*
	Pendidikan Tinggi	-0,449	0,638	<,0001*
Jenis Kelamin Kepala Rumah Tangga (X_5)	Laki-laki	Ref		
	Perempuan	-0,059	0,943	<,0001*
Indeks Kekayaan Rumah Tangga (X_6)	Sangat Miskin	Ref		
	Miskin	-0,098	0,907	<,0001*
	Menengah	-0,153	0,858	<,0001*
	Kaya	-0,168	0,846	<,0001*
	Sangat Kaya	-0,140	0,869	<,0001*
Pengetahuan Kontrasepsi (X_8)	Tidak Tahu Kontrasepsi	Ref		
	Tahu Kontrasepsi <i>Folklore</i> (KF)	0,406	1,501	0,0003*
	Tahu Kontrasepsi Tradisional (KT)	0,077	1,080	0,5621
	Tahu Kontrasepsi Modern (KM)	0,179	1,196	<,0001*
Penggunaan Kontrasepsi (X_9)	Tidak Kontrasepsi	Ref		
	Kontrasepsi <i>Folklore</i> (KF)	0	0	0
	Kontrasepsi Tradisional (KT)	0,217	1,242	<,0001*
	Kontrasepsi Modern (KM)	0,234	1,263	<,0001*

Sumber: Hasil olah data, 2025

*signifikan dengan $P\text{-Value} < 0,05$

Berdasarkan hasil estimasi parameter model regresi Generalized Poisson, diketahui bahwa beberapa variabel memiliki pengaruh signifikan terhadap jumlah anak yang pernah dilahirkan oleh perempuan menikah di Indonesia. Variabel usia perempuan menunjukkan hubungan positif terhadap jumlah anak, yang berarti semakin tinggi usia perempuan, semakin banyak jumlah anak yang dilahirkan. Hal ini wajar karena usia yang lebih tinggi mencerminkan durasi pernikahan yang lebih lama sehingga peluang melahirkan anak lebih banyak juga meningkat. Variabel provinsi menunjukkan adanya variasi jumlah anak antar wilayah. Beberapa provinsi seperti Jawa Barat dan Banten memiliki koefisien negatif yang signifikan, yang menunjukkan bahwa perempuan di provinsi tersebut cenderung memiliki jumlah anak lebih sedikit dibandingkan provinsi referensi. Sebaliknya, ada pula provinsi dengan pengaruh positif meskipun tidak semua signifikan secara statistik. Wilayah tempat tinggal juga memiliki pengaruh, di mana perempuan yang tinggal di pedesaan cenderung memiliki lebih banyak anak dibandingkan mereka yang tinggal di perkotaan. Hal ini dapat dikaitkan dengan norma sosial dan akses terhadap layanan keluarga berencana yang berbeda antar wilayah.

Tingkat pendidikan perempuan menunjukkan pengaruh negatif terhadap jumlah anak yang dilahirkan. Semakin tinggi pendidikan yang dimiliki, maka semakin sedikit jumlah anak yang dilahirkan. Ini menunjukkan bahwa pendidikan berperan dalam mengubah preferensi dan perilaku reproduksi, seperti menunda usia pernikahan atau lebih memahami manfaat program keluarga berencana. Jenis kelamin kepala rumah tangga juga berpengaruh, di mana rumah tangga dengan kepala keluarga perempuan cenderung memiliki jumlah anak lebih sedikit dibandingkan rumah tangga yang dikepalai oleh laki-laki. Selain itu, tingkat kekayaan menunjukkan pengaruh negatif terhadap jumlah anak. Semakin tinggi tingkat kekayaan rumah tangga, cenderung semakin sedikit jumlah anak yang dilahirkan. Hal ini mungkin terkait dengan gaya hidup, akses terhadap pendidikan, dan penggunaan kontrasepsi yang lebih baik di kalangan rumah tangga kaya. Pengetahuan tentang kontrasepsi ternyata memiliki pengaruh yang positif terhadap jumlah anak, namun tidak signifikan. Artinya, meskipun pengetahuan ada, belum tentu diikuti dengan praktik penggunaannya secara efektif. Sementara itu, penggunaan kontrasepsi berhubungan negatif dan signifikan terhadap jumlah anak, yang berarti perempuan yang menggunakan kontrasepsi memiliki kecenderungan melahirkan lebih sedikit anak.

