

## Analisis Perbandingan Metode Distribusi Binomial Negatif Dan Poisson (Studi Kasus: Faktor-Faktor Terjadinya Perceraian di Indonesia Tahun 2023)

Lidya Ananda Talalu<sup>a\*</sup>, Bagus Kusuma<sup>b</sup>

<sup>a,b</sup> Prodi Matematika Militer, Universitas Pertahanan Republik Indonesia  
email: [lidyaananda2033@gmail.com](mailto:lidyaananda2033@gmail.com)\*

### INFO ARTIKEL

#### Riwayat Artikel:

Received : 7 Desember 2024

Revised : 31 Desember 2024

Accepted : 5 Januari 2025

#### Keywords:

Divorce; Integrated Nested Laplace Approximation (INLA); Widely Applicable Information Criterion (WAIC)

#### Kata Kunci:

Perceraian; Integrated Nested Laplace Approximation (INLA); Widely Applicable Information Criterion (WAIC)

DOI: 10.62335

### ABSTRACT

The increasing number of divorce cases in Indonesia is suspected to influence the country's social structure. Factors such as the Human Development Index, Regional Gross Domestic Product (RGDP), and education quality are believed to play significant roles. Negative binomial distribution and Poisson distribution were utilized to analyze divorce data in Indonesia using the Integrated Nested Laplace Approximation (INLA) approach. The Widely Applicable Information Criterion (WAIC) metric was employed to determine the most suitable distribution for the data. The study recommends using the negative binomial distribution with the INLA approach for further analysis in modeling divorce cases in Indonesia.

### ABSTRAK

Semakin banyaknya kasus perceraian di Indonesia diduga bisa mempengaruhi sistem ketatanan negara, diantaranya dapat dipengaruhi oleh Indeks Pembangunan Manusia, Produk Domestik Regional Bruto (PDRB), dan kualitas pendidikan. Distribusi binomial negatif dan distribusi poisson digunakan untuk menganalisis data jumlah perceraian di Indonesia menggunakan pendekatan *Integrated Nested Laplace Approximation (INLA)*. Metrik *Widely Applicable Information Criterion (WAIC)* bertujuan untuk memilih distribusi mana yang cocok digunakan pada data tersebut. Hasil penelitian merekomendasikan penggunaan distribusi binomial negatif dengan pendekatan INLA untuk analisis lebih lanjut dalam memodelkan kasus perceraian di Indonesia.

### LATAR BELAKANG

Pada umumnya, keluarga terdiri dari ayah, ibu, dan anak-anak. Namun, seringkali dijumpai dalam suatu keluarga, kondisi orangtuanya tidak lengkap. Keadaan itu disebut orang tua tunggal, dimana salah satu orang tua hanya membesarkan anak-anaknya secara mandiri tanpa ada bantuan, dukungan, tanggung jawab dari pasangannya. Hal tersebut terjadi karena adanya "perceraian".

Perceraian adalah putusannya atau berakhirnya suatu hubungan suami istri yang diputuskan secara hukum maupun agama (talak) karena sudah tidak ada rasa kecocokan satu sama lain yang menyebabkan ketidakharmonisan dalam keluarga tersebut (Untari, Dhini, & Hafiddudin, 2018).

Berdasarkan data yang didapat dari Badan Pusat Statistika (BPS) tahun 2023, terdapat 408.347 kasus perceraian yang tersebar di 34 provinsi Indonesia. Faktor-faktor perceraian diantaranya, zina, mabuk, judi, meninggalkan salah satu pihak, poligami, kekerasan dalam rumah tangga, dll. Dan kasus terbanyak dengan jumlah 251.282 kasus perceraian terjadi dikarenakan adanya perselisihan dan pertengkaran terus menerus dalam rumah tangga. Banyak sekali penelitian yang mengungkapkan dengan adanya ketidakharmonisan dalam suatu keluarga mampu menyebabkan dampak-dampak negatif, yang dimana hal tersebut bisa mempengaruhi ketatanan suatu negara (Hasanah, 2019). Remaja-remaja yang mengalami kasus perceraian orang tuanya bisa membawa dampak ke lingkungan sekitar, seperti menjadi seorang yang pembangkang, tidak taat aturan, pergaulan bebas, dll. Dikarenakan kebutuhan seorang anak tidak terpenuhi, yaitu tidak mendapatkan kasih sayang dan perhatian dari orang terdekatnya.

