JURNAL FOURIER | April 2018, Vol. 7, No. 1, 35-43 ISSN 2252-763X
DOI: 10.14421/fourier.2018.71.35-43 E-ISSN 2541-5239

Zero Inflated Negative Binomial (ZINB) untuk Pemodelan Frekuensi Bepergian Penduduk Kabupaten Tapanuli Selatan Tahun 2016

Eko Yulian

Pusdiklat Badan Pusat Statistik, Jl. Jagakarsa no. 70, Jagakarsa, Jakarta Selatan, Indonesia

Korespondensi; Email: okeyulian@gmail.com

Abstrak

Sektor pariwisata memiliki peran penting dalam pembangunan ekonomi wilayah karena dianggap mampu meningkatkan penerimaan daerah, terciptanya lapangan kerja dan peluang usaha serta pembangunan infrastruktur. Sumatera Utara menjadi tujuan wisata utama di pulau Sumatera bila dilihat dari besarnya persentase penerimaan dari wisatawan domestik. Salah satu kabupaten yang memiliki potensi lokasi wisata di Sumatera Utara adalah Kabupaten Tapanuli Selatan. Pemodelan aktivitas perjalanan wisata penduduk Kabupaten Tapanuli Selatan menggunakan *Zero Inflated Binomial Negative* (ZINB) karena mayoritas penduduk banyak yang tidak melakukan perjalanan (*excess zero*) dan terjadi underdispersi jika dilakukan pemodelan dengan *Zero Inflated Poisson* (ZIP). Pada *ZINB* bagian regresi Binomial Negatif *(log-link)* variabel yang berpengaruh signifikan terhadap frekuensi bepergian adalah variabel usia dan jenis kelamin, sedangkan pada *zero inflation (logistic-link) part* variabel yang berpengaruh adalah variabel tingkat pendidikan.

Kata Kunci: Pariwisata; Underdispersi Poisson; ZIP; ZINB

Abstract

Tourism Sectors has an important role in regional economic formation. It's affected on increasing regional income, job creations and business opportunities, and infrastructure development. Sumatera Utara is a primary tourism destination in Sumatera Island. It can be seen in the income share of GDRP domestic tourism sector. One of region that has a lot of tourism sites is Tapanuli Selatan Regency. Tourism trip of Tapanuli Selatan people can be analyzed in Zero Inflated Binomial Negative (ZINB) because most of people didn't do any trip (excess zero) in a periode of time is really possible and this case may drive the data to under or overdispersion. ZINB outcomes is consists two parts, Negative Binomial part and zero inflation part. The significant variable in Negative Binomial part (log-link) are age and sex. In zero inflation (logistic-link) part, education is significant variable.

Keywords: Tourism; Poisson Underdispersion ZIP; ZINB

Pendahuluan

Aktivitas pariwisata memilik peran penting dalam pembangunan ekonomi suatu wilayah karena menurut Kemenpar (2014) industri pariwisata saat ini telah menjadi penggerak utama kemajuan sosial ekonomi suatu negara melalui penerimaan devisa, terciptanya lapangan pekerjaan dan peluang usaha, serta pembangunan infrastruktur. Pesatnya industri pariwisata saat ini tak bisa terlepas dari peran para pengunjung baik wisatawan domestik maupun wisatawan mancanegara di lokasi wisata yang melakun transaksi atas berbagai penawaran di lokasi wisata. Hal tersebut dikuatkan pula dengan data PDB menurut pengeluaran yang dirilis oleh BPS pada publikasi NESPARNAS bahwa konsumsi rumah tangga di sektor pariwisata sebesar 183,73 triliun rupiah atau sekitar 3,62 persen terhadap PDB Nasional pada tahun 2013.

© 2018 JURNAL FOURIER

Tabel 1. PDB Nasional Menurut Pengeluaran Tahun 2013 (Trilyun Rupiah).

Komponen	Konsumsi Rumah Tangga	Konsumsi Pemerintah	Investasi	Expor	Impor
Pariwisata	183,73	7,12	121,30	129,75	81,72
PDB Nasional	5.071,09	827,24	2.876,25	2.156,81	2.338,12
Share Pariwisata (%)	3,62	0,86	4,22	6,02	3,50

Berdasarkan hal tersebut diatas, potensi wisatawan lokal tentunya akan terus menjadi perhatian dalam memajukan perekonomian negara dan daerah pada khususnya. Memahami bagaimana karakteristik dan perilaku masyarakat dalam melakukan perjalanan wisata menjadi hal yang penting untuk mengembangkan sektor pariwisata di suatu daerah agar dapat dirumuskan berbagai kebijakan dalam membangun industri pariwisata di daerah yang sesuai dengan karakteristik dan kebutuhan masyarakat.

Sumatera Utara merupakan propinsi di pulau Sumatera yang menjadi daerah tujuan wisata utama bagi para wisatawan domestik bila dibandingkan dengan propinsi lainnya di pulau tersebut. BPS menyebutkan dalam publikasi NESPARNAS bahwa pada tahun 2014 Sumatera Utara menempati posisi pertama di pulau Sumatera sebagai propinsi dengan tingkat pendapatan tertinggi dari wisatawan domestik atau sekitar 3,71 persen berkontribusi terhadap pendapatan dari wisatawan domestik secara nasional. Salah satu daerah di Sumatera Utara yang memiliki potensi pariwisata adalah Kabupaten Tapanuli Selatan yang terletak di bagian selatan. Kabupaten ini awalnya merupakan kabupaten dengan wilayah yang sangat luas dan telah dimekarkan menjadi 5 kabupaten/ kota ketika otonomi daerah. Sebagai kabupaten induk yang posisinya berada ditengah-tengah daerah pemekarannya, menjadi menarik untuk menganalisis aktivitas perjalanan wisata penduduknya apalagi Tapanuli Selatan memiliki objek wisata alam yang paling banyak dibandingkan kabupaten/ kota di sekitarnya. Aktivitas perjalanan yang dilakukan oleh masyarakat Tapanuli Selatan dalam penelitian ini dibatasi pada aspek sosial demografi karena terbatasnya variabel pada data primer terkait aktivitas perjalanan wisata.

Regresi Poisson merupakan suatu *Generlize Linier Model* (GLM) yang mengasumsikan peluang kejadian pada variabel respon sangat kecil terjadi. Variabel respon dalam hal aktivitas perjalanan wisata yang dilakukan penduduk bisa saja banyak mengandung nilai nol (*excess zero*) atau sangat sedikit penduduk yang berpergian sehingga jika dianalisis dengan regresi Poisson maka estimasi yang diperoleh akan cenderung *over/ underestimate*. Untuk mengatasi hal tersebut maka digunakan regresi *Zero Inflated Poisson* yang akan memecah model menjadi 2 bagian yaitu model *poisson part* untuk mengakomodir data cacah bernilai tidak nol pada variabel respon dan model logit yang akan mengakomodir data cacah bernilai nol pada variabel respon. Model regersi *Zero Inflated Poisson* mensyaratkan asumsi equidispersi antara *means* dan varians dari variabel respon. Namun pada kenyataannya hal tersebut hampir tidak pernah terpenuhi karena yang sering terjadi adalah kasus overdispersi (nilai means lebih kecil daripada varians) maupun underdispersi. Dalam mengatasi masalah tersebut dapat diatasi dengan model regresi *Zero Inflated Negative Binomial* (ZINB). Berdasarkan paparan tersbut maka tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor yang berpengaruh terhadap frekuensi bepergian penduduk di Tapanuli selatan dengan menggunakan metode ZINB.

Landasan Teori

Survei Sosial Ekonomi Nasional

Data yang digunakan dalam penelitian ini adalah data sekunder, yang diperoleh dari data mentah Survei Sosial Ekonomi Nasional (SUSENAS) Tahun 2016. SUSENAS 2016 merupakan survei tingkat nasional yang dilaksanakan oleh Badan Pusat Statistik (BPS), mencakup 75.000 rumah tangga yang tersebar di 497 kabupaten/kota di seluruh provinsi di Indonesia. Tujuan utama SUSENAS 2016 adalah untuk menyediakan informasi yang rinci bagi pemerintah tentang keadaan sosial dan ekonomi masyarakat Indonesia. Di Kabupaten Tapanuli Selatan, Susenas dilakukan di 56 blok sensus dengan jumlah responden 560 rumah tangga.

Variabel yang berpengaruh terhadap frekuensi perjalanan

Pada umumnya seseorang yang melakukan perjalanan bisa dimotivasi oleh beberapa hal, Menurut McIntosh dan Murphy (1977)), motivasi melakukan perjalanan dapat dibagi menjadi empat macam, yaitu sebagai berikut:

1. *Physical or physiological motivation* atau motivasi yang mempunyai sifat fisik atau fisiologis, contohnya seperti relaksasi, kesehatan, kenyamanan, berpartisipasi dalam kegiatan olahraga, bersantai, dan yang berhubungan dengan sifat fisik lain.
2. *Cultural motivation* atau motivasi budaya, adalah keinginan untuk mengetahui budaya, adat, tradisi, dan kesenian di daerah lain. Hal itu juga termasuk ketertarikan dari berbagai objek peninggalan budaya atau bangunan yang bersejarah.
3. *Social motivation* atau *interpersonal motivation* atau motivasi yang mempunyai sifat sosial, seperti mengunjungi teman dan keluarga, menemui mitra kerja, melakukan hal yang dianggap mendatangkan gengsi, seperti nilai prestise, melakukan ziarah, dan melakukan pelarian dari situasi-situasi yang membosankan.
4. *Fantasy motivation* atau motivasi karena fantasi, merupakan adanya fantasi bahwa di daerah lain, seseorang akan mampu lepas dari rutinitas keseharian yang menjemukan, dan *ego-enhancement* yang memberi kepuasan psikologis. Hal seperti itu juga disebut dengan status and *prestige motivation.*

Menurut Richardison dalam Pitana (2006) keputusan seseorang untuk melakukan perjalanan dipengaruhi oleh kuatnya faktor pendorong dan faktor penarik. Faktor pendorong dan penarik ini sesungguhnya merupakan faktor internal dan eksternal yang memotivasi seseorang untuk mengambil keputusan untuk melakukan perjalanan. Faktor pendorong umumnya bersifat sosial-psikologis, atau merupakan *person-spesific motivation*. Sedangkan menurut Assael, 1987 dalam (Tejo, 2008) pengambilan keputusan untuk melakukan perjalanan dapat didekati oleh teori pemasaran. Ada tiga faktor yang mempengaruhi pengambilan keputusan pengunjung, yaitu: (a) faktor individual pengunjung yang meliputi pendidikan dan penghasilan pengunjung serta psikologis, (b) pengaruh lingkungan, dan (c) strategi pemasaran. Strategi pemasaran merupakan variabel yang dapat dikontrol oleh pemasar dalam usaha memberi informasi dan mempengaruhi pengunjung. Variabel ini meliputi produk, harga dan promosi.

Pada penelitian ini variabel yang digunakan dalam mencari variabel prediktor dibatasi pada faktor internal pribadi yang melakukan perjalanan yaitu jenis kelamin, umur, tingkat pendidikan dan sektor pekerjaan.

Zero-Inflated Poisson

Suatu peristiwa count mengacu pada berapa kali peristiwa itu terjadi. Metode utama yang dikembangkan untuk memahami distribusi dari data cacah (count) adalah regresi Poisson, dan kemudian menjadi metode standar yang digunakan untuk memodelkan data cacah. Peubah acak Y dikatakan mengikuti distribusi Poisson, jika fungsi kepadatan peluangnya berbentuk:

$$p(y) = \frac{e^{-\mu}\mu^{y}}{y!}$$

Pada regresi Poisson terdapat asumsi yang harus dipenuhi yaitu mean dan varian harus bernilai sama, namun seringkali dijumpai data count dengan varian lebih besar dari meannya yang disebut overdispersi. Overdispersi pada regresi Poisson dapat menyebabkan standard error dari taksiran parameter regresi yang dihasilkan memiliki kecenderungan untuk menjadi lebiih rendah dari seharusnya. Salah satu metode yang dikembangkan untuk mengatasi masalah overdispersi adalah model regresi Binomial Negatif. Model regresi Binomial Negatif dapat digunakan baik dalam keadaan equidispersion ataupun overdispersi. Model regresi Binomial Negatif yang dibangun memiliki sebaran Binomial Negatif dengan parameter μ dan k, dimana $\mu = \alpha\beta$ dan $k = 1 / \alpha$, sehingga mean dan varian masing-masing adalah sebagai berikut.

$$E(Y) = \mu \text{ dan } Var(Y) = \mu + k\mu^2$$

Salah satu permasalahan regresi Poisson yaitu excess zeros, dimana jumlah respon "0" yang diamati melebihi jumlah respon "0" yang diperkirakan oleh model. Data yang overdispersi terkadang disebabkan oleh excess zeros. Dalam hal ini, munculnya kelebihan nol dapat diatasi dengan menggunakan model yang mengakomodasi overdispersi. Excess zeros dapat dilihat pada proporsi variabel respon yang bernilai nol lebih besar dari data diskrit lainnya. Model Zero-Inflated Poisson pertama kali diperkenalkan oleh Lambert (1992). Jika Yi merupakan variabel acak independen yang berdistribusi ZIP, nilai nol pada observasi diduga muncul dalam dua cara yang sesuai untuk keadaan (state) yang terpisah. Keadaan pertama disebut zero-state terjadi dengan probabilitas ω dan menghasilkan hanya observasi bernilai nol, sementara keadaan kedua disebut Poisson state terjadi dengan probabilitas $(1 - \omega)$. Sehingga model regresi *Zero-Inflated Poisson* didefinisikan sebagai berikut:

$$P(Y_i = y_i) = \begin{cases} \omega + (1-\omega)e^{-\lambda}, & untuk\ Y_i = 0 \\ (1-\omega)\dfrac{e^{-\lambda}\lambda^{y_i}}{y_i!}, & untuk\ Y_i = 1,2,\dots,n \end{cases}$$

atau bisa kita tuliskan dalam bentuk logit sebagai berikut

$$\omega = \frac{\exp(Z_i^T\lambda)}{1+\exp(Z_i^T\lambda)}$$

model regresi ZIP dengan parameter $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)^{\boldsymbol{T}}$ dan $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_n)^{\boldsymbol{T}}$ dibentuk oleh

$$log\ (\boldsymbol{\lambda}) = \boldsymbol{X\beta} \qquad logit\ (\boldsymbol{\omega}) = log\ (\boldsymbol{\omega}/(1-\ \boldsymbol{\omega})) = \boldsymbol{X\alpha}$$

dengan X adalah matriks kovariat yang terdiri dari variabel-variabel prediktor, β dan $\boldsymbol{\alpha}$ merupakan parameter koefisien regresi untuk poisson part dan zero inflation (logistic) part serta log merupakan fungsi logaritma natural (ln). Mean dan varians dari model ZIP sebagai berikut:

$$E(Y) = (1-\omega)\lambda;\ Var\ (Y) = E(Y) + E(Y)(\ \lambda - E(Y))$$

Pengujian Overdispersi

Model standar Poisson GLM mensyaratkan $E(y) = \mu$ yang diasumsikan sama dengan varians $VAR(y) = \mu$. Test Dispersi memberikan hipotesis bahwa asumsi ini berlaku (equidispersion) terhadap alternatif bahwa variansnya adalah dari bentuk:

$$VAR(y) - \mu + \alpha.trafo(\mu)$$

Overdispersion terjadi jika nilai $\alpha > 0$ dan underdispersion terjadi jika nilai $\alpha < 0$. Koefisien dapat diperkirakan dengan regresi OLS tambahan dan diuji dengan statistik t (atau z) yang sesuai dengan standar normal asimtotik berdasarkan hipotesis nol.

Zero Inflated Negative Binomial

Berdasarkan analisis yang dilakukan dengan ZIP dapat diketahui apakah asumsi *mean* sama dengan variansinya atau tidak. Jika ternyata diketahui rata-ratanya tidak sama dengan variansinya (overdispersi), maka masalah ini dapat ditanggulangi dengan menerapkan model *Zero Inflated Negative Binomial* (ZINB). Pada model ini data cacahan (*counting*) diasumsikan diperoleh dari dua proses, yaitu proses pertama merupakan munculnya data *count* yang hanya berisi nilai nol ditentukan dengan probabilitas p. Proses kedua yaitu nilai nol dan nilai positif pada data cacahan, kedua-duanya dihasilkan oleh suatu proses yang mengikuti distribusi *Negative Binomial* (NB). Model ZINB diperoleh dari persamaan:

$$F_i = \frac{\exp(\mathbf{z}_i\gamma)}{1+\exp(\mathbf{z}_i\gamma)}$$

atau

$$\log it(F_i) = \log\left(\frac{F_i}{1-F_i}\right) = \mathrm{z}_i\gamma;\ (i\ =\ 1,2,\dots,n)$$

Dengan $g(y_i)$ adalah fungsi distribusi NB, sehingga dihasilkan model ZINB sebagai berikut:

$$Pr(Y_i = y_i|\mathrm{x}_i,\mathrm{z}_i) = \begin{cases} F_i + (1-F_i)(1+\alpha\mu_i)^{-\alpha^{-1}} & ;y_i = 0 \\ (1-F_i)\dfrac{\Gamma(y_i+\alpha^{-1})}{y_i!\,\Gamma(\alpha^{-1})}\left(\dfrac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}}\left(\dfrac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y_i} & ;y_i > 0 \end{cases}$$

Parameter-parameter model ZINB adalah $\mu_i = (\mu_1,\mu_2,\dots,\mu_n)'$, $F_i = (F_1,F_2,\dots,F_n)'$ dan merupakan parameter dispersi. Bentuk μ_i dan F_i, masing-masing memenuhi $\mu_i = \exp(\mathrm{x}_i\beta)$ atau $\log(\mu_i) = \mathrm{x}_i\beta;\ i\ =\ 1,2,\dots,n$ dan

$$F_i = F(\mathbf{z}_i\gamma) = \frac{\exp(\mathbf{z}_i\gamma)}{1+\exp(\mathbf{z}_i\gamma)}$$

Dengan $\boldsymbol{x_i}$ adalah vektor kovariat, dalam hal ini terdiri dari variabel-variabel penjelas/prediktor yang masing-masing mempengaruhi rata-rata distribusi *Negative Binomial* pada kelompok risiko tinggi dengan parameter $\boldsymbol{\beta} = (\beta_1,\beta_2,\dots,\beta_p)$, sedangkan $\boldsymbol{z_i}$ adalah vektor kovariat yang mempengaruhi probabilitas pada kelompok risiko rendah dengan parameter $\boldsymbol{\gamma} = (\gamma_1,\gamma_2,\dots,\gamma_p)$. Dalam pemodelan regresi ZINB, jika $\boldsymbol{F} = 0$, maka model akan tereduksi menjadi model NB, jika $\alpha = 0$, maka model akan tereduksi menjadi model ZIP, sedangkan jika $\boldsymbol{F} = 0$ dan $\alpha = 0$, maka model akan tereduksi menjadi model Poisson (Kibria, 2006). Rata-rata dan varians model ZINB berturut-turut adalah $E(y_i|\mathrm{x}_i,\mathrm{z}_i) = (1-F_i)\mu_i$ dan $ar\,(y_i|\mathrm{x}_i,\mathrm{z}_i) = (1-F_i)\mu_i(1+F_i\mu_i+\mu_i\alpha)$. Dengan mensubstitusikan persamaan diatas maka model ZINB dapat ditulis sebagai berikut:

$$Pr(Y_i = y_i|\mathrm{x}_i,\mathrm{z}_i) = \begin{cases} \dfrac{\exp(\mathrm{z}_i\gamma)}{1+\exp(\mathrm{z}_i\gamma)} + \dfrac{1}{1+\exp(\mathrm{z}_i\gamma)}(1+\exp(\mathrm{x}_i\beta))^{-\alpha^{-1}} & ;y_i = 0 \\ \dfrac{1}{1+\exp(\mathrm{z}_i\gamma)}\dfrac{\Gamma(y_i+\alpha^{-1})}{y_i!\,\Gamma(\alpha^{-1})}\left(\dfrac{\alpha^{-1}}{\alpha^{-1}+\exp(\mathrm{x}_i\beta)}\right)^{-\alpha} & ;y_i > 0 \end{cases}$$

Pemilihan Model Terbaik

Model terbaik dapat ditentukan dengan beberapa ukuran/statistik, di antaranya adalah dengan melihat nilai Akaike's Information Criterion (AIC). Model dengan nilai AIC terkecil merupakan model yang terbaik.

$$AIC = -2l(\Omega) + 2p$$

Dengan l adalah nilai *ln likelihood* dari model, p adalah banyaknya parameter dalam model dan $\Omega = (\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\alpha})$.

Bahan dan Metode

Sumber Data dan Variabel Penelitian

Data yang digunakan dalam makalah ini adalah hasil Survei Sosial Ekonomi Nasional (SUSENAS) bulan Maret 2016 Kabupaten Tapanuli Selatan. Survei ini dilakukan oleh Badan Pusat Statistik (BPS). Variabel-variabel yang digunakan dalam makalah ini merupakan karakteristik yang didapat dari SUSENAS Maret 2016 Kabupaten Tapanuli Selatan. Variabel respon yang digunakan yaitu jumlah perjalanan yang dilakukan anggota rumah tangga yang merupakan data frekuensi (*count*). Variabel bebas yang digunakan yaitu jenis kelamin anggota rumah tangga (ART), usia ART, tingkat pendidikan ART, dan sektor/lapangan usaha dari pekerjaan ART. Secara detail, variabel-variabel yang digunakan dalam makalah ini disajikan dalam tabel berikut:

Tabel 2. Variabel-variabel penelitian.

Variabel		Deskripsi	Kategori
Respon	Y	Jumlah perjalanan yang dilakukan ART dalam 6 bulan terakhir	
Bebas	X_1	Jenis Kelamin ART	0: perempuan 1: laki-laki
	X_2	Usia ART	Data kontinu
	X_3	Tingkat Pendidikan ART	0: SLTP kebawah 1: SLTA keatas
	X_4	Sektor pekerjaan/ lapangan usaha ART	0: pertanian 1: non pertanian

Langkah-Langkah Analisis

Langkah-langkah analisis dalam penelitian adalah sebagai berikut:

- Melakukan pemeriksaan multikolinieritas pada variabel bebas
- Memodelkan data dengan menggunakan metode *Zero Inflated Poisson* (ZIP)
- Memeriksa overdispersi/underdispersi pada model ZIP
- Memodelkan data dengan menggunakan metode *Zero Inflated Negative Binomial* (ZINB) jika terjadi overdispersi/underdispersi pada model ZIP
- Membuat kesimpulan

Hasil dan Pembahasan

Analisis Deskriptif

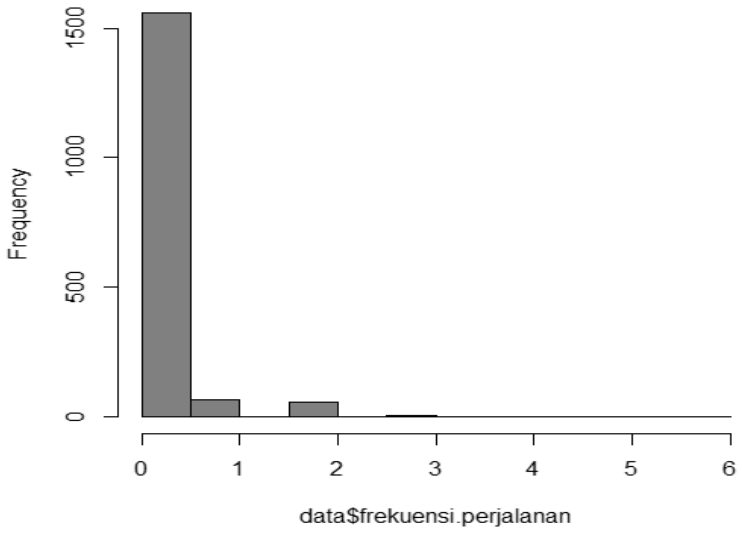
Dari data diketahui jumlah sampel yang digunakan sebanyak 554 rumah tangga yang terdiri dari 1.694 individu hasil SUSENAS Bulan Maret 2016 di Kabupaten Tapanuli Selatan. Ringkasan data sebagai berikut:

Tabel 3. Statistik deskriptif variabel-variabel penelitian.

Variabel	Jumlah Observasi	*Rataan*	Simpangan Baku	*Min*	*Max*
Jumlah Perjalanan (Y)	1.694	0,13	0,50	0	6
Jenis Kelamin (X_1)	1.694	0,50	0,5	0	1
Usia (X_2)	1.694	34,98	18,05	10	90
Tingkat Pendidikan (X_3)	1.694	0,28	0,45	0	1
Sektor Pekerjaan (X_4)	1.694	0,52	0,50	0	1

Rataan dan standar deviasi X_1, X_3, dan X_4 diperoleh dengan kaidah penghitungan *mean* dan simpangan baku untuk data proporsi. Dari informasi di atas dapat diketahui bahwa kabupaten Tapanuli Selatan secara umum masih memiliki karakteristik perdesaan yang ditanda dengan dominasi yang relatif tinggi pada tingkat pendidikan rendah, pekerjaan di sektor pertanian masih cukup tinggi dan

frekuensi bepergian yang masih rendah. Respon pada pertanyaan frekuensi bepergian dalam SUSENAS Maret 2016 banyak bernilai nol, proporsinya mencapai lebih dari 90 persen dengan distribusi sebagai berikut:



Gambar 1. Grafik Jumlah Perjalanan ART.

Uji Multikolinieritas

Dengan VIF diperoleh hasil uji sebagai berikut:

Tabel 4. Nilai VIF variabel.

Variabel	Nilai VIF
Jