

**PEMODELAN MORTALITAS PENDERITA DIFTERI
DI INDONESIA MENGGUNAKAN REGRESI *ZERO-
INFLATED NEGATIVE BINOMIAL***

Skripsi

**Disusun untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Statistika**



**Syifa Fadia Salsabila
1314618012**


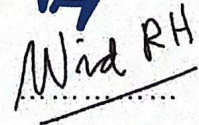
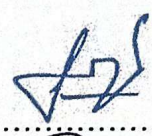
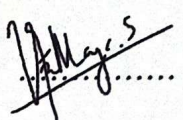
**PROGRAM STUDI STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA
2022**

LEMBAR PENGESAHAN UJIAN SKRIPSI

PEMODELAN MORTALITAS PENDERITA DIFTERI DI INDONESIA
MENGUNAKAN REGRESI *ZERO-INFLATED NEGATIVE BINOMIAL*

Nama : Syifa Fadia Salsabila

No. Registrasi : 1314618012

	Nama	Tanggal
Penanggung Jawab		
Dekan	: <u>Prof. Dr. Muktiningsih N, M.Si.</u> NIP. 196405111989032001	 29-08-2022
Wakil Penanggung Jawab		
Wakil Dekan I	: <u>Dr. Esmar Budi, S.Si., MT.</u> NIP. 197207281999031002	 29-08-2022
Ketua Penguji	: <u>Dian Handayani, M.Si.</u> NIP. 197404151998032001	 21-08-2022
Sekretaris	: <u>Dra. Widyanti Rahayu, M.Si.</u> NIP. 196611032001122001	 19-08-2022
Anggota		
Pembimbing I	: <u>Dr. Ir. Bagus Sumargo, M.Si.</u> NIP. 196309221986011001	 19-08-2022
Pembimbing II	: <u>Faroh Ladayya, M.Si.</u> NIP. 199401282020122018	 23-08-2022
Penguji Ahli	: <u>Vera Maya Santi, M.Si.</u> NIP. 197905312005012006	 22-08-2022

Dinyatakan lulus ujian skripsi tanggal 16 Agustus 2022

LEMBAR PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul **“Pemodelan Mortalitas Penderita Difteri di Indonesia Menggunakan Regresi *Zero-Inflated Negative Binomial*”** yang disusun sebagai syarat untuk memperoleh gelar Sarjana Statistika dari Program Studi Statistika Universitas Negeri Jakarta adalah karya ilmiah saya dengan arahan dari dosen pembimbing,

Sumber informasi yang diperoleh dari penulis lain yang telah dipublikasikan yang disebutkan dalam teks skripsi ini telah dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini sesuai dengan norma, kaidah, dan etika penulisan ilmiah pada umumnya dan ketentuan yang berlaku di Universitas Negeri Jakarta.

Jika dikemudian hari ditemukan sebagian besar skripsi ini bukan hasil karya saya sendiri dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sanding dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku.

Jakarta, 23 Agustus 2022



Syifa Fadia Salsabila

ABSTRAK

SYIFA FADIA SALSABILA. Pemodelan Mortalitas Penderita Difteri di Indonesia Menggunakan Regresi *Zero-Inflated Negative Binomial*. Skripsi, Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Agustus 2022.

Indonesia merupakan salah satu negara penyumbang kasus kejadian difteri tertinggi di dunia yang masih disertai dengan kematian. Penelitian ini bertujuan untuk memodelkan jumlah mortalitas difteri di Indonesia dan mengetahui faktor-faktor yang signifikan mempengaruhinya. Data yang dianalisis dalam penelitian ini mengindikasikan terjadinya overdispersi yang disebabkan oleh *excess zeros* karena pada variabel respon, yaitu data jumlah mortalitas difteri, terdapat 73% pengamatan bernilai nol dan persentase nilai nol pada data lebih besar dari persentase nilai nol berdasarkan perhitungan distribusi Poisson sehingga regresi Poisson menjadi kurang sesuai dalam memodelkan data tersebut. Model alternatif yang mampu mengatasi permasalahan overdispersi yang disebabkan oleh *excess zeros* adalah regresi *Zero-Inflated Negative Binomial* (ZINB). Berdasarkan kriteria uji Vuong dan nilai AIC, model regresi ZINB memiliki performa yang relatif lebih cocok digunakan daripada regresi Poisson. Berdasarkan pengujian hipotesis parameter model regresi ZINB, pada model *count* diperoleh kelima variabel penjelas yang digunakan dalam penelitian ini berpengaruh signifikan terhadap jumlah mortalitas difteri di Indonesia, sementara pada model *logit* tidak ada variabel penjelas yang signifikan.