Penelitian terdahulu membahas tentang faktor-faktor penyebab terjadinya perceraian di Kab. Pinrang menggunakan metode regresi logistik. Regresi logistik biner merupakan metode yang digunakan untuk mengetahui hubungan antar variabel independen dengan variabel dependen (Irwan, Ibtisam, & Azisa, 2023). Dengan tujuan memperoleh model dan mengetahui faktor apa saja yang berpengaruh signifikan terhadap penyebab terjadinya kasus perceraian di Kab. Pinrang. Pada penelitian ini dibagi menjadi dua model parameter, yaitu uji simultan dan uji parsial. Untuk hasil uji simultan, keputusan  $H_0$  diterima artinya parameter bernilai sama dengan 0 berarti tidak ada satupun pengaruh signifikan terhadap variabel independen. Sedangkan, hasil uji parsial didapatkan bahwa variabel independen tidak berpengaruh terhadap variabel dependen.

Sehingga, tujuan penelitian ini adalah menyempurnakan metode regresi logistik biner pada artikel analisis kasus perceraian di Kab. Pinrang. Penulis berupaya untuk menganalisis kasus perceraian yang lebih meluas, yaitu di Indonesia. Dengan membandingkan metode negatif binomial dan gamma diharapkan mampu memberikan kesimpulan metode mana yang lebih efisien dan signifikan untuk digunakan pada perhitungan kasus perceraian di Indonesia. Bantuan *software* R studio digunakan dalam penelitian ini untuk menganalisis distribusi binomial negatif dengan gamma menggunakan metode *INLA*.

## **KAJIAN PUSTAKA DAN PERUMUSAN HIPOTESIS**

### **Distribusi Binomial Negatif**

Distribusi binomial negatif adalah suatu distribusi peluang acak  $x$  yang terdiri dari  $n$  usaha berulang yang akan menghasilkan dua kemungkinan antara gagal atau berhasil, dimana  $x$  dinyatakan sebagai banyaknya percobaan yang diperlukan untuk mendapatkan  $k$  sukses (Sauddin, Auliah, & Alwi, 2020). Regresi binomial negatif memiliki tiga komponen, diantaranya komponen random yang dihasilkan dari distribusi mixture Poisson-Gamma dan komponen sistematis dinyatakan dalam bentuk

kombinasi linear antara parameter ( $\eta$ ). Dalam teori probabilitik, regresi binomial negatif merupakan suatu distribusi yang memodelkan jumlah kegagalan dalam uji coba Bernoulli yang independen sebelum jumlah keberhasilan (non-acak) tertentu terjadi. Fungsi masa probabilitik dari regresi binomial negatif sebagai berikut.

$$f(k; r, p) \equiv \Pr(X = k) \left( \frac{K + r + 1}{k} \right) (1 - p)^k p^r$$

Keterangan:

$r$  = jumlah keberhasilan

$k$  = jumlah kegagalan

$p$  = peluang keberhasilan pada setiap percobaan

Kuantitas dalam tanda kurung fungsi probabilitik adalah koefisien binomial, dan sama dengan:

$$\binom{k+r-1}{k} = \frac{(k+r-1)!}{(r-1)!(k)!} = \frac{(k+r-1)(k+r-2) \dots (r)}{k!} = \frac{\Gamma(k+r)}{k! \Gamma(r)}$$

Keterangan:

$\Gamma(r)$  = fungsi gamma

Ada  $k$  kegagalan yang dipilih dari  $(k+r-1)$  percobaan dan bukan  $(k+r)$  karena percobaan terakhir dari  $(k+r)$  percobaan menurut definisinya adalah sebuah keberhasilan. Campuran antara distribusi Poisson dan Gamma ini menghasilkan distribusi negatif binomial dan memberikan model yang masuk akal untuk mendukung penggunaan distribusi negatif binomial (White & Bennetts, 1996).

