

## MODELING THE NUMBER OF DIARRHEA INCIDENTS IN KENDARI USING GENERALIZED POISSON REGRESSION

Angga Saputra <sup>a)</sup>, Makkulau <sup>b)</sup>, Bahridin Abapihi <sup>c)</sup>, Irma Yahya <sup>d)</sup>, Ruslan <sup>e)</sup>,  
Lilis Laome <sup>f)</sup>

Department of Statistics, Halu Oleo University, Kendari, Indonesia  
e-mail: <sup>a)</sup>[anggasp0808@gmail.com](mailto:anggasp0808@gmail.com), <sup>b)</sup>[makkulau@gmail.com](mailto:makkulau@gmail.com), <sup>c)</sup>[bahridinabapihi@yahoo.com](mailto:bahridinabapihi@yahoo.com),  
<sup>d)</sup>[awwalud@gmail.com](mailto:awwalud@gmail.com), <sup>e)</sup>[ruslantibandung@gmail.com](mailto:ruslantibandung@gmail.com), <sup>f)</sup>[lilis.laome@uho.ac.id](mailto:lilis.laome@uho.ac.id)

### Abstract

The number of diarrhea incidents in Southeast Sulawesi Province, Indonesia, remains high, especially in the capital city, Kendari. As we can see in 2019 and 2020, there were 5.559 and 6.582 cases of diarrhea, respectively. For this reason, this article aims to address the problem by proposing the application of Generalized Poisson Regression (GPR) model. The adoption of GPR model is due to the existing of overdispersion on the dataset. The results of this research show that the population density variable is significant in the data analysis. Which means that the main factor results in high number of diarrhea incidents in Southeast Sulawesi Province is the population density.

**Key words:** Diarrhea, Poisson distribution, Generalized Poisson Regression, Overdispersion, Population density

### PENDAHULUAN

Diare merupakan penyakit yang tidak asing terdengar di kalangan masyarakat. Penyakit ini bisa terjadi kapan saja dan pada siapa saja, mulai dari kalangan anak-anak sampai kalangan dewasa. Diare dalam bahasa Indonesia biasa disebut dengan "buang air besar" atau *diarrhea* dalam bahasa Inggris. Diare merupakan sebuah gangguan pada saluran pencernaan dimana tinja atau *feses* berubah menjadi cair dan terjadi 3-7 kali dalam waktu 24 jam. Penyakit ini biasanya disertai dengan sakit perut, rasa mulas yang berkepanjangan, dehidrasi dan sering mual serta muntah.

Di Provinsi Sulawesi Tenggara, kabupaten/kota dengan jumlah kasus Diare terbanyak adalah Kota Kendari. Jumlah penyakit Diare di Kota Kendari mengalami kenaikan, dimana pada tahun 2019 jumlah kasus Diare sebanyak 5559 kasus dan pada tahun 2020 sebanyak 6582 kasus (Kota Kendari Dalam Angka, 2021).

Distribusi Poisson merupakan suatu distribusi untuk peristiwa yang probabilitas kejadiannya kecil di suatu daerah tertentu dengan hasil pengamatan berupa variabel diskrit dan antar variabel prediktor saling bebas. Ciri-ciri percobaan yang mengikuti sebaran distribusi Poisson yaitu (Cameron & Trivedi, 1998) kejadian dengan probabilitas kecil yang terjadi pada populasi dengan jumlah anggota yang besar, bergantung pada interval daerah tertentu, kejadian termasuk dalam proses stokastik (*counting process*).

Salah satu asumsi yang harus dipenuhi dalam regresi Poisson adalah variabel dependen (Y) diskrit dan asumsi equidispersi. Equidispersi yaitu nilai rata-rata sama dengan nilai varian.

Namun, terkadang pada analisis dengan model regresi Poisson terjadi pelanggaran asumsi equidispersi tersebut.