D. Kesimpulan

Berdasarkan hasil estimasi regresi *Generalized Poisson*, dapat disimpulkan bahwa faktor-faktor yang secara signifikan memengaruhi terhadap jumlah anak yang pernah dilahirkan oleh wanita menikah di Indonesia meliputi usia wanita, sebagian besar provinsi, jenis tempat tinggal, tingkat pendidikan, jenis kelamin kepala rumah tangga, indeks kekayaan rumah tangga, serta pengetahuan dan penggunaan kontrasepsi. Dengan demikian, berdasarkan temuan dari hasil tersebut diharapkan bagi pemerintah untuk merancang kebijakan perlu difokuskan pada peningkatan akses pendidikan dan layanan kontrasepsi, khususnya di wilayah pedesaan dan kelompok masyarakat dengan tingkat kesejahteraan rendah, guna mengendalikan jumlah kelahiran dan meningkatkan kesejahteraan keluarga secara berkelanjutan.

Ucapan Terimakasih

Penulis menyampaikan terima kasih kepada Universitas Islam Bandung atas dukungan akademik yang telah diberikan dalam pelaksanaan penelitian ini. Ucapan terima kasih juga disampaikan kepada dosen pembimbing atas bimbingan dan arahnya, serta kepada instansi terkait yang telah menyediakan data sebagai dasar analisis. Apresiasi turut diberikan kepada seluruh pihak yang telah membantu dan mendukung kelancaran penelitian ini hingga selesai.

Daftar Pustaka

- Al-Balushi, M. S., Ahmed, M. S., Islam, M. M., & Khan, M. H. R. (2020). Multilevel poisson regression modeling to identify factors influencing the number of children ever born to married women in Oman. *Journal of Statistics and Management Systems*, 23(8), 1357–1373. <https://doi.org/10.1080/09720510.2019.1709328>
- Badan Perencanaan dan Pembangunan Nasional. (2023). *Penduduk Berkualitas Menuju Indonesia Emas*.
- Badan Pusat Statistik. (2011). Fertilitas Penduduk Indonesia: Hasil Sensus Penduduk 2010. In *Badan Pusat Statistik*.
- Cherie, N., Getacher, L., Belay, A., Gultie, T., Mekuria, A., Sileshi, S., & Degu, G. (2023). Modeling on number of children ever born and its determinants among married women of reproductive age in Ethiopia: A Poisson regression analysis. *Heliyon*, 9(3), e13948. <https://doi.org/10.1016/j.heliyon.2023.e13948>
- Damayanti CR, M., & Yanti, T. S. (2022). Regresi Poisson Invers Gaussian (PIG) untuk Pemodelan Jumlah Kasus Pneumonia pada Balita di Provinsi Jawa Tengah Tahun 2019. *Jurnal Riset Statistika*, 1(2), 143–151. <https://doi.org/10.29313/jrs.v1i2.523>
- Dewi, E. F., & Hajarisman, N. (n.d.). *Penanganan Data Hilang pada Pemodelan Persamaan Terstruktur melalui Metode Full Information Maximum Likelihood (FIML)*. <https://journals.sangkurianpublisher.com/datamath>
- Fauziah, G., & Sunendiari, S. (2021). Estimasi Pseudo Poisson Maximum Likelihood untuk Mengatasi Masalah dalam Model Log-Linear pada Kasus Kusta di Jawa Barat Tahun 2018. *Jurnal Riset Statistika*, 1(1), 57–62. <https://doi.org/10.29313/jrs.v1i1.147>
- Gebre, M. N. (2024). Number of children ever-born and its associated factors among currently married Ethiopian women: evidence from the 2019 EMDHS using negative binomial regression. *BMC Women's Health*, 24(1). <https://doi.org/10.1186/s12905-024-02883-w>
- Hajarisman, N. (2022). Analisis Data Kategorik. In *Universitas Islam Bandung*.
- Hilbe, J. M. (2014). *Modeling Count Data*. Cambridge University Press.
- United Nations. (2024). World Population Prospects 2024: Summary of Results. In *United Nation* (Issue 9). www.un.org/development/desa/pd/.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). Probability and Statistics for Engineers and Scientists. In *Pearson* (9th ed.).
- Yang, Z., Hardin, J. W., & Addy, C. L. (2009). A score test for overdispersion in Poisson regression based on the generalized Poisson-2 model. *Journal of Statistical Planning and Inference*, 139(4), 1514–1521. <https://doi.org/10.1016/j.jspi.2008.08.018>