### Distribusi Poisson

Distribusi poisson merupakan distribusi yang menyatakan probabilitas atau peluang jumlah peristiwa yang terjadi dalam periode waktu tertentu. Merepresentasikan probabilitas terjadinya suatu peristiwa dalam eksperimen A yang terkait dengan jumlah total percobaan (Zhao, Zhang, Zhao, Cao, & Wu, 2020), dengan rumus sebagai berikut:

$$f(k; \lambda) = \frac{(\lambda^k e^{-\lambda})}{k!}$$

dengan,

- $e$  adalah basis logaritma natural
- $k$  adalah jumlah kejadian suatu peristiwa

- $\lambda$  adalah bilangan riil positif

Distribusi ini memiliki rata-rata varians sebesar  $\lambda$ . Distribusi Poisson cocok digunakan untuk memodelkan kejadian langka, seperti jumlah panggilan masuk dalam satu jam, kesalahan produksi per unit, atau gempa bumi dalam setahun. Ketika nilai rata-rata kecil, distribusinya condong ke kiri, sedangkan pada rata-rata besar, distribusi ini mendekati normal.

### ***Integrated Nested Laplace Approximation (INLA)***

Integrated Nested Laplace Approximation (INLA) merupakan suatu metode untuk menghasilkan estimasi yang lebih akurat dan lebih cepat waktu komputasinya. Sesuai namanya, metode ini menggunakan pendekatan Laplace (Firda, Djuraidah, & Kurnia, 2019). Tujuan utama dari pendugaan Bayes dengan *Aproximation Bayesian with INLA* adalah mencari sebaran posterior marjinal untuk setiap elemen dari vektor parameter sebagai berikut.

$$p(\theta_i|y) = \int p(\theta_i, \varphi|y)d\varphi = \int p(\theta_i|\varphi, y)p(\varphi|y)d\varphi$$

Metode INLA ini dapat digunakan untuk pemodelan distribusi spesies, pemodelan spasial, pemodelan hierarkis data spasial, mengimplementasikan model spasial dan spasio-temporal untuk data referensi titik.

### ***Widely Applicable Information Criterion (WAIC)***

*Watanabe Akaike Information Criterion (WAIC)* atau *Widely Applicable Information Criterion* merupakan salah satu pendekatan Bayesian yang lebih lengkap untuk menduga data *out-of-sample* (Abdullah, Iriawan, & Irhamah, 2020). Berbeda dari AIC atau DIC yang hanya mengandalkan perkiraan parameter titik, WAIC mempertimbangkan ketidakpastian model dengan mengintegrasikan distribusi posterior dari parameter tersebut, memberikan gambaran lebih menyeluruh tentang kecocokan model. Nilai WAIC yang lebih rendah menunjukkan model yang lebih tepat dalam memprediksi data tanpa cenderung overfit, dimana rumusnya sebagai berikut:

$$WAIC = -2 \cdot LPPD + 2 \cdot P_{WAIC}$$

Dimana,

$$LPPD = \sum_{i=1}^n \log \left( \frac{1}{S} \sum_{s=1}^S f(y_i|\theta_s) \right)$$

Keunggulan utama WAIC terletak pada kemampuannya untuk diterapkan pada berbagai model, termasuk yang tidak memerlukan asumsi distribusi normal atau yang memiliki struktur lebih rumit. Hal ini membuatnya lebih serbaguna dibandingkan dengan metrik lainnya seperti AIC atau DIC yang lebih terbatas. Meskipun begitu, karena WAIC melibatkan perhitungan log-likelihood untuk setiap data dan varians dalam distribusi posterior, proses perhitungannya bisa lebih berat secara komputasi. Meski demikian, WAIC tetap banyak digunakan dalam analisis model Bayesian karena

kemampuannya memberikan evaluasi yang komprehensif terhadap performa model.