Pelanggaran asumsi terjadi apabila nilai varian lebih besar dari nilai rata-rata yang disebut overdispersi dan underdispersi apabila nilai varian lebih kecil dari nilai rata-rata. Pelanggaran asumsi equidispersi dapat diatasi dengan menggunakan salah satu metode statistika yaitu *Generalized Poisson Regression* (GPR). Model dari *Generalized Poisson Regression* adalah salah satu model regresi yang merupakan perluasan dari regresi Poisson yang dikenal dengan equidispersi. Model *Generalized Poisson Regression* merupakan sebuah model yang lebih baik dari regresi Poisson untuk mengatasi apabila terjadi masalah overdispersi/underdispersi (Ismail dan Jemain, 2005).

Penelitian ini menggunakan jumlah penderita Diare di Kota Kendari sebagai variabel respon yang merupakan data cacahan yang mengikuti distribusi Poisson dan berpotensi mengalami overdispersi. Berdasarkan latar belakang tersebut, peneliti menggunakan analisis *Generalized Poisson Regression* untuk mengetahui faktor-faktor yang mempengaruhi jumlah penderita Diare di Kota Kendari.

## METODE PENELITIAN

Penelitian ini dilakukan sejak bulan Juni 2022 sampai dengan Oktober 2022 bertempat di Laboratorium Statistika, Lantai 3 Gedung A Fakultas Matematika dan Ilmu Pengetahuan Alam (MIPA) Universitas Halu Oleo, Kota Kendari Provinsi Sulawesi Tenggara. Data yang digunakan dalam penelitian ini merupakan data sekunder dari Jumlah penyakit Diare di Kota Kendari pada tahun 2020 yang diperoleh dari Dinas Kesehatan Kota Kendari.

Pengolahan data pada penelitian ini yaitu dengan menggunakan *Software R*, *Software Minitab*, dan *Software Microsoft Excel* untuk pengolahan data sehingga menghasilkan running program yang diinginkan.

Langkah analisis yang dilakukan dalam penelitian ini yang didasarkan pada tujuan penelitian adalah sebagai berikut.

1. Mendeskripsikan jumlah kasus Diare dan faktor-faktor yang diduga mempengaruhinya di Kota Kendari pada tahun 2020 menggunakan analisis statistika deskriptif.
2. Menguji data jumlah kasus Diare apakah berdistribusi poisson atau tidak dengan uji Kolomogrov Smirnov.
3. Pengujian kasus multikolinieritas berdasarkan kriteria korelasi dan VIF.
4. Mendeteksi adanya overdispersion pada regresi Poisson.
5. Menganalisis model *Generalized Poisson Regression* (GPR) dengan langkah sebagai berikut:
  - a. Penaksiran parameter model GPR dengan metode MLE.
  - b. Menguji signifikansi parameter model GPR secara serentak dan parsial.
  - c. Menghitung nilai AIC model GPR.
  - d. Penaksiran parameter model GPR pada variabel signifikan
  - e. Menguji signifikansi parameter model GPR secara serentak dan parsial
  - f. Pemilihan model terbaik berdasarkan nilai AIC terkecil.
  - g. Menarik kesimpulan.

## HASIL DAN PEMBAHASAN

### Deskripsi Penelitian

Hal pertama yang dilakukan dalam penelitian yaitu melakukan analisis deskriptif dari variabel yang digunakan. Manfaat dari analisis deskriptif sendiri adalah untuk memberikan keterangan-keterangan mengenai suatu data. Peneliti menggunakan jumlah kasus Diare di Kota Kendari sebagai variabel respon ( $Y$ ), peneliti menggunakan empat variabel prediktor yang meliputi: Kepadatan Penduduk ( $X_1$ ), Rasio Fasilitas Kesehatan ( $X_2$ ), Rasio Petugas Kesehatan ( $X_3$ ), dan Persentasi Rumah Tangga Berperilaku Hidup Bersih dan Sehat ( $X_4$ ).