Kata Kunci: regresi Poisson, overdispersi, *excess zeros*, *Zero-Inflated Negative Binomial* (ZINB)

ABSTRACT

SYIFA FADIA SALSABILA. Modeling of Diphtheria Patients Mortality in Indonesia Using Zero-Inflated Negative Binomial Regression. Thesis, Department of Statistics, Faculty of Mathematics and Natural Science, State University of Jakarta. August 2022.

Indonesia is one of the countries that contribute to the highest incidence of diphtheria cases in the world which is still accompanied by death. This study aims to model the number of diphtheria mortality in Indonesia and to determine the factors that significantly influence it. The data analyzed in this study indicates the occurrence of overdispersion caused by excess zeros because the response variable, namely the data on the number of diphtheria mortality, there are 73% of observations with zero values and the percentage of zero values in the data is greater than the percentage of zero values based on the calculation of the Poisson distribution so that the Poisson regression is not suitable for modeling the data. An alternative model that is able to overcome the overdispersion caused by excess zero is the Zero-Inflated Negative Binomial (ZINB) regression. Based on the Vuong test and the AIC value criteria, the ZINB regression model has a relatively more suitable performance than the Poisson regression. Based on the hypothesis testing of the ZINB regression model parameters, the count model obtained five explanatory variables used in this study that had a significant effect on the number of diphtheria mortality in Indonesia, while in the logit model there is no significant explanatory variable.

Keywords: Poisson regression, overdispersion, excess zeros, Zero-Inflated Negative Binomial (ZINB)

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Allah Subhanahu Wa Ta'ala atas berkat rahmat dan hidayah-Nya, penulis dapat menyelesaikan skripsi dengan baik. Salam dan shalawat semoga selalu tercurah pada Rasulullah Shallallahu 'Alaihi Wasallam.

Skripsi berjudul “Pemodelan Mortalitas Penderita Difteri di Indonesia Menggunakan Regresi *Zero-Inflated Negative Binomial*” ini merupakan syarat untuk menyelesaikan akhir program strata 1 yang telah ditetapkan di Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.

Selama tahap penyelesaian skripsi, penulis banyak menerima bimbingan, bantuan, dukungan, serta doa dari berbagai pihak. Oleh karena itu, melalui kesempatan ini penulis ingin menyampaikan ucapan rasa terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Ir. Bagus Sumargo, M.Si. selaku Dosen Pembimbing I sekaligus Koordinator Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Jakarta yang telah meluangkan waktunya untuk memberikan banyak ilmu, bimbingan, serta arahan sehingga skripsi ini dapat menjadi lebih baik dan terarah.
2. Ibu Faroh Ladayya, M.Si. selaku Dosen Pembimbing II yang telah meluangkan waktunya untuk memberikan banyak ilmu, bimbingan, serta arahan sehingga skripsi ini dapat menjadi lebih baik dan terarah.
3. Ibu Vera Maya Santi, M.Si. selaku Pembimbing Akademik atas segala bimbingan dan masukan yang telah diberikan selama perkuliahan.
4. Seluruh Bapak/Ibu dosen pengajar dan staf administrasi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Jakarta atas segala dukungan, bimbingan, dan petunjuk selama pelaksanaan pendidikan dan penyusunan skripsi.
5. Kedua orang tua tercinta, Bapak Noviansyah dan Ibu Rokayah serta Zahra Salsabila dan Muhammad Fahri selaku adik-adik terkasih yang selalu memberikan dukungan kepada penulis secara moril maupun materil, doa, kasih sayang, dan motivasi.

6. Nia Rahayu Ningsih, Afifah Nur Mutia, Rita Septia NR, Defina Ambarwati, Rifa Kamilia, Anjuita, Indah Lestari, Yayang Nurul Aulia Azizah, dan Muhamad Rifan yang telah menjadi tempat bertukar pikiran, pengetahuan, serta memberikan masukan dan semangat sampai dengan selesainya skripsi ini.
7. Abdul Jalil dan Firya Amanda yang telah menjadi teman cerita dan keluh kesah serta memberikan semangat dalam menyusun skripsi ini.
8. Seluruh teman-teman Prodi Statistika Angkatan 2018 yang telah memberikan bantuan dan dukungan serta berjuang bersama-sama selama masa perkuliahan.
9. Seluruh pihak yang tidak dapat penulis sebutkan satu persatu, yang telah memberikan dukungan, motivasi, dan doa selama proses penyelesaian proposal skripsi ini.