## **METODE PENELITIAN**

Data yang digunakan dalam penelitian ini adalah kumpulan dari sumber sumber seperti, wikipedia, BPS, databoks, dll. Sehingga didapatkan satu tabel dengan variabel dependen jumlah perceraian, dan variabel independennya adalah Indeks Pembangunan Manusia (IPM), Produk Domestik Regional Bruto (PDRB), dan kualitas pendidikan.

### **Prosedur analisis**

Analisis data dalam penelitian ini adalah sebagai berikut.

1. Mengumpulkan data-data perceraian di 34 provinsi Indonesia dari berbagai website sesuai dengan variabel independennya.
2. Membuat deskriptif data pada Ms. Excel.
3. Menghitung mean dan varians.
4. Menganalisis hubungan antar variabel (perceraian-IPM, perceraian-PDRB, dan perceraian-KP).
5. Menghitung uji statistik sebaran binomial negatif dan distribusi poisson dengan metode *INLA* dengan bantuan software R studio.
6. Menentukan distribusi terbaik dengan metrik *Widely Applicable Information Criterion (WAIC)*.
7. Menarik kesimpulan berdasarkan analisis yang diperoleh.

## **ANALISIS DAN PEMBAHASAN**

Dari data yang terkumpul pada tahun 2023 terhitung lebih dari 500.000 kasus perceraian yang terjadi di 34 provinsi Indonesia, diantaranya disebabkan oleh faktor Indeks Pembangunan Manusia (IPM), Produk Domestik Regional Bruto (PDRB), dan kualitas pendidikan (KP). Dengan jumlah kasus perceraian terbanyak disebabkan oleh Produk Domestik Regional Bruto (PDRB) yang terjadi di DKI Jakarta dengan jumlah 322.615 kasus.

### **1. Menghitung mean dan varians**

Berdasarkan hasil perhitungan dengan bantuan software R studio, didapat mean dari kasus perceraian ini sebesar 9,125 menunjukkan rata-rata jumlah perceraian di seluruh wilayah yang diamati cenderung tinggi. Sedangkan, untuk varians sebesar 675,784 menunjukkan penyebaran jumlah perceraian yang sangat besar.

**Tabel 1.** Hasil mean dan varians dengan bantuan software R studio

Mean	Varians
9,125	675,784

## 2. Menganalisis hubungan antar variabel

Analisis korelasi antar variabel dilakukan untuk mengetahui variabel independen mana yang lebih signifikan atau berhubungan besar dengan variabel dependen (jumlah perceraian). Sehingga didapat data sebagai berikut.

**Tabel 2.** Analisis korelasi

Variabel Dependen	Variabel Independen	Korelasi
Jumlah perceraian	IPM	0.0218
Jumlah perceraian	PDRB	-0.128
Jumlah perceraian	KP	-0.0331

Dapat dilihat pada tabel di atas, dimana antara jumlah perceraian dan IPM memiliki korelasi sebesar 0.0218 yang berarti berkorelasi lemah sehingga hubungan antar keduanya hampir tidak signifikan. Korelasi antara jumlah perceraian dengan PDRB sebesar -0.128 yang berarti hubungan antar keduanya kecil ke arah negatif. Sedangkan, untuk jumlah perceraian dengan kualitas pendidikan menunjukkan hasil sebesar -0.0331 yang artinya hubungan keduanya hampir tidak signifikan.