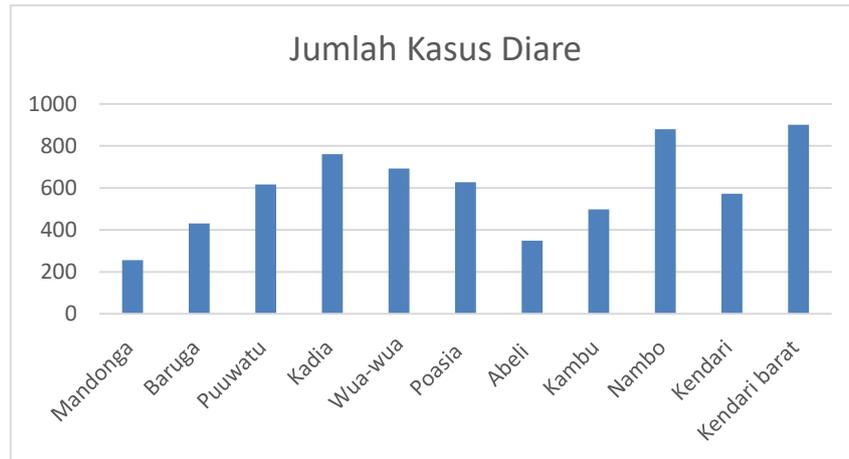
Kota Kendari memiliki 11 Kecamatan yang terdiri atas Kecamatan Mandonga, Baruga, Puuwatu, Kadia, Wua-wua, Poasia, Abeli, Kambu, Nambo, Kendari, dan Kendari Barat. Karakteristik dari kasus Diare di Kota Kendari dan faktor-faktor yang diduga mempengaruhinya dapat dilihat pada Tabel 1.

Tabel 1. Deskripsi jumlah penderita Diare dan variabel yang diduga berpengaruh

Variabel	Minimum	Maksimum	Rata-rata	Variansi	Standar Deviasi
$Y$	256	901	598,4	42615,4	206,4
$X_1$	441	5650	1797	2199048	1483
$X_2$	14,95	44,76	26,76	75,42	8,68
$X_3$	32,24	96,13	65,41	393,04	19,83
$X_4$	73	87	81,32	25,31	5,03

Dari Tabel 1 di atas ditemukan bahwa jumlah kasus Diare terendah berada di kecamatan Mandonga dengan jumlah kasus sebanyak 256, jumlah kasus Diare terbesar di kecamatan Kendari Barat sebanyak 901 jiwa dengan rata-rata kasus Diare yaitu 598 jiwa. Kemudian untuk jumlah Kepadatan Penduduk terkecil berada di Kecamatan Nambo sebanyak 441 jiwa/km<sup>2</sup> dan terbesar di Kecamatan Kadia sebanyak 5650 jiwa/km<sup>2</sup>. Untuk Rasio Fasilitas Kesehatan terkecil 14,95 persen di Kecamatan wua-wua dan terbesar di Kecamatan Nambo sebesar 44,76 persen. Untuk Rasio Tenaga Kesehatan terkecil di Kecamatan Mandonga 32,24 persen dan terbesar di Kecamatan Baruga 96,13 persen untuk Rumah Tangga yang Berperilaku Hidup Bersih dan Sehat terkecil di Kecamatan Kambu dengan 73 persen dan terbesar di Kecamatan Puuwatu dengan 87 persen dengan rata-rata 81,32 persen dan varian 25,31.

Pola sebaran data  $Y$  dapat dilihat pada Gambar 1. Dari Diagram batang Gambar 3.1 terlihat bahwa secara visual data jumlah kasus Diare di Kota Kendari berdistribusi poisson, selanjutnya akan dibuktikan dengan uji Kolomogrov-Smirnov.



Gambar1 Diagram batang dari jumlah penderita Diare(Y)

### Uji Distribusi Poisson

Untuk mengetahui apakah data yang diamati berdistribusi Poisson atau tidak, dilakukan dengan uji Kolmogorov-smirnov, di mana hipotesis pengujiannya sebagai berikut:

$H_0$  : Data Y berdistribusi Poisson

$H_1$  : Data Y tidak berdistribusi Poisson.

Kriteria pengujian dalam uji Kolmogorov-Smirnov adalah tolak  $H_0$  jika nilai signifikan  $< \alpha$ . Metode Kolmogorov-Smirnov juga dapat dihitung dengan nilai D yang didapat dihitung sebagai berikut:

$$D = \max|F_0(y) - S_n(y)|$$

$$D = 0,3636$$

dengan:

D adalah nilai standar deviasi absolut maksimum antara  $F_0(y)$  dan  $S_n(y)$ ,  $F_0(y)$  adalah fungsi distribusi frekuensi kumulatif relatif dari distribusi teoritis dibawah  $H_0$ , dan  $S_n(y)$  adalah distribusi frekuensi kumulatif pengamatan sebanyak sampel.

Dengan nilai D sebesar 0,3636 dan nilai  $D_{tabel}$  sebesar 0,3912 yang berarti bahwa nilai  $D < D_{(11;0,05)}$  maka dapat disimpulkan bahwa terima  $H_0$  yang berarti bahwa data Y berdistribusi Poisson. Diperoleh pula nilai  $p$  pada hasil output di R yaitu  $0,0825 > \alpha$ . Maka dapat disimpulkan bahwa variabel respon (Y) berdistribusi Poisson.

### Pemeriksaan Multikolinieritas

Salah satu syarat yang harus dipenuhi dalam pembentukan model regresi dengan beberapa variabel prediktor adalah tidak ada kasus multikolinieritas. Pendeteksian kasus multikolinieritas yaitu kriteria nilai VIF (*Variance Inflation Factor*). Terdapat kasus multikolinieritas jika nilai VIF lebih besar dari 10. Berikut adalah nilai VIF untuk 4 variabel yang digunakan dalam analisis.

Tabel 2. Nilai VIF

VARIABEL	VIF
$X_1$	1,346
$X_2$	1,661
$X_3$	1,471
$X_4$	1,637

Berdasarkan Tabel 2 menunjukkan bahwa nilai VIF dari seluruh variabel yang digunakan memiliki nilai kurang dari 10, sehingga dapat disimpulkan bahwa tidak terjadi kasus multikolinieritas pada variabel yang akan digunakan dalam analisis dan asumsi multikolinieritas telah terpenuhi

### Pengujian Overdispersi

Overdispersi merupakan kondisi dimana model regresi Poisson memiliki nilai varians lebih besar dari nilai rata-ratanya dan underdispersi yaitu kondisi dimana data variabel respon menunjukkan nilai varians lebih kecil dari nilai rata-ratanya.

Uji statistik yang dapat digunakan untuk mendeteksi overdispersi pada suatu data yaitu dengan menggunakan *package* AER dari *software* R, dengan hipotesis pengujiannya adalah sebagai berikut:

$H_0$  = Tidak mengalami overdispersi

$H_1$  = Mengalami overdispersi

Berdasarkan hasil output R, dapat dilihat nilai dispersi sebesar 2,9599 dan p-value = 0,0015 dimana nilai  $\alpha = 0,05$ . Oleh karena nilai p-value  $< \alpha$  atau nilai dispersi  $> 1$  maka tolak  $H_0$  sehingga dapat disimpulkan bahwa terjadi overdispersi.

### Pemodelan *Generalized Poisson Regression* (GPR)

Dalam penanganan kasus overdispersi pada penelitian ini akan menggunakan *Generalized Poisson Regression* (GPR). Model GPR merupakan suatu model yang digunakan untuk data diskrit yang dimana data tersebut terjadi overdispersi. Model GPR mempunyai bentuk yang sama dengan regresi Poisson. Pada model GPR dilakukan penaksiran parameter yang tidak diketahui dengan metode *Maximum Likelihood Estimation* (MLE). Selanjutnya diperoleh taksiran parameternya yaitu sebagai berikut:

$$\mu_i = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4)$$

Hasil estimasi parameter jumlah penderita Diare di Kota Kendari dapat dilihat dalam Tabel 3.3

Tabel 3. Estimasi parameter model *Generalized Poisson Regression*

Parameter	Estimasi	Standard Error
$\beta_0$	3,639	3,020
$\beta_1$	$7,148 \times 10^{-5}$	$9,407 \times 10^{-6}$
$\beta_2$	$1,980 \times 10^{-2}$	$1,958 \times 10^{-3}$
$\beta_3$	$-1,343 \times 10^{-3}$	$7,021 \times 10^{-4}$
$\beta_4$	$2,671 \times 10^{-2}$	$3,455 \times 10^{-3}$

Dalam pengujian parameter terdapat 2 (dua) cara yaitu uji serentak parameter (simultan) dan uji parsial parameter untuk mengetahui signifikan dari masing-masing parameter yang diperoleh:

Pengujian serentak parameter model GPR dilakukan dengan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) seperti pada pengujian parameter model Regresi Poisson, dengan hipotesis sebagai berikut.