Penulis menyadari bahwa penyusunan skripsi ini masih jauh dari sempurna karena keterbatasan pengetahuan dan kemampuan yang penulis miliki. Oleh karena itu, segala kritik dan saran yang membangun akan penulis terima dengan senang hati. Semoga skripsi ini dapat bermanfaat bagi semua yang membaca.

Jakarta, 19 Agustus 2022



Syifa Fadia Salsabila

DAFTAR ISI

	Halaman
ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	5
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
BAB II KAJIAN PUSTAKA	7
2.1 Penyakit Difteri	7
2.2 Multikolinearitas	9
2.3 <i>Generalized Linear Model</i>	10
2.4 Regresi Poisson	11
2.4.1 Pendugaan Parameter Regresi Poisson	12
2.5 Uji Kecocokan (<i>Goodness of Fit</i>) pada Regresi Poisson	15
2.6 Permasalahan pada Regresi Poisson	15
2.6.1 Overdispersi	16
2.6.2 <i>Excess Zeros</i>	16
2.7 Regresi <i>Zero-Inflated Negative Binomial</i>	17
2.7.1 Pendugaan Parameter Regresi <i>Zero-Inflated Negative Binomial</i>	21
2.8 Pengujian Kelayakan Model	24
2.9 Pengujian Signifikansi Parameter Model	24
2.9.1 Pengujian Signifikansi Parameter Secara Simultan	25
2.9.2 Pengujian Signifikansi Parameter Secara Parsial	25
2.10 Pemilihan Model Terbaik	26
BAB III METODOLOGI PENELITIAN	27

3.1 Sumber Data	27
3.2 Variabel Penelitian	27
3.3 Definisi Operasional Variabel	28
3.3.1 Variabel Respon.....	28
3.3.2 Variabel Penjelas	28
3.3.3 Variabel <i>Exposure</i>	29
3.4 Tahapan Analisis Data.....	29
BAB IV HASIL DAN PEMBAHASAN	32
4.1 Eksplorasi Data.....	32
4.2 Pemeriksaan Multikolinearitas	35
4.3 Pemodelan Regresi Poisson.....	35
4.4 Pemeriksaan <i>Excess Zeros</i> pada Variabel Respon	37
4.5 Pemodelan Regresi <i>Zero-Inflated Negative Binomial</i>	38
4.6 Uji Kelayakan Model	39
4.7 Uji Signifikansi Parameter Model ZINB.....	40
4.7.1 Uji Signifikansi Parameter Secara Simultan.....	40
4.7.2 Uji Signifikansi Parameter Secara Parsial	40
BAB V KESIMPULAN DAN SARAN.....	44
5.1 Kesimpulan.....	44
5.2 Saran	44
DAFTAR PUSTAKA	45
LAMPIRAN.....	49

DAFTAR TABEL

	Halaman
Tabel 3.1 Variabel Penelitian.....	27
Tabel 4.1 Statistik Deskriptif Mortalitas Difteri dan Variabel Penjelasnya.....	33
Tabel 4.2 Nilai VIF Variabel Penjelasa	35
Tabel 4.3 Hasil Pendugaan Parameter Model Regresi Poisson	36
Tabel 4.4 Nilai Parameter Dispersi	36
Tabel 4.5 Pemeriksaan <i>Excess Zeros</i>	37
Tabel 4.6 Hasil Pendugaan Parameter Model Regresi ZINB	38
Tabel 4.7 Nilai AIC Model Regresi Poisson dan ZINB.....	39
Tabel 4.8 Hasil Pengujian Parameter Model Regresi ZINB Secara Parsial	41

DAFTAR GAMBAR

	Halaman
Gambar 1.1 Grafik Jumlah Kasus Difteri di Indonesia Tahun 1985-2020	2
Gambar 3.1 Diagram Alir	31
Gambar 4.1 Peta Sebaran Jumlah Mortalitas Difteri di Indonesia.....	32
Gambar 4.2 <i>Bar Chart</i> Jumlah Mortalitas Difteri di Indonesia	34
Gambar 4.3 <i>Pie Chart</i> Jumlah Mortalitas Difteri di Indonesia.....	37

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Jumlah Mortalitas Difteri dan Jumlah Penduduk.....	50
Lampiran 2. Data Variabel Penjelas.....	51
Lampiran 3. Pemeriksaan Multikolinearitas	52
Lampiran 4. Model Regresi Poisson	52
Lampiran 5. Model Regresi <i>Zero-Inflated Negative Binomial</i>	53
Lampiran 6. Nilai AIC Model Regresi ZINB	53
Lampiran 7. Pengujian Kelayakan Model.....	54
Lampiran 8. Pengujian Signifikansi Parameter Secara Simultan.....	54
Lampiran 9. Tabel <i>Chi-Square</i>	55