## 3. Uji statistik dengan INLA

### a. Uji distribusi binomial negatif dengan INLA

**Tabel 3.** Hasil uji distribusi binomial negatif dengan INLA

	mean	std. deviasi	0.025quant	0.5quant	0.975quant	mode
(Intercept)	-62.9574	21.6861	-105.6875	-62.9195	-20.4224	-62.9147
PDRB	-0.0274	0.0075	-0.0422	-0.0274	-0.0127	-0.0274
IPM	1.3116	0.4254	0.4772	1.3109	2.1497	1.3108
KP	-0.3544	0.1546	-0.659	-0.3542	-0.0509	-0.3541

Berdasarkan tabel 3 hasil konsisten menunjukkan bahwa seluruh variabel independen memiliki korelasi dengan variabel dependen. Seperti yang dapat kita analisis, dimana jika pada tabel

$0.025quant$ ;  $0.5quant$ ; dan  $0.975quant$  tidak terdapat nilai nol. Sehingga, tiap variabel memiliki hubungan yang signifikan terhadap jumlah kasus perceraian di Indonesia.

Koefisien intercept (-62.957) menunjukkan nilai negatif yang mengindikasikan bahwa tanpa pengaruh variabel independen dan angka perceraian berada pada level yang sangat rendah. Koefisien PDRB (-0.0274) menunjukkan angka perceraian akan menurun jika terdapat peningkatan pada angka PDRB, meskipun pengaruhnya relatif kecil. Koefisien IPM (1.3116) menunjukkan dengan meningkatnya IPM cenderung meningkatkan angka perceraian dengan pengaruh yang signifikan. Sedangkan, dengan koefisien kualitas pendidikan (-0.3544) semakin tinggi kualitas pendidikan akan mampu menurunkan kasus perceraian di Indonesia.

#### b. Uji distribusi poisson dengan INLA

**Tabel 4.** Hasil uji distribusi poisson dengan INLA

	mean	std. deviasi	0.025quant	0.5quant	0.975quant	mode
(Intercept)	-15.4021	0.0928	-15.5839	-15.4021	-15.2203	-15.4021
PDRB	-0.0132	0.0001	-0.0133	-0.0132	-0.0131	-0.013
IPM	0.4380	0.0016	0.4349	0.4380	0.4411	0.4380
KP	-0.1073	0.0004	-0.1081	-0.1073	-0.1065	-0.1073

Hasil analisis dengan model INLA menggunakan distribusi Poisson menunjukkan bahwa variabel independen seperti PDRB (Produk Domestik Regional Bruto), IPM (Indeks Pembangunan Manusia), dan KP (Kualitas Pendidikan) secara signifikan memengaruhi tingkat perceraian. Intersep model memiliki rata-rata -15.4021 dengan standar deviasi 0.0928, mengindikasikan tingkat perceraian yang rendah ketika semua variabel independen bernilai nol. Koefisien PDRB sebesar -0.0132 mengartikan bahwa setiap peningkatan PDRB berkaitan dengan sedikit penurunan dalam tingkat perceraian.

Sebaliknya, IPM memiliki koefisien positif 0.4380, yang menunjukkan hubungan langsung antara peningkatan IPM dan kenaikan tingkat perceraian. Sementara itu, KP dengan koefisien -0.1073 mengindikasikan bahwa peningkatan kualitas pendidikan berhubungan dengan penurunan tingkat perceraian. Rentang nilai  $0.025quant$  hingga  $0.975quant$  yang sempit pada setiap variabel menunjukkan konsistensi hasil model.

#### 4. Pemilihan distribusi terbaik

Setelah dilakukan uji analisis kedua distribusi dengan INLA, dilakukan pemilihan distribusi yang paling baik dengan metode *Widely Applicable Information Criterion (WAIC)*. Dengan bantuan software didapat hasil sebagai berikut:

**Tabel 5.** Hasil uji metrik WAIC

Negative Binomial	Poisson
13.70679	Inf

Berdasarkan tabel di atas, ditunjukkan bahwa dengan metode *Widely Applicable Information Criterion (WAIC)* didapat hasil distribusi negatif binomial sebesar 13.70679 dan distribusi poisson bernilai *inf*. Sehingga, distribusi yang paling baik digunakan pada kasus jumlah perceraian di Indonesia tahun 2023 adalah dengan distribusi Negatif Binomial.