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  (tidak ada parameter yang signifikan dalam model)

$H_1: \beta_j \neq 0 ; j = 1, 2, 3, 4$  (paling sedikit ada satu parameter yang signifikan dalam model)

Statistika Uji:

$$D(\widehat{\beta}_j) = -2 \ln \left( \frac{L(\widehat{\omega})}{L(\widehat{\Omega})} \right)$$

Dengan  $L(\widehat{\Omega})$  yaitu nilai Likelihood untuk model lengkap dengan melibatkan variabel prediktor dan  $L(\widehat{\omega})$ , yaitu nilai Likelihood untuk model sederhana tanpa melibatkan variabel prediktor.

Daerah penolakan adalah tolak  $H_0$  jika statistik nilai  $D(\widehat{\beta}_j) > X^2_{(k,\alpha)}$ . Artinya paling sedikit ada satu variabel yang memberikan pengaruh pada model.

Tabel 4 Pengujian parameter secara serentak (simultan)

Model	Nilai $D(\widehat{\beta}_j)$	Chi- Kuadrat
GPR	576,331	9,487

Berdasarkan hasil output R Nilai  $D(\widehat{\beta}_j)$  adalah 576,331 dan nilai tabel Chi-kuadrat  $X^2_{(0,05;4)}$  adalah 9,487. Oleh karena  $D(\widehat{\beta}_j) > X^2_{(0,05;4)}$  maka tolak  $H_0$ . Sehingga dapat disimpulkan variabel Kepadatan Penduduk, Rasio Tenaga Kesehatan, Rasio Fasilitas Kesehatan, Rumah Tangga Berperilaku Hidup Bersih dan Sehat secara simultan mempengaruhi penyakit Diare di Kota Kendari.

Setelah dilakukan uji serentak, dilanjutkan dengan pengujian parsial setiap variabel prediktornya dengan hipotesis sebagai berikut:

$H_0: \beta_j = 0$

$H_1: \beta_j \neq 0 ; j = 1, 2, 3, 4$

Statistika Uji:

$$Z = \frac{\widehat{\beta}_j}{SE(\widehat{\beta}_j)}$$

Daerah penolakan adalah tolak  $H_0$  jika nilai  $|Z_{hit}| > Z_{\alpha/2}$  dimana  $\alpha$  adalah tingkat signifikansi yang digunakan..

1. Hipotesis

$H_0: \beta_1 = 0$  (pengaruh variabel  $X_1$  tidak signifikan terhadap jumlah penderita)

$H_1: \beta_1 \neq 0$  (pengaruh variabel  $X_1$  signifikan terhadap jumlah penderita)

Statistika Uji:

$$Z_1 = \frac{7,148 \times 10^{-5}}{9,407 \times 10^{-6}}$$

$$Z_1 = 7,598$$

Hasil pengujian seluruh variabel secara parsial dapat dilihat pada tabel 5. Berdasarkan Tabel 5 dapat disimpulkan bahwa variabel yang berpengaruh secara parsial adalah Kepadatan Penduduk ( $X_1$ ), Rasio Fasilitas Kesehatan ( $X_2$ ), dan Persentasi Rumah Tangga Berperilaku Hidup Bersih dan Sehat ( $X_4$ ).

Tabel 5. Pengujian parameter secara parsial

Variabel	Estimate	Standard Error	Z <sub>hit</sub>	Z <sub>a/2</sub>
X <sub>1</sub>	7,148 x 10 <sup>-5</sup>	9,407 x 10 <sup>-6</sup>	7,598	1,96
X <sub>2</sub>	1,980 x 10 <sup>-2</sup>	1,958 x 10 <sup>-3</sup>	10,113	1,96
X <sub>3</sub>	-1,343 x 10 <sup>-3</sup>	7,021 x 10 <sup>-4</sup>	-1,913	1,96
X <sub>4</sub>	2,671 x 10 <sup>-2</sup>	3,455 x 10 <sup>-3</sup>	7,730	1,96

### Pemodelan Lanjutan terhadap Variabel yang Signifikan

Berdasarkan hasil perhitungan di *software* R pada pemodelan dengan menggunakan tiga variabel signifikan, di dapatkan variabel X<sub>2</sub> dan X<sub>4</sub>, memiliki arah yang tidak logis. Kemudian dilakukan pemodelan dengan menggunakan variabel X<sub>1</sub>.

$$\mu_i = \exp(\beta_0 + \beta_1 X_1)$$

Hasil estimasi parameter variabel yang berpengaruh terhadap jumlah penderita Diare di Kota Kendari dapat dilihat pada Tabel 6.

Tabel 6. Estimasi parameter model *Generalized Poisson Regression*

Parameter	Estimasi	Standard error
$\beta_0$	6,287	1,991 x 10 <sup>-1</sup>
$\beta_1$	5,752 x 10 <sup>-5</sup>	8,138 x 10 <sup>-6</sup>

### Pemilihan Model Terbaik

Dalam pemilihan model yang baik dalam penelitian ini yaitu dengan menggunakan nilai AIC yaitu seperti yang ditunjukkan pada Tabel 3.7.

Tabel 7. Nilai AIC

Model	AIC
X <sub>1</sub> , X <sub>2</sub> , X <sub>4</sub>	677,846
X <sub>1</sub> , X <sub>4</sub>	788,556
X <sub>1</sub>	795,194

Hasil pada Tabel 4.7 dapat diketahui bahwa nilai AIC terkecil pada model GPR yaitu pada model pertama namun pada model pertama terdapat variabel yang tidak signifikan dan model kedua dan ketiga memiliki arah yang tidak logis. Sehingga, model terbaik yaitu model keempat X<sub>1</sub>.

Untuk hasil estimasi parameter *Generalized Poisson Regression* untuk jumlah penderita Diare di Kota Kendari dengan model yang terbentuk yaitu :

$$\hat{\mu} = \exp(\hat{\beta}_0 + \hat{\beta}_1 X_1)$$

$$\hat{\mu} = \exp(6,287 + 0,00005752X_1)$$

Variabel yang berpengaruh signifikan dalam model adalah Kepadatan Penduduk (X<sub>1</sub>). Dari model yang diperoleh menggambarkan bahwa untuk setiap kenaikan 1% Kepadatan Penduduk dengan asumsi bahwa nilai variabel lainnya tetap, maka jumlah penderita Diare akan meningkat sebesar  $\exp(0,00005752) = 1$  di Kota Kendari. Berdasarkan model yang diperoleh dapat diketahui bahwa dari 4 koefisien variabel prediktor ada 1 yang memiliki arah hubungan yang logis terhadap banyaknya kasus Diare di Kota Kendari yaitu Kepadatan Penduduk (X<sub>1</sub>).

Hasil penelitian ini searah dengan penelitian yang telah dilakukan oleh Fitri ernawati dalam tugas akhir/thesis yang berjudul “Pemodelan Dan Pemetaan Kejadian Diare di Kota Surabaya Menggunakan *Geographically Weighted Negative Binomial Regression*” dalam penelitian tersebut variabel Kepadatan Penduduk ( $X_1$ ).

## KESIMPULAN

Berdasarkan hasil dan pembahasan di atas, dapat di Pemodelan *Generalized Poisson Regresion* (GPR) pada kasus jumlah penderita Diare di Kota Kendari pada tahun 2020 diperoleh model yang signifikan yaitu  $\hat{\mu} = \exp(6,287 + 0,00005752X_1)$  yang menunjukkan bahwa variabel prediktor yang diduga mempengaruhi variabel respon yang terbukti berpengaruh signifikan dalam model yaitu variabel Kepadatan Penduduk ( $X_1$ ). Diperoleh pula model yang terbaik dengan nilai AIC yaitu sebesar 795,194

## BIBLIOGRAFI

- Cahyandari, R. (2014). Pengujian Overdispersi pada Model Regresi Poisson (Studi Kasus: Laka Lantas Mobil Penumpang di Provinsi Jawa Barat). *Jurnal Statistika*. **14(2)**: 69-76.
- Desiana, C.M. & Melaniani, S. (2022). Pemodelan *Generalized Poisson Regression* pada Jumlah Kasus Penyakit Difteri di Kota Surabaya Tahun 2017, *Jurnal Ilmiah Permas: Jurnal Ilmiah STIKES Kendal*. **15(3)**: 547-558.
- Dinas Kesehatan Kota Kendari. (2021). Profil Kesehatan Kota Kendari 2020. Kendari: Dinas Kesehatan Kota Kendari.
- Fauziyyah, A., Anna C., Gede N.M.J, (2022) Pemodelan dan Pemetaan Penyakit Diare di Kabupaten Sumedang Tahun 2022, *Jurnal Statistika*. **16(2)**:51-62.
- Hardiyanti, F., Tambunan, H.S. & Saragih, I.S. (2019). Penerapan Metode K-Medoids Clustering Pada Penanganan Kasus Diare di Indonesia. *KOMIK*. **3(1)**.
- Margarethy, I., Suryaningtyas, N.H. & Yahya, Y. (2020). Kejadian Diare Ditinjau Dari Aspek Jumlah Penduduk dan Sanitasi Lingkungan.
- Marita, V.E. (2017). Pemodelan Terhadap Faktor-Faktor yang Mempengaruhi Jumlah Penduduk Miskin di Jawa Timur Menggunakan *Generalized Poisson Regression*. [Skripsi]. FMIPA:Institut Teknologi Sepuluh Nopember. Surabaya.
- Meilania, I. (2016). Faktor-Faktor yang Mempengaruhi Jumlah Penderita Tuberkulosis di Kabupaten Probolinggo Tahun 2014 Menggunakan Pendekatan *Generalized Poisson Regression* (GPR). [Skripsi]. FMIPA;Institut Teknologi Sepuluh Nopember. Surabaya.
- Nasution, L.M. (2017). Statistik Deskriptif. *Jurnal Hikmah*. **14(1)**: 49-55.
- Prahutama, A. (2017). Analisis Faktor-Faktor yang Mempengaruhi Angka Kematian Bayi di Jawa Tengah Menggunakan Regresi *Generalized Poisson* dan *Binomial Negatif*. *Jurnal Statistika*. **5(2)**.
- Priambodo, B.W.Y. & Irhamah. (2019). Pemetaan Jumlah *Property Crime* di Provinsi Jawa Timur Menggunakan Metode *Geographically Weighted Negative Binomial Regression* (GWNBR) dan *Geographically Weighted Poisson Regression* (GWPR). *Jurnal INFERENSI*, **2(2)**: 53-62.
- Rahmadeni & Jannah, F.K. (2019). Pemodelan Kasus Kematian Neonatal dengan Menggunakan *Generalized Poisson Regression* (GPR). *Jurnal Sains Matematika dan Statistika*. **5(2)**.

- Ratri, N.H & Purhadi. (2015). Analisis Faktor-Faktor yang Mempengaruhi Jumlah Kasus Malaria di Jawa Timur Tahun 2014 dengan *Geographically Weighted Negative Binomial Regression* (GWNBR). *Jurnal Sains dan Seni*. ITS. **4(2)**.
- Ruliana, Hendikawati, P. dan Agoestanto, A. (2016). Pemodelan *Generalized Poisson Regression* Suryana, N., Pratiwi dan Prasetyo, R.T. 2021. Penanganan Ketidakseimbangan Data Pada Prediksi Customer Churn Menggunakan Kombinasi SMOTE Dan Boosting. *Indonesian Journal On Computer And Information Technology*.
- Setiyono, A. (2019). Faktor Risiko Kejadian Diare pada Masyarakat Kota Tasikmalaya, *Jurnal Kesehatan Komunitas Indonesia*. **15(2)**: 49-49.
- Syadifa, M.Z. (2016). Pemodelan Jumlah Penderita Kusta di Kabupaten Probolinggo Tahun 2014 Menggunakan *Generalized Poisson Regression* (GPR). [Skripsi]. FMIPA: Institut Teknologi Sepuluh Nopember. Surabaya.