

Pemodelan Jumlah Kasus Kusta di Kabupaten Mojokerto dan Kabupaten Jombang Tahun 2019 Menggunakan Regresi Zero-Inflated Poisson Inverse Gaussian

Rahmania Azwarini¹, Purhadi²

^{1,2}Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember
Email: rahmaniaazwarini@gmail.com¹, purhadi@statistika.its.ac.id²

Received: 30 August 2021

Accepted: 27 September 2021

ABSTRAK— Kusta atau lepra adalah penyakit yang menyerang berbagai bagian tubuh diantaranya saraf dan kulit yang disebabkan oleh infeksi bakteri *Mycobacterium leprae*. Jawa Timur merupakan provinsi dengan jumlah penderita kusta tertinggi di Indonesia hingga tahun 2019 sebanyak 3.306 kasus. Stigma negatif masyarakat terhadap penderita kusta menyebabkan munculnya perkampungan kusta di dusun Sumberglagah, Kabupaten Mojokerto. Selain adanya kampung kusta, pada Kabupaten Mojokerto juga terdapat rumah sakit kusta terbesar di Jawa Timur yang menjadi pusat pengobatan kusta baik di Kabupaten Mojokerto maupun di daerah sekitarnya seperti Kabupaten Jombang. Penelitian ini menggunakan data jumlah kasus kusta di Kabupaten Mojokerto dan Kabupaten Jombang tahun 2019 sebagai variabel respon dan enam variabel lainnya sebagai variabel prediktor. Data jumlah kasus kusta tersebut memiliki proporsi nilai nol sebesar 30,77%, lalu nilai mean sebesar 2,179 serta varians sebesar 6,625. Hal ini mengindikasikan bahwa adanya extra zeros serta terdapat pelanggaran asumsi equidispersi. Regresi Zero Inflated-Poisson Inverse Gaussian (ZIPIG) merupakan metode pengembangan regresi yang mampu menangani overdispersi serta extra zeros pada variabel respon data observasi. Faktor yang berpengaruh signifikan terhadap jumlah kasus kusta berdasarkan hasil pemodelan regresi ZIPIG yaitu persentase pelayanan kesehatan untuk penduduk usia lanjut (X_6).

Kata Kunci— *Extra Zeros, Kusta, Overdispersi, Regresi Zero-Inflated Poisson Inverse Gaussian (ZIPIG)*.

I. PENDAHULUAN

Kusta atau disebut juga dengan lepra merupakan penyakit yang menyerang berbagai bagian tubuh diantaranya saraf dan kulit yang disebabkan oleh infeksi bakteri *Mycobacterium leprae*. Penularan penyakit kusta dapat melalui pernapasan, udara, maupun kontak langsung dengan penderita yang belum diobati. Kusta telah dikenal hampir 2000 tahun sebelum masehi. Penderita kusta seringkali mendapatkan stigma negatif dari masyarakat dikarenakan kusta dianggap sebagai penyakit kutukan, keturunan maupun penyakit aib. Adanya perlakuan diskriminatif masyarakat tentunya memiliki dampak buruk pada penderita kusta yang harus menanggung beban fisik serta moral dalam waktu yang berkepanjangan [1].

Saat ini tidak sedikit dari masyarakat Indonesia yang menganggap penyakit kusta telah menghilang, namun pada kenyataan tidak demikian. *World Health Organization* (2020) mencatat bahwa selama 5 tahun berturut-turut mulai tahun 2015 hingga 2019, Indonesia berada pada posisi 3 teratas sebagai negara dengan penderita kusta terbanyak setelah India dan Brazil. Adapun provinsi di Indonesia dengan predikat kasus kusta terbanyak hingga tahun 2019 yaitu Provinsi Jawa Timur, dengan penderita terdaftar sebanyak 3.306 kasus dan angka prevalensi per *rate* sebesar 0,83/10.000 penduduk. Jumlah kasus kusta di Jawa Timur telah memenuhi target Renstra (Rencana Strategi) Kementerian Kesehatan Republik Indonesia dengan angka prevalensi per *rate* bernilai dibawah 1/10.000, meskipun demikian Jawa Timur merupakan provinsi di Indonesia dengan jumlah kasus terbanyak, sehingga sudah sepatutnya jumlah kasus kusta diupayakan menurun [2].

Stigma negatif terhadap penderita kusta seringkali menjadi hambatan dalam upaya penanggulangan penyakit kusta. Keberadaan kampung kusta di Indonesia merupakan bentuk pengasingan masyarakat terhadap penderita kusta, dimana penderita kusta dari berbagai daerah dikumpulkan dan diisolasi dalam suatu tempat terpencil untuk memutus mata rantai penularan. Secara umum terdapat sekitar 60 kampung kusta di Indonesia, salah satunya yaitu kampung kusta di Dusun Sumberglagah, Kabupaten Mojokerto. Masyarakat penderita kusta mampu bertahan hidup di kampung kusta dengan pengobatan seadanya, selain itu di kampung kusta terjadi pernikahan antara sesama penderita kusta dimana hal tersebut berlangsung secara turun temurun [3].

Hingga pada tahun 1952 dibangun rumah sakit kusta di Dusun Sumberglagah Kabupaten Mojokerto sebagai rumah sakit khusus kusta terbesar di Jawa Timur. Rumah sakit kusta Sumberglagah berdasarkan surat keputusan gubernur Jawa Timur nomor 188/359/KPTS/013/2015 ditetapkan sebagai rumah sakit rujukan regional kasus kusta untuk 20 kabupaten/kota di Jawa Timur [1]. Hal ini tentunya berdampak positif pada pengobatan dan perawatan penderita kusta menjadi lebih baik khususnya di Kabupaten Mojokerto serta daerah sekitarnya seperti Kabupaten Jombang.

Jumlah kasus kusta merupakan data cacah yang bernilai non negatif. Metode yang dapat digunakan untuk menganalisis data cacah salah satunya yaitu regresi *Poisson*. Syarat asumsi yang harus dipenuhi pada regresi *Poisson* adalah equidispersi, yaitu adanya kesamaan antara nilai mean dan varians pada variabel respon. Apabila asumsi equidispersi tidak terpenuhi yaitu ketika nilai varians lebih besar dari mean (overdispersi) atau varians lebih kecil dari

mean (underdispersi), hal ini dapat menyebabkan standar eror akan menjadi bias dan statistik uji dari model menjadi kurang tepat, sehingga kesimpulan yang didapatkan menjadi kurang valid dan regresi *Poisson* tidak dapat digunakan [4]. Selain itu data jumlah kasus kusta seringkali mengandung nilai nol berlebih atau *excess zero*. Apabila kondisi *excess zero* diabaikan juga dapat menyebabkan hasil analisis yang kurang tepat.

Pengembangan dari model regresi *Poisson* yang dapat digunakan untuk mengatasi nilai nol berlebih pada data yaitu dengan model regresi *Zero-Inflated Poisson* (ZIP). Model ZIP terdiri atas dua bagian (state) yaitu model *Poisson state* dan model logit. Model *Poisson state* digunakan untuk mewakili data cacah yang tidak bernilai nol pada variabel respon, sedangkan model logit digunakan untuk mewakili data cacah yang bernilai nol pada variabel respon. Selanjutnya pengembangan lainnya dari model regresi *Poisson* yang dapat digunakan untuk mengatasi pelanggaran asumsi equidispersi yaitu menggunakan regresi *Poisson Inverse Gaussian* (PIG) yang terdiri atas dua parameter yaitu λ (rata-rata) dan τ (parameter dispersi).

Penelitian ini selanjutnya dikembangkan dengan membentuk suatu model baru yaitu model regresi *Zero-Inflated Poisson Inverse Gaussian* (ZIPIG) yang merupakan gabungan dari dua model yaitu *Zero Inflated Poisson* (ZIP) dan *Poisson Inverse Gaussian* (PIG), dimana model ZIP berperan untuk menangani *extra zeros* sedangkan model PIG untuk menangani kasus pelanggaran asumsi equidispersi. Kajian terkait metode *Zero-Inflated Poisson Inverse Gaussian* (ZIPIG) telah diteliti oleh Rahayuning, untuk kasus kematian ibu nifas di Karesidenan Pekalongan dimana data terdiri atas nilai nol sebesar 60,44%, varians sebesar 0,912, dan mean sebesar 0,67 yang menandakan bahwa data mengalami overdispersi [5].

Kabupaten Mojokerto dan Kabupaten Jombang dipilih dalam penelitian ini dikarenakan pada Kabupaten Mojokerto terdapat kampung kusta yang terletak pada Dusun Sumberglagah. Sebagian besar penghuni kampung kusta merupakan penderita dan mantan penderita kusta yang berasal dari berbagai daerah, kemudian menetap dan melanjutkan keturunan, dimana hal tersebut berlangsung secara turun temurun hingga saat ini. Selain itu pada Kabupaten Mojokerto terdapat rumah sakit kusta terbesar di Jawa Timur yang menjadi pusat pengobatan kusta baik di Kabupaten Mojokerto maupun di daerah sekitarnya seperti Kabupaten Jombang, dengan mempertimbangkan pula ketersediaan data di Profil Kesehatan 2019 maka dipilihlah kedua kabupaten tersebut. Penelitian ini diharapkan mampu menjadi bahan pertimbangan bagi Dinas Kesehatan Kabupaten Mojokerto dan Kabupaten Jombang dalam menentukan kebijakan terkait penanganan jumlah kasus kusta sehingga kasus kusta dapat ditekan seminimal mungkin.

II. TINJAUAN PUSTAKA

A. Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu data sehingga memberikan informasi yang berguna [6]. Statistika deskriptif yang digunakan dalam penelitian ini yaitu *mean*, varians, nilai maksimum dan minimum. Rumus *mean* dapat dituliskan pada Persamaan (1) sebagai berikut.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \tag{1}$$

dimana, \bar{x} adalah rata-rata (*mean*), x_i adalah data ke- i dan n adalah banyaknya data. Sedangkan rumus varians dapat dituliskan pada Persamaan (2).

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \tag{2}$$

dimana s^2 adalah varians dari sampel data, n adalah banyaknya data, dan \bar{x} adalah rata-rata (*mean*).

B. Uji Korelasi

Korelasi merupakan suatu perhitungan untuk menentukan besar keeratan hubungan serta arah hubungan antara dua variabel. Indikator yang digunakan untuk melihat keeratan hubungan antara dua variabel yaitu melalui koefisien korelasi [6]. Korelasi yang digunakan untuk menghitung korelasi antara variabel kontinyu atau diskrit yaitu korelasi Pearson sebagaimana disajikan pada Persamaan (3).

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \tag{3}$$

Uji korelasi memiliki hipotesis sebagai berikut.

Hipotesis :

$H_0 : \rho_{x,y} = 0$ (Tidak ada hubungan antara kedua variabel)

$H_1 : \rho_{x,y} \neq 0$ (Terdapat hubungan antara kedua variabel)

Statistik uji :

$$t = \frac{r_{x,y} \sqrt{n-2}}{\sqrt{1-r^2}} \tag{4}$$

Keputusan yang didapatkan akan Tolak H_0 jika nilai $|t_{hitung}| > t_{\alpha/2;(n-2)}$ atau $p\text{-value} < \alpha$.

C. Multikolinearitas

Suatu keadaan dimana terdapat korelasi atau hubungan linier antara dua variabel prediktor atau lebih dinamakan multikolinearitas. Syarat asumsi yang harus dipenuhi untuk menyatakan bahwa suatu model regresi baik adalah tidak adanya gejala multikolinearitas antara variabel independen.

Konsekuensi dari adanya multikolinearitas yaitu pengujian secara parsial menjadi tidak signifikan meskipun pengujian secara serentak memberikan hasil yang signifikan, selain itu multikolinearitas menyebabkan model regresi terdapat standar eror yang besar namun dengan signifikansi yang kecil [7].

Salah satu cara untuk mendeteksi adanya gejala multikolinearitas adalah dengan melihat nilai *Variance Inflation Factor* (VIF) [7]. Multikolinearitas terjadi apabila nilai VIF lebih dari 10, VIF dihitung berdasarkan rumus sebagai berikut.

$$VIF_j = \frac{1}{1-R_j^2} \tag{5}$$

dimana R_j^2 adalah koefisien determinasi dari variabel independen yang diregresikan terhadap variabel dependen serta nilainya berada pada rentang 0 sampai 1 yang dihitung dari rumus sebagai berikut.

$$R_j^2 = \frac{\sum_{i=1}^n (x_{ji} - \hat{x}_{ji})^2}{\sum_{i=1}^n (x_{ji} - \bar{x}_j)^2} \tag{6}$$

D. Equidispersi

Equidispersi merupakan asumsi yang harus terpenuhi pada pemodelan regresi *Poisson* yaitu ketika nilai rata-rata dan varians pada variabel respon bernilai sama. Apabila asumsi equidispersi tidak terpenuhi dapat menyebabkan terjadinya overdispersi. Overdispersi merupakan suatu kondisi dimana pada variabel respon nilai varians melebihi nilai rata-rata. Overdispersi muncul ketika terjadi pelanggaran asumsi distribusi data, misal ketika data dikelompokkan sehingga melanggar asumsi observasi kebebasan *likelihood* [7]. Salah satu cara untuk mendeteksi keberadaan overdispersi adalah melihat nilai *The Variance Test* (VT). Rumus VT adalah sebagai berikut.

$$VT = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{\bar{y}} = (n-1) \frac{s^2}{\bar{y}} \tag{7}$$

Apabila nilai indeks dispersi < 1 maka terjadi underdispersi, sedangkan apabila nilai indeks dispersi > 1 maka terjadi overdispersi [8].

E. Regresi Zero-Inflated Poisson Inverse Gaussian

Regresi *Zero-Inflated Poisson Inverse Gaussian* merupakan suatu pemodelan regresi gabungan dari distribusi *Zero-Inflated Poisson* dan distribusi *Inverse Gaussian*. Fungsi kepadatan peluang $Y \sim ZIPIG(\lambda, \tau, p)$ dapat dituliskan pada Persamaan (9).

$$P(Y = y | \lambda, \tau, p) = \begin{cases} p + (1-p)P(Y = 0 | \tau, p) & ; y = 0 \\ (1-p)P(Y = y | \tau, p) & ; y = 1, 2, 3, \dots \end{cases} \tag{8}$$

dimana untuk $y=0$ dituliskan

$$P(Y = 0 | \lambda, \tau, p) = p + (1-p) e^{\frac{1}{\tau}} \left(\frac{2}{\pi \tau} \right)^{\frac{1}{2}} (2\lambda\tau + 1)^{\frac{1}{4}} \left(\frac{\pi}{2 \left(\frac{1}{\tau} \sqrt{(2\lambda\tau + 1)} \right)} \right)^{\frac{1}{2}} \exp\left(-\frac{1}{\tau} \sqrt{(2\lambda\tau + 1)} \right) \tag{9}$$

$$P(Y = 0 | \lambda, \tau, p) = p + (1-p) e^{\frac{1}{\tau}} \left(\frac{2}{\pi \tau} \right)^{\frac{1}{2}} (2\lambda\tau + 1)^{\frac{1}{4}} \left(\frac{\pi}{2 \left(\frac{1}{\tau} \sqrt{(2\lambda\tau + 1)} \right)} \right)^{\frac{1}{2}} \exp\left(-\frac{1}{\tau} \sqrt{(2\lambda\tau + 1)} \right)$$

Kemudian untuk $y= 1, 2, 3, \dots$ dapat dituliskan

$$P(Y = y | \lambda, \tau, p) = (1-p) \frac{\lambda^y e^{-\frac{1}{\tau}}}{y!} \left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}} (2\lambda\tau + 1)^{-\frac{(y-1)}{2}} \left(\frac{\pi}{2\left(\frac{1}{\tau}\sqrt{(2\lambda\tau + 1)}\right)}\right)^{y-\frac{1}{2}} \exp\left(-\frac{1}{\tau}\sqrt{(2\lambda\tau + 1)}\right) \quad (10)$$

Misalkan sebuah variabel respon $Y \sim \text{ZIPIG}(\lambda, \tau, p)$ maka model ZIPIGR dapat dituliskan pada dua komponen model, yaitu komponen untuk model *Poisson state* (λ) dan komponen model *zero inflated*. dituliskan:

Untuk model λ

$$\begin{aligned} \lambda &= \exp(\mathbf{x}^T \boldsymbol{\beta}) \\ \ln(\lambda) &= \mathbf{x}^T \boldsymbol{\beta} \end{aligned} \quad (11)$$

Untuk model *zero inflated*,

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = -\gamma \mathbf{x}^T \boldsymbol{\beta} \quad (12)$$

dengan mensubstitusikan kedua persamaan (12) dan (13) pada fungsi kepadatan peluang ZIPIG, diperoleh fungsi peluang untuk model regresi ZIPIG yang dapat dituliskan sebagai berikut.

Untuk $y=0$

$$\begin{aligned} P(Y = 0 | \boldsymbol{\beta}, \gamma, \tau) &= \frac{\exp(-\gamma \mathbf{x}^T \boldsymbol{\beta})}{1 + \exp(-\gamma \mathbf{x}^T \boldsymbol{\beta})} + \frac{1}{1 + \exp(-\gamma \mathbf{x}^T \boldsymbol{\beta})} e^{\frac{1}{\tau}} \left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}} (2 \exp(\mathbf{x}^T \boldsymbol{\beta})\tau + 1)^{\frac{1}{4}} \\ &\quad \left(\frac{\pi}{2\left(\frac{1}{\tau}\sqrt{(2 \exp(\mathbf{x}^T \boldsymbol{\beta})\tau + 1)}\right)}\right)^{\frac{1}{2}} \exp\left(-\frac{1}{\tau}\sqrt{(2 \exp(\mathbf{x}^T \boldsymbol{\beta})\tau + 1)}\right) \end{aligned} \quad (13)$$

Untuk $y=1,2,3,\dots$

$$\begin{aligned} P(Y = y | \boldsymbol{\beta}, \gamma, \tau) &= \frac{1}{1 + \exp(-\gamma \mathbf{x}^T \boldsymbol{\beta})} \frac{(\exp(\mathbf{x}^T \boldsymbol{\beta}))^y e^{-\frac{1}{\tau}}}{y!} \left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}} (2 \exp(\mathbf{x}^T \boldsymbol{\beta})\tau + 1)^{-\frac{(y-1)}{2}} \\ &\quad \left(\frac{\pi}{2\left(\frac{1}{\tau}\sqrt{(2 \exp(\mathbf{x}^T \boldsymbol{\beta})\tau + 1)}\right)}\right)^{y-\frac{1}{2}} \exp\left(-\frac{1}{\tau}\sqrt{(2 \exp(\mathbf{x}^T \boldsymbol{\beta})\tau + 1)}\right) \end{aligned} \quad (14)$$

F. Estimasi Parameter

Estimasi parameter regresi ZIPIG dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Model ZIPIG memiliki dua model yang saling bebas yakni model untuk variabel respon bernilai nol $Y_i = 0$ dan model untuk variabel respon lebih dari nol ($Y_i = 1, 2, \dots$). Oleh karena sifat saling bebas pada kedua model, persamaan likelihood untuk model ZIPIG dapat dituliskan sebagai berikut.

$$\begin{aligned} L(\boldsymbol{\beta}; \tau; \gamma) &= \left(\prod_{\substack{i=1 \\ y_i=0}}^n P(Y_i = y_i | \mathbf{x}_i; \boldsymbol{\beta}; \tau; \gamma)\right) \left(\prod_{\substack{i=1 \\ y_i>0}}^n P(Y_i = y_i | \mathbf{x}_i; \boldsymbol{\beta}; \tau; \gamma)\right) \\ L(\boldsymbol{\beta}; \tau; \gamma) &= (L_1(\boldsymbol{\beta}; \tau; \gamma))(L_2(\boldsymbol{\beta}; \tau; \gamma)) \end{aligned} \quad (15)$$

dimana

$$L_1(\boldsymbol{\beta}; \tau; \gamma) = \prod_{\substack{i=1 \\ y_i=0}}^n \frac{1}{1 + \exp(-\gamma \mathbf{x}_i^T \boldsymbol{\beta})} \left(\exp(-\gamma \mathbf{x}_i^T \boldsymbol{\beta}) + \exp\left(\frac{1}{\tau} - \frac{1}{\tau}\sqrt{2\tau \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + 1}\right)\right) \quad (16)$$

dan

$$L_2(\boldsymbol{\beta}; \tau; \gamma) = \prod_{i=1}^n \left\{ \frac{1}{1 + \exp(-\gamma \mathbf{x}_i^T \boldsymbol{\beta})} \right\} \left(\frac{\exp(\mathbf{x}_i^T \boldsymbol{\beta}) \exp\left(\frac{1}{\tau}\right)}{y_i!} \right) \left(\frac{2}{\pi\tau}\right)^{-y_i+1} (2\tau \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + 1)^{-y_i+\frac{1}{2}}$$

$$\exp\left(-\frac{1}{\tau} \sqrt{(2\tau \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + 1)}\right)$$
(17)

Kemudian kedua fungsi likelihood pada persamaan (16) dan (17), diubah menjadi bentuk ln likelihood. Fungsi ln likelihood untuk kedua fungsi dapat dituliskan sebagai berikut.

Untuk $L_1(\boldsymbol{\beta}; \tau; \gamma)$

$$l_1 = \ln L_1(\boldsymbol{\beta}; \tau; \gamma)$$

$$= \sum_{\substack{i=1 \\ y_i=0}}^n \ln \left(\exp(-\gamma \mathbf{x}_i^T \boldsymbol{\beta}) + \exp\left(\frac{1}{\tau} - \frac{1}{\tau} \sqrt{(2\tau \exp(\mathbf{x}_i^T \boldsymbol{\beta}) + 1)}\right) \right) - \sum_{\substack{i=1 \\ y_i=0}}^n \ln(1 + \exp(-\gamma \mathbf{x}_i^T \boldsymbol{\beta}))$$
(18)

Untuk $L_2(\boldsymbol{\beta}; \tau; \gamma)$

$$l_2 = \ln L_2(\boldsymbol{\beta}; \tau; \gamma)$$

$$= \sum_{\substack{i=1 \\ y_i > 0}}^n (1 - y_i) (\ln 2 - \ln \pi - \ln \tau) + \sum_{\substack{i=1 \\ y_i > 0}}^n y_i \mathbf{x}_i^T \boldsymbol{\beta} + \frac{n}{\tau} - \sum_{\substack{i=1 \\ y_i > 0}}^n \ln(1 + \exp(-\gamma \mathbf{x}_i^T \boldsymbol{\beta})) - \sum_{\substack{i=1 \\ y_i > 0}}^n \ln(y_i!)$$

$$+ \sum_{\substack{i=1 \\ y_i > 0}}^n \ln\left(\frac{1}{2} - y_i\right) \ln(2 \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \tau + 1) - \sum_{\substack{i=1 \\ y_i > 0}}^n \left(-\frac{1}{\tau} \sqrt{(2 \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \tau + 1)}\right)$$
(19)

Fungsi likelihood yang diperoleh kemudian diturunkan terhadap parameter $\boldsymbol{\beta}$, τ dan γ . Apabila hasil turunan pertama terhadap masing-masing parameter tidak dapat diselesaikan, estimasi dilanjutkan dengan algoritma Fisher Scoring dapat dituliskan pada persamaan sebagai berikut.

$$\hat{\boldsymbol{\theta}}_{(r+1)} = \hat{\boldsymbol{\theta}}_{(r)} + \mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}_{(m)}) \mathbf{D}(\hat{\boldsymbol{\theta}}_{(m)})$$
(20)

dimana

$$\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^T, \hat{\tau}, \hat{\gamma})^T$$

$$\mathbf{D}\hat{\boldsymbol{\theta}} = \left(\frac{\partial l}{\partial \hat{\gamma}}, \frac{\partial l}{\partial \hat{\tau}}, \frac{\partial l}{\partial \hat{\boldsymbol{\beta}}^T} \right)^T$$

$$\mathbf{I}(\hat{\boldsymbol{\theta}}_{(m)}) = -E[\mathbf{H}(\hat{\boldsymbol{\theta}}_{(m)})]$$

$$\mathbf{H}(\hat{\boldsymbol{\theta}}_{(m)})_{(k+1)(k+1)} = \begin{bmatrix} \frac{\partial^2 l}{\partial \hat{\gamma}^2} & \frac{\partial^2 l}{\partial \hat{\gamma} \partial \hat{\tau}} & \frac{\partial^2 l}{\partial \hat{\gamma} \partial \hat{\boldsymbol{\beta}}} \\ \frac{\partial^2 l}{\partial \hat{\tau} \partial \hat{\gamma}} & \frac{\partial^2 l}{\partial \hat{\tau}^2} & \frac{\partial^2 l}{\partial \hat{\tau} \partial \hat{\boldsymbol{\beta}}} \\ \frac{\partial^2 l}{\partial \hat{\boldsymbol{\beta}} \partial \hat{\gamma}} & \frac{\partial^2 l}{\partial \hat{\boldsymbol{\beta}} \partial \hat{\tau}} & \frac{\partial^2 l}{\partial \hat{\boldsymbol{\beta}} \partial \hat{\boldsymbol{\beta}}^T} \end{bmatrix}$$

$$\text{Sehingga } \mathbf{I}(\hat{\boldsymbol{\theta}}_{(m)}) = -E \begin{bmatrix} \frac{\partial^2 l}{\partial \hat{\gamma}^2} & \frac{\partial^2 l}{\partial \hat{\gamma} \partial \hat{\tau}} & \frac{\partial^2 l}{\partial \hat{\gamma} \partial \hat{\boldsymbol{\beta}}} \\ \frac{\partial^2 l}{\partial \hat{\tau} \partial \hat{\gamma}} & \frac{\partial^2 l}{\partial \hat{\tau}^2} & \frac{\partial^2 l}{\partial \hat{\tau} \partial \hat{\boldsymbol{\beta}}} \\ \frac{\partial^2 l}{\partial \hat{\boldsymbol{\beta}} \partial \hat{\gamma}} & \frac{\partial^2 l}{\partial \hat{\boldsymbol{\beta}} \partial \hat{\tau}} & \frac{\partial^2 l}{\partial \hat{\boldsymbol{\beta}} \partial \hat{\boldsymbol{\beta}}^T} \end{bmatrix}$$

Matriks hessian berisi turunan kedua dari fungsi likelihood terhadap parameter $\boldsymbol{\beta}$, τ , dan γ . Langkah-langkah Fisher Scoring Algorithm sebagai berikut.

1. Menentukan vektor awal parameter θ_0 dengan mengasumsikan data memenuhi model regresi linier berganda: $Y^* = \beta_{0(0)} + x_{i1}\beta_{1(0)} + \dots + x_{ip}\beta_{p(0)} + \varepsilon_i$ dimana $i = 1, 2, \dots, n$ dan dengan metode kuadrat terkecil diperoleh $\hat{\beta}_{(0)} = (X^T X)^{-1} X^T Y^*$
2. Membentuk vektor gradien $D(\hat{\theta})$.
3. Membentuk matriks hessian $H(\hat{\theta}_{(0)})$.
4. Membentuk matriks informasi Fisher $I(\hat{\theta}_{(0)})$.
5. Memasukkan nilai $\hat{\theta}_{(0)}$ sehingga diperoleh vektor gradien $D(\hat{\theta}_{(0)})$ dan matriks hessian $H(\hat{\theta}_{(0)})$.
6. Mulai dari $m = 0$ dilakukan iterasi pada $\hat{\theta}_{(r+1)} = \hat{\theta}_{(r)} + I^{-1}(\hat{\theta}_{(m)})D(\hat{\theta}_{(m)})$, nilai $\hat{\theta}_{(m)}$ merupakan sekumpulan penaksir parameter yang konvergen saat iterasi ke- m .
7. Jika belum diperoleh penaksiran parameter yang konvergen saat iterasi ke- m , maka dilanjutkan kembali ke langkah 5 hingga iterasi ke- $m+1$. Iterasi akan berhenti apabila nilai dari $\|\hat{\theta}_{(m+1)} - \hat{\theta}_{(m)}\| \leq \varepsilon$ dan $\varepsilon > 0$ adalah bilangan yang sangat kecil.

Fisher Scoring Algorithm digunakan di dalam algoritma yang digunakan dalam *Package Gamlss* pada *software R-studio* yaitu algoritma *Rigby Stasinopoulos (RS Algorithm)*. *RS Algorithm* merupakan generalisasi dari algoritma yang digunakan *Rigby dan Stasinopoulos (1996)* dalam melakukan *fitting* terhadap *Mean and Dispersion Additive Models (MADAM)*.

G. Pengujian Parameter

Pengujian hipotesis secara serentak pada parameter model ZIPIGR dilakukan dengan menggunakan metode *Maximum Likelihood Ratio Test (MLRT)*. Pengujian hipotesis secara serentak yaitu sebagai berikut.

Hipotesis :

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

$H_1 : \text{minimal ada satu } \beta_j \neq 0, \text{ dengan } j = 1, 2, \dots, k$

Statistik uji :

$$G = -2 \ln \left(\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \tag{21}$$

G berdistribusi $\chi^2_{(v)}$, dengan taraf signifikansi α sedangkan v menunjukkan jumlah parameter dibawah populasi dikurangi jumlah parameter dibawah H_0 dengan rumus v yaitu $v = n(\Omega) - n(\omega)$. Pengujian parameter terjadi tolak H_0 apabila nilai $G > \chi^2_{(v, \alpha)}$.

Parameter model regresi *Zero-Inflated Poisson Inverse Gaussian* yang telah dihasilkan dari penaksiran parameter belum tentu mempunyai pengaruh yang signifikan terhadap model secara parsial, sehingga diperlukan uji parsial untuk parameter β , parameter τ , dan parameter γ dengan hipotesis sebagai berikut.

- Parameter β

Hipotesis :

$H_0 : \beta_j = 0$

$H_1 : \beta_j \neq 0 \text{ dengan } j = 1, 2, \dots, k$

-Parameter τ

Hipotesis :

$H_0 : \tau = 0$

$H_1 : \tau \neq 0$

-Parameter γ

Hipotesis :

$H_0 : \gamma = 1$

$H_1 : \gamma \neq 1$

Statistik uji untuk uji parsial adalah sebagai berikut :

$$Z = \frac{\hat{\theta}_j}{se(\hat{\theta}_j)} \tag{22}$$

dimana $se(\hat{\theta}_j)$ merupakan standar error dari $\hat{\theta}_j$, sedangkan θ merupakan parameter yang diuji. Nilai $se(\hat{\theta}_j)$ didapatkan dari elemen diagonal yang bersesuaian dari invers matriks informasi Fisher. Daerah kritis uji coba adalah tolak H_0 jika $|Z| > Z_{(\alpha/2)}$, dimana α adalah tingkat signifikansi.

H. Kriteria Keباikan Model

Pemodelan regresi ZIPIG dilakukan menggunakan package dari software R yaitu GAMLSS (*Generalized Additive Model for Location and Shape*) berdasarkan metode *stepwise*. Pemilihan model terbaik ZIPIG dapat dilihat berdasarkan nilai AICc. *Akaike Information Criterion* (AIC) pada saat $(n/p) < 40$. Kriteria pengukuran AICc digunakan untuk mengatasi potensi *overfitting* pada AIC. Nilai AICc dapat dihitung berdasarkan rumus sebagai berikut.

$$AICc = AIC + \frac{2p(p+1)}{n-p-1} \tag{23}$$

$$= -2\log(L(\hat{\beta})) + 2p + \frac{2p(p+1)}{n-p-1}$$

dimana L merupakan maksimum *likelihood*, n merupakan jumlah unit penelitian, dan k merupakan banyaknya parameter pada model. Suatu model merupakan model terbaik apabila diantara seluruh model, model tersebut memiliki nilai AICc terkecil [9].

Namun selain berdasarkan nilai AICc yang terkecil, model terbaik juga harus memperhatikan parameter dispersi τ yang signifikan. Parameter dispersi τ didapatkan dari GAMLSS pada metode *stepwise*, dimana munculnya parameter ini dikarenakan tidak terjadinya kondisi *equidispersi* pada data. Bagi data yang mengalami *overdispersi* maka nilai dari parameter dispersi adalah $\tau > 0$, dimana parameter dispersi τ signifikan ketika *p-value* bernilai kurang dari taraf signifikansi α [10].

Selanjutnya pemilihan model regresi ZIPIG terbaik juga dilihat berdasarkan persentase ketepatan klasifikasi tertinggi. Nilai ketepatan klasifikasi dihitung berdasarkan *Apparent Error Rate* (APER) menggunakan tabel ketepatan klasifikasi sebagai berikut.

Tabel 1 Ketepatan Klasifikasi

Aktual (Y)	Prediksi (Y)		Total
	0	1	
0	n_{00}	n_{01}	$n_{0.}$
1	n_{10}	n_{11}	$n_{1.}$
Total	$n_{.0}$	$n_{.1}$	n

Nilai persentase ketepatan klasifikasi diperoleh berdasarkan persamaan sebagai berikut.

$$\text{Ketepatan Klasifikasi} = (1 - \text{APER}) \times 100\% \tag{24}$$

$$= \left(1 - \left(\frac{n_{01} + n_{10}}{n} \right) \right) \times 100\%$$

Suatu model dikatakan model terbaik apabila diantara seluruh model, model tersebut memiliki persentase ketepatan klasifikasi tertinggi. Sehingga dalam pemilihan model regresi ZIPIG terbaik yaitu berdasarkan nilai parameter dispersi (τ) yang signifikan, *Akaike Information Criterion Correction* (AICc) terkecil, dan persentase ketepatan klasifikasi tertinggi.

I. Penyakit Kusta

Penyakit kusta atau disebut juga dengan lepra merupakan penyakit yang menyerang berbagai bagian tubuh diantaranya saraf tepi, kulit, dan jaringan tubuh lainnya kecuali susunan saraf pusat [11]. Penyakit kusta terdiri dari dua tipe yaitu *Paucibacillary* (PB) atau kusta kering serta *Multibacillary* (MB) atau kusta basah [12].

Penyebab dari penyakit kusta yaitu bakteri *Mycobacterium leprae* yang dapat menular langsung kepada manusia melalui kontak langsung dengan penderita (kontak yang lama dan berulang-ulang) serta melalui pernafasan.

Perkembangbiakan bakteri kusta mengalami proses selama 2-3 minggu, dalam tubuh manusia pertahanan bakteri yaitu selama 9 hari, kemudian kuman membelah dalam jangka 14-21 hari dengan masa inkubasi rata-rata dua hingga lima tahun. Penyakit ini merupakan tipe penyakit *granulomatosa* pada saraf tepi dan mukosa dari saluran pernafasan atas dan lesi pada kulit. Kusta dapat menjadi sangat progresif apabila tidak segera ditangani dapat menyebabkan kerusakan pada kulit, saraf-saraf, dan anggota gerak mata [1]. Pedoman klasifikasi jenis kusta menurut WHO ditampilkan sebagai berikut [13].

Tabel 1 Tanda Utama Kusta (PB) dan Kusta Basah (MB)

Tanda Utama	Kusta Kering (PB)	Kusta Basah (MB)
Jumlah bercak kusta	1-5	>5
Penebalan saraf tepi disertai gangguan fungsi	Hanya 1 saraf	>1 saraf
Distribusi	Unilateral atau bilateral	Bilateral simetris
Konsistensi	Kasar	Halus dan berkilat
Kerokan jaringan kulit	Basil Tahan Asam (BTA) negatif	Basil Tahan Asam (BTA) positif

III. METODOLOGI PENELITIAN

A. Sumber Data

Penelitian ini menggunakan sumber data sekunder yang diperoleh dari Dinas Kesehatan Kabupaten Mojokerto, Kabupaten Jombang tahun 2019 serta Badan Pusat Statistik Kabupaten Mojokerto, Kabupaten Jombang. Tingkat observasi pada penelitian ini yaitu data tingkat kecamatan, dengan unit penelitian yaitu 39 kecamatan di Kabupaten Mojokerto dan Kabupaten Jombang, Provinsi Jawa Timur.

B. Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri atas variabel respon dan variabel prediktor sebagai berikut.

Tabel 3 Variabel Penelitian

Variabel	Keterangan	Skala Data
Y	Jumlah kasus kusta	Rasio
X ₁	Persentase penduduk dengan sanitasi layak	Rasio
X ₂	Rasio tenaga medis	Rasio
X ₃	Persentase sarana air minum memenuhi syarat kesehatan	Rasio
X ₄	Persentase tempat-tempat umum memenuhi syarat kesehatan	Rasio
X ₅	Persentase pelayanan kesehatan untuk penduduk usia produktif	Rasio
X ₆	Persentase pelayanan kesehatan untuk penduduk usia lanjut	Rasio

C. Langkah Analisis

Langkah-langkah analisis dalam penelitian ini dituliskan sebagai berikut.

- Melakukan analisis statistika deskriptif pada data sesuai dengan persamaan 1 dan 2.
- Menghitung korelasi antara variabel prediktor dan variabel respon sesuai dengan persamaan 3.
- Mendeteksi kasus multikolinearitas dari variabel prediktor menggunakan VIF sesuai dengan persamaan 5.
- Melakukan deteksi overdispersi dengan *variance test* sesuai dengan persamaan 7.
- Melakukan analisis data menggunakan Regresi ZIPIG.
 - Menentukan nilai penaksir parameter pada model menggunakan MLE untuk beberapa kombinasi model dengan metode *stepwise* sesuai dengan persamaan 15 sampai dengan persamaan 20.
 - Melakukan pengujian hipotesis model ZIPIG secara serentak sesuai dengan persamaan 21.
 - Melakukan pengujian hipotesis model ZIPIG secara parsial sesuai dengan persamaan 22.
 - Menentukan model terbaik dari regresi ZIPIG dengan membandingkan nilai AICc setiap model, nilai parameter dispersi yang signifikan, serta persentase ketepatan klasifikasi tertinggi sesuai dengan persamaan 23 dan 24.
 - Melakukan interpretasi model.
- Menarik kesimpulan dan memberikan saran.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Data

Analisis karakteristik data pada penelitian ini dilakukan menggunakan statistika deskriptif untuk masing-masing variabel prediktor serta variabel respon.

Tabel 4 Analisis Deskriptif

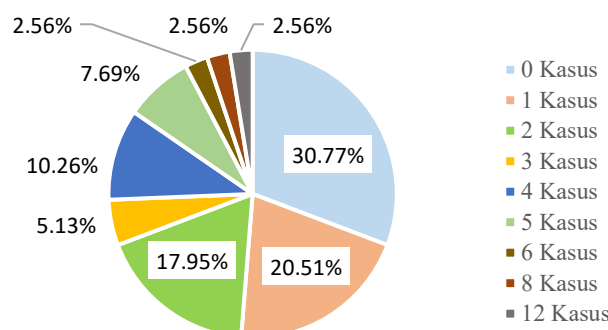
Variabel	Mean	Varians	Minimum	Maksimum
Jumlah kasus kusta (Y)	2,179	6,625	0,000	12,000
Persentase penduduk dengan sanitasi layak (X ₁)	90,214	99,981	62,100	100,000
Rasio tenaga medis (X ₂)	23,769	1080,164	2,924	176,415
Persentase sarana air minum memenuhi syarat kesehatan (X ₃)	81,891	783,654	0,000	100,000
Persentase tempat-tempat umum memenuhi syarat kesehatan (X ₄)	72,941	221,668	39,400	100,000
Persentase pelayanan kesehatan untuk penduduk usia produktif (X ₅)	38,875	414,578	1,900	98,000

Berdasarkan Tabel 4 dapat diketahui bahwa nilai rata-rata variabel jumlah kasus kusta (Y) untuk 39 kecamatan di Kabupaten Mojokerto dan Kabupaten Jombang sebesar 2,179 dengan varians sebesar 6,625. Nilai varians yang lebih tinggi daripada nilai rata-rata mengindikasikan adanya overdispersi pada data.

Variabel prediktor X_1 memiliki nilai *mean* tertinggi dibandingkan variabel-variabel prediktor lainnya yaitu sebesar 90,214. Nilai ini dapat diartikan rata-rata persentase penduduk dengan sanitasi layak pada setiap kecamatan di Kabupaten Mojokerto dan Kabupaten Jombang sangat tinggi yaitu sebesar 90,214 %. Sedangkan variabel prediktor dengan nilai *mean* terendah yaitu X_2 sebesar 23,769 , hal ini dapat diartikan rata-rata tenaga medis untuk setiap 100.000 penduduk pada setiap kecamatan di Kabupaten Mojokerto dan Kabupaten Jombang yaitu sebanyak 24 tenaga medis.

Selanjutnya diketahui bahwa variabel prediktor X_1 memiliki varians terendah jika dibandingkan variabel-variabel prediktor lainnya sebesar 99,981. Nilai ini dapat diartikan tingkat keragaman persentase penduduk dengan sanitasi layak di Kabupaten Mojokerto dan Kabupaten Jombang yaitu sebesar 99,981. Sebaliknya untuk variabel prediktor dengan nilai varians tertinggi yaitu variabel X_6 sebesar 2786,376 ,artinya tingkat keragaman persentase pelayanan kesehatan untuk penduduk usia lanjut di Kabupaten Mojokerto dan Kabupaten Jombang sangat beragam yaitu sebesar 2786,376. Nilai varians yang tinggi pada variabel X_6 dapat terjadi dikarenakan terdapat kesenjangan yang cukup besar antara pelayanan kesehatan untuk penduduk usia lanjut antara kecamatan yang satu dengan yang lainnya. Semakin besar nilai varians menandakan data yang semakin beragam, namun jika semakin kecil varians maka dapat diartikan data semakin homogen.

Selain itu pada Tabel 4 diketahui pula variabel prediktor dengan nilai minimum terendah yaitu persentase tempat-tempat umum yang memenuhi syarat kesehatan (X_4) sebesar 0%, artinya terdapat kecamatan dengan tempat umum yang belum memenuhi syarat kesehatan. Apabila ditinjau berdasarkan nilai maksimum diketahui bahwa variabel dengan nilai maksimum tertinggi bahkan melebihi seratus persen yaitu variabel persentase pelayanan kesehatan untuk penduduk usia lanjut (X_6). Fenomena yang terjadi pada variabel X_6 dapat terjadi dikarenakan, masyarakat usia lanjut pada kecamatan dengan persentase X_6 melebihi 100% lebih terpantau kesehatannya oleh petugas kesehatan, hal ini dapat terjadi dikarenakan masyarakat usia lanjut tersebut tinggal bersama dengan anak-anak nya sehingga pendataan oleh petugas kesehatan menjadi lebih mudah dibandingkan masyarakat usia lanjut di kecamatan-kecamatan lainnya dengan persentase di bawah 100%.



Gambar 1 Jumlah Kasus Kusta di Kabupaten Mojokerto dan Kabupaten Jombang

Visualisasi pada Gambar 1 memberikan informasi bahwa sebanyak 30,77% atau 12 kecamatan di Kabupaten Mojokerto dan Kabupaten Jombang tidak terdapat kasus kusta, sedangkan 27 kecamatan lainnya terdapat kasus kusta. Sebanyak 8 kecamatan terdapat kasus kusta sebesar 1 kasus yaitu pada kecamatan Sooko, Bangsal, Gedeg, Perak, Wonosalam, Tembelang, Kudu, Ngusikan. Adapun kecamatan dengan jumlah kasus tertinggi yaitu sebanyak 12 kasus terdapat di Kecamatan Mojowarno, Kabupaten Jombang.

B. Korelasi

Indikator untuk melihat keeratan hubungan antara dua variabel yaitu melalui koefisien korelasi *Pearson* sesuai dengan persamaan 2.3. Koefisien korelasi pada masing-masing variabel disajikan pada Tabel 4.2 sebagai berikut.

Tabel 5 Koefisien Korelasi Variabel Respon dan Variabel Prediktor

Korelasi	Y	X_1	X_2	X_3	X_4	X_5
X_1	0,130					
X_2	0,119	0,124				
X_3	-0,176	0,058	0,082			
X_4	-0,035	-0,192	0,182	0,151		
X_5	-0,313	0,162	-0,122	0,490	0,046	
X_6	-0,317	-0,053	-0,087	0,459	0,178	0,292

Berdasarkan Tabel 5. Diketahui bahwa terdapat hubungan antara variabel prediktor dan variabel respon. Variabel persentase penduduk dengan sanitasi layak (X₁), rasio tenaga medis (X₂) memiliki hubungan yang positif terhadap jumlah kasus kusta (Y).

Apabila ditinjau berdasarkan nilai koefisien korelasi secara keseluruhan pada Tabel 4.2 diketahui bahwa hubungan antara variabel persentase sarana air minum memenuhi syarat kesehatan (X₃) dengan variabel pelayanan kesehatan untuk penduduk usia produktif (X₅) memiliki koefisien korelasi tertinggi sebesar 0,490. Sedangkan untuk nilai koefisien korelasi terendah terdapat pada hubungan antara variabel persentase tempat-tempat umum memenuhi syarat kesehatan (X₄) dan variabel jumlah kasus kusta (Y) sebesar -0,035.

C. Pemeriksaan Multikolinearitas

Multikolinearitas merupakan suatu kondisi ketika terdapat korelasi atau hubungan linier antara dua variabel prediktor atau lebih. Multikolinearitas dapat dideteksi berdasarkan nilai VIF sesuai dengan persamaan 2.44 dan didapatkan hasil sebagai berikut.

Tabel 6 Nilai VIF Variabel Prediktor

Variabel	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆
VIF	1,116	1,138	1,604	1,136	1,419	1,333

Berdasarkan Tabel 4.3 diketahui bahwa nilai VIF untuk seluruh variabel prediktor bernilai kurang dari 10, hal ini dapat diartikan bahwa tidak terdapat multikolinearitas antara variabel prediktor serta data telah memenuhi asumsi regresi *Zero-Inflated Poisson Inverse Gaussian*.

D. Pemeriksaan Equidispersi

Variabel respon pada data penelitian diketahui memiliki nilai rata-rata sebesar 2,179 dan nilai varians sebesar 6,625, sehingga dapat dikatakan data mengalami overdispersi. Namun perlu dilakukan pemeriksaan equidispersi lebih lanjut berdasarkan nilai *variance to mean ratio* (VT) atau indeks dispersi sesuai dengan persamaan 7 .

Tabel 7 Pemeriksaan Equidispersi

Varians	Rata-Rata	VT
6,625	2,179	115,506

Berdasarkan Tabel 7. diketahui bahwa untuk indeks dispersi atau nilai VT bernilai sebesar 115,506, dikarenakan nilai ini lebih besar dari 1 maka dapat dikatakan terjadi overdispersi pada data.

E. Pemodelan Regresi ZIPIG

Pemodelan regresi ZIPIG dilakukan pada 6 variabel prediktor dengan metode stepwise mengacu pada persamaan 15 sampai dengan persamaan 20. Hasil yang didapatkan berupa enam kombinasi kemungkinan model regresi yang konvergen sebagai berikut.

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i})$$

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_5 x_{5i} + \beta_6 x_{6i})$$

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_5 x_{5i})$$

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_3 x_{3i} + \beta_5 x_{5i})$$

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_5 x_{5i})$$

$$\lambda_i = \exp(\beta_0 + \beta_5 x_{5i})$$

Penaksiran parameter dari setiap kemungkinan model yang telah didapatkan disajikan pada Tabel 6. sebagai berikut.

Tabel 8 Estimasi Parameter Kemungkinan Model Regresi ZIPIG

Variabel	β ₀	β ₁	β ₂	β ₃	β ₄	β ₅	β ₆	τ	γ
X ₁ X ₂ X ₃ X ₄ X ₅ X ₆	-1,079	0,031	-0,004	0,010	0,005	-0,027	-0,013	0,657	-32,580
X ₁ X ₂ X ₃ X ₅ X ₆	-0,830	0,030	-0,002	0,011		-0,027	-0,012	0,326	-36,040
X ₁ X ₂ X ₃ X ₅	-0,902	0,027	0,000	0,004		-0,025		-1,019	-1,741
X ₁ X ₃ X ₅	-0,884	0,027		0,004		-0,025		-1,092	-1,702
X ₁ X ₅	-0,642	0,026				-0,021		-1,389	-1,613
X ₅	1,681					-0,024		-0,218	-2,765

F. Pengujian Parameter Secara Serentak

Pengujian parameter secara serentak dilakukan untuk memeriksa peran koefisien β secara keseluruhan menggunakan statistik uji devians (G) atau *likelihood ratio*. Nilai statistik uji G akan dibandingkan dengan nilai $\chi^2_{(\alpha, v)}$ dengan taraf signifikansi sebesar α = 0,1. Pengujian parameter secara serentak pada setiap kemungkinan model disajikan pada Tabel 9 sebagai berikut.

Tabel 9 Pengujian Parameter Secara Serentak

Variabel dari Model	Statistik G	$\chi^2_{(\alpha, v)}$	Keputusan
X ₁ X ₂ X ₃ X ₄ X ₅ X ₆	150,098	31	41,422
X ₁ X ₂ X ₃ X ₅ X ₆	148,326	32	42,585
X ₁ X ₂ X ₃ X ₅	147,955	33	43,745
X ₁ X ₃ X ₅	148,012	34	44,903
X ₁ X ₅	148,668	35	46,059
X ₅	150,328	36	47,212

Berdasarkan Tabel 9 dapat diketahui bahwa setiap kemungkinan model memiliki nilai statistik G lebih besar dari $\chi^2_{(\alpha, v)}$ sehingga diambil keputusan Tolak H₀. Hasil ini dapat diartikan untuk masing-masing model terdapat minimal satu variabel yang berpengaruh signifikan terhadap model ZIPIG.

G. Pengujian Parameter Secara Parsial

Pengujian parameter secara parsial dilakukan untuk mencari variabel prediktor yang berpengaruh signifikan terhadap model. Pengujian parameter secara parsial pada setiap kemungkinan model disajikan sebagai berikut.

Tabel 10 Pengujian Parameter Secara Parsial

Model	Variabel dari Model	Parameter Signifikan
1	X ₁ X ₂ X ₃ X ₄ X ₅ X ₆	β_6, τ
2	X ₁ X ₂ X ₃ X ₅ X ₆	β_6
3	X ₁ X ₂ X ₃ X ₅	β_5, γ
4	X ₁ X ₃ X ₅	β_5, γ
5	X ₁ X ₅	β_5, γ
6	X ₅	β_5

Berdasarkan Tabel 10 diketahui bahwa terdapat 1 model dengan parameter τ yang signifikan pada α sebesar 0,1 yaitu model 1. Sedangkan model dengan parameter γ yang signifikan yaitu model 4, model 5, dan model 6. Parameter τ merupakan parameter dispersi yang menandakan adanya overdispersi. Sedangkan parameter γ nantinya digunakan dalam penentuan model logit regresi ZIPIG.

H. Pemilihan Model Terbaik

Hasil pemodelan regresi ZIPIG didapatkan enam kemungkinan model. Selanjutnya dari enam kemungkinan model tersebut dilakukan pemilihan model regresi ZIPIG terbaik berdasarkan nilai *Akaike Information Criterion Correction* (AICc) terkecil, kemudian berdasarkan adanya parameter dispersi (τ) yang signifikan dalam model ZIPIG, serta berdasarkan nilai persentase ketepatan klasifikasi tertinggi.

Tabel 11 Pemilihan Model Terbaik

Model	Variabel dari Model	AICc	τ	P-value τ	Ketepatan Klasifikasi
1	X ₁ X ₂ X ₃ X ₄ X ₅ X ₆	172,898	0,657	0,090	69,23%
2	X ₁ X ₂ X ₃ X ₅ X ₆	167,939	0,326	0,414	66,67%
3	X ₁ X ₂ X ₃ X ₅	164,580	-1,019	0,198	33,33%
4	X ₁ X ₃ X ₅	161,830	-1,092	0,175	33,33%
5	X ₁ X ₅	159,844	-1,389	0,124	33,33%
6	X ₅	159,013	-0,218	0,784	64,10%

Berdasarkan Tabel 11 didapatkan hasil bahwa model 6 dengan variabel prediktor X₅ memiliki nilai AICc terkecil sebesar 159,013, namun untuk model ini memiliki nilai τ yang tidak signifikan ditandai dengan p-value τ bernilai lebih besar dari tingkat signifikansi α 0,1. Selain itu dikarenakan parameter τ pada model 6 bernilai negatif nilai ini menandakan underdispersi, padahal jika dilihat berdasarkan pada Tabel 4.4 seharusnya model menggambarkan terjadinya overdispersi pada data, maka untuk model 6 tidak dipilih sebagai model terbaik.

Selanjutnya apabila ditinjau berdasarkan nilai τ yang signifikan hanya terdapat pada model 1 yang terdiri atas variabel prediktor X₁, X₂, X₃, X₄, X₅, X₆ dengan nilai parameter τ sebesar 0,657 dan AICc sebesar 172,898. Kemudian apabila ditinjau berdasarkan ketepatan klasifikasi tertinggi terdapat pula pada model 1 dengan persentase sebesar 69,23%.

Sehingga untuk penelitian ini, model yang dipilih sebagai model terbaik yaitu model 1 yang terdiri atas variabel prediktor X₁, X₂, X₃, X₄, X₅, X₆, hal ini dikarenakan pada model ini meskipun memiliki nilai AICc yang tidak cukup rendah, model ini memiliki nilai τ yang signifikan. Nilai parameter τ sebesar 0,657 menandakan adanya overdispersi dikarenakan nilai ini lebih besar dari 0. Selain itu model 1 memiliki persentase ketepatan klasifikasi tertinggi dibandingkan model lainnya. Hasil estimasi parameter model regresi ZIPIG terbaik ditampilkan pada Tabel 12 sebagai berikut.

Tabel 12 Penaksiran Parameter Model Regresi ZIPIG

Parameter	Taksiran	Standar Error	P-value
β_0	-1,079	4,328	0,805
β_1	0,031	0,036	0,393
β_2	-0,004	0,012	0,762
β_3	0,010	0,017	0,549
β_4	0,005	0,023	0,815
β_5	-0,027	0,023	0,243
β_6	-0,013	0,007	0,054*
τ	0,657	0,377	0,090*
γ	-32,580	16012,820	0,998

Berdasarkan hasil yang didapatkan pada Tabel 12 dapat dilihat untuk parameter γ (*zero-inflated*) tidak signifikan maka hal ini berarti nilai $\gamma = 1$. Sehingga model regresi ZIPIG untuk jumlah kasus kusta di Kabupaten Mojokerto dan Kabupaten Jombang adalah sebagai berikut.

a. Model *Poisson state* untuk $\hat{\lambda}$ pada jumlah kasus kusta.

$$\ln(\hat{\lambda}_i) = -1,079 + 0,031x_{1i} - 0,004x_{2i} + 0,010x_{3i} + 0,005x_{4i} - 0,027x_{5i} - 0,013x_{6i}$$

dimana $i = 1, 2, \dots, 91$.

Hasil pada model diketahui bahwa variabel yang berpengaruh signifikan terhadap jumlah kasus kusta yaitu variabel variabel persentase pelayanan kesehatan untuk penduduk usia lanjut (X_6). Sedangkan untuk variabel-variabel lainnya tidak signifikan pada model.

Interpretasi dari model $\ln(\hat{\lambda}_i)$ yaitu untuk variabel X_6 dengan koefisien sebesar -0,013 dapat diartikan bahwa setiap kenaikan persentase pelayanan kesehatan untuk penduduk usia lanjut sebesar 1% akan menurunkan rata-rata jumlah kasus kusta sebesar $\exp(-0,013) = 0,987$ kali dengan asumsi variabel lain bernilai konstan.

b. Model logit pada jumlah kasus kusta.

$$\begin{aligned} \logit(\hat{p}_i) &= -1(-1,079 + 0,031x_{1i} - 0,004x_{2i} + 0,010x_{3i} + 0,005x_{4i} - 0,027x_{5i} - 0,013x_{6i}) \\ &= 1,079 - 0,031x_{1i} + 0,004x_{2i} - 0,010x_{3i} - 0,005x_{4i} + 0,027x_{5i} + 0,013x_{6i} \end{aligned}$$

dimana $i = 1, 2, \dots, 91$.

Hasil pada model logit regresi ZIPIG diketahui bahwa peluang respon bernilai nol dipengaruhi oleh koefisien *zero-inflated* (γ) sebesar negatif satu pada masing-masing variabel di dalam model. Variabel yang signifikan pada model

$\logit(\hat{p}_i)$ yaitu variabel persentase pelayanan kesehatan untuk penduduk usia lanjut (X_6). Sedangkan untuk variabel-variabel lainnya tidak signifikan pada model.

Berdasarkan model $\logit(\hat{p}_i)$ diinterpretasikan bahwa setiap kenaikan persentase pelayanan kesehatan untuk penduduk usia lanjut (X_6) sebesar 1% maka akan menaikkan peluang tidak terdapat kasus kusta sebesar $\exp(0,013) = 1,013$ kali jika dibandingkan adanya kasus kusta dengan menganggap variabel lainnya bernilai konstan.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Kesimpulan yang diperoleh berdasarkan analisis dan pembahasan yaitu sebagai berikut.

1. Kabupaten Mojokerto dan Kabupaten Jombang pada tahun 2019 memiliki rata-rata jumlah kasus kusta sebesar 2,179 dan varians sebesar 6,625, nilai varians yang lebih besar dari nilai rata-rata mengindikasikan adanya overdispersi. Selain itu persentase nilai nol pada kasus kusta sebesar 30,77% mengindikasikan adanya *extra zeros*. Variabel prediktor dengan nilai rata-rata tertinggi yaitu persentase penduduk dengan sanitasi layak. Sedangkan variabel prediktor dengan varians tertinggi yaitu persentase pelayanan kesehatan untuk penduduk usia lanjut, nilai varians yang tinggi dapat terjadi disebabkan pada kecamatan tertentu memiliki persentase yang sangat tinggi.
2. Hasil pemodelan terbaik dengan regresi ZIPIG adalah sebagai berikut.

a. Model *Poisson state* untuk $\hat{\lambda}$ pada jumlah kasus kusta.

$$\ln(\hat{\lambda}_i) = -1,079 + 0,031x_{1i} - 0,004x_{2i} + 0,010x_{3i} + 0,005x_{4i} - 0,027x_{5i} - 0,013x_{6i}$$

b. Model logit pada jumlah kasus kusta.

$$\begin{aligned} \logit(\hat{p}_i) &= -1(-1,079 + 0,031x_{1i} - 0,004x_{2i} + 0,010x_{3i} + 0,005x_{4i} - 0,027x_{5i} - 0,013x_{6i}) \\ &= 1,079 - 0,031x_{1i} + 0,004x_{2i} - 0,010x_{3i} - 0,005x_{4i} + 0,027x_{5i} + 0,013x_{6i} \end{aligned}$$

Berdasarkan pemodelan menggunakan regresi ZIPIG, diperoleh hasil bahwa faktor yang berpengaruh signifikan terhadap jumlah kasus kusta yaitu variabel persentase pelayanan kesehatan untuk penduduk usia lanjut (X_6).

B. Saran

Saran bagi Dinas Kesehatan Kabupaten Mojokerto dan Dinas Kesehatan Kabupaten Jombang yaitu sebaiknya lebih memperhatikan variabel prediktor yang berpengaruh secara signifikan pada model ZIPIG yaitu persentase pelayanan kesehatan untuk penduduk usia lanjut sehingga diharapkan jumlah kasus kusta kedepannya dapat ditekan seminimal mungkin.

REFERENSI

- [1] Kemenkes RI. (2018). Hapuskan Stigma dan Diskriminasi Terhadap Kusta. *Pusat Data dan Informasi Kementerian Kesehatan Republik Indonesia*, 2.
- [2] Dinas Kesehatan Provinsi Jawa Timur. (2019). *Profil Kesehatan Provinsi Jawa Timur*. Surabaya: Dinas Kesehatan Provinsi Jawa Timur.
- [3] Andrian, R. N. (2015). Penularan dan Penyembuhan Penyakit Kusta pada Keluarga Anak Mantan Penderita Kusta di Dusun Sumberglagah, Desa Tanjungkkenongo, Kecamatan Pacet. *Departemen Antropologi, Fakultas Ilmu Sosial dan Ilmu Politik, Universitas Airlangga*.
- [4] Hilbe, J. (2007). *Negative Binomial Regression* (1st ed.). New York: Cambridge University Press.
- [5] Rahayuning, H. A. (2020). *Pemodelan Jumlah Kematian Ibu Nifas di Karesidenan Pekalongan Provinsi Jawa Tengah Tahun 2017 Menggunakan Regresi Zero-Inflated Poisson Inverse Gaussian*. Surabaya.
- [6] Walpole, R. (1995). *Pengantar Metode Statistika Edisi Ketiga. Diterjemahkan: Bambang Sumantri*. Jakarta: PT Gramedia Pustaka Utama.
- [7] Gujarati, D. N. (2003). *Basic Economics (Fourth Edition)*. New York: McGraw-Hill Inc.
- [8] Karlis, D., & Xekalaki, E. (2000). A Simulation Comparison of Several Procedures for Testing the *Poisson* Assumption. *The Statistician*, 355-382.
- [9] Burnham, K., & Anderson, D. (2002). *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach*. New York: Springer-Verlag.
- [10] Wang, W., & Famoye, F. (1997). Modelling Household Fertility Decision with Generalized *Poisson* Regression. *Journal of Population Economics*, 273-283.
- [11] Weng, X. (2007). Identification and Distribution of *Mycobacterium leprae* Genotypes in a Region of High Leprosy Prevalence in China. *a 3-Year Molecular Epidemiological Study*. *J. Clin. Microbiol*, 45(6):1728-1734.
- [12] Depkes RI. (2005). *Buku Pedoman Pemberantasan Penyakit Kusta*. Jakarta: Depkes RI dan Ditjen PPM & PL.
- [13] Kemenkes RI. (2012). *Pedoman Nasional Pengendalian Penyakit Kusta*. Jakarta: Depkes.