## PENUTUP / KESIMPULAN

Jika dilihat pada hasil pembahasan dan asumsi data, maka dapat disimpulkan untuk menganalisis dan mencari korelasi antara variabel dependen dengan variabel independen pada kasus perceraian di Indonesia tahun 2023 akan lebih sesuai jika menggunakan distribusi negatif binomial. Hal tersebut dibuktikan dengan pemilihan model terbaik dengan metrik *Widely Applicable Information Criterion (WAIC)*. Dengan maksud, semakin rendah kualitas pendidikan masyarakat maka semakin tinggi tingkat perceraian. Sedangkan, apabila IPM dan PDRB meningkat maka jumlah perceraian akan menurun. Hal ini membuktikan bahwa ketiga variabel tersebut memiliki pengaruh yang cukup besar terhadap kasus perceraian di Indonesia.

## DAFTAR PUSTAKA

- Abdullah, M. N., Iriawan, N., & Irhamah. (2020). Pendekatan Bayesian untuk Analisis Survival pada Kasus Demam Berdarah Dengue Pasien RSUD Dr. Soetomo Surabaya. *Jurnal Sains dan Seni*, 9, 138-146.
- Firda, R., Djuraidah, A., & Kurnia, A. (2019). Pemodelan Kemiskinan Di Jawa Menggunakan Bayesian Spasial Probit Pendekatan Integrated Nested Laplace Approximation (INLA). *Media Statistika*, 140-151. doi:10.14710/medstat.12.2.140-151
- Gelman, Hwang, & Vehtari. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 997-1016. doi:10.1007/s11222-013-9408-2
- Hasanah, U. (2019). Pengaruh Perceraian Orang Tua Bagi Psikologis Anak. *Jurnal Analisis Gender dan Agama*, 2, 18-24. doi:DOI: <http://dx.doi.org/10.31958/agenda.v2i1.1983>
- Irwan, M., Ibbnas, R., & Azisa, N. (2023). Identifikasi Faktor-Faktor Penyebab Terjadinya Perceraian Di Kabupaten Pinrang Menggunakan Metode Regresi Logistik. 11. doi:<https://doi.org/10.24252/msa.v11i1.18458>

- Sauddin, A., Auliah, N. I., & Alwi, W. (2020). Pemodelan Jumlah Kematian Ibu di Provinsi Sulawesi Selatan Menggunakan Regresi Binomial Negatif. *Semantic Scholar*, 1-6. doi:<https://doi.org/10.24252/MSA.V8I2.17409>
- Simarmata, R. T., & Ispriyanti, D. (2011). Penanganan Overdispersi Pada Model Regresi Poisson Menggunakan Model Regresi Binomial Negatif. *Media Statistika*, 4, 95-99. doi:<https://doi.org/10.14710/medstat.4.2.95-104>
- Untari, I., Dhini, K. P., & Hafiddudin, M. (2018). Dampak Perceraian Orang Tua Terhadap Kesehatan Psikologis Remaja. *Jurnal Profesional Islam*, 1-8.
- White, G. C., & Bennetts, R. E. (1996). Analysis of Frequency Count Data Using the Negative Binomial Distribution. *Ecological Society of America*, 2549-2557. doi:[10.2307/2265753](https://doi.org/10.2307/2265753)
- Zhao, J., Zhang, F., Zhao, C., Cao, X., & Wu, G. (2020). The Properties and Application of Poisson Distribution. *Journal of Physics: Conferences Series*, 1-4. doi:[10.1088/1742-6596/1550/3/032109](https://doi.org/10.1088/1742-6596/1550/3/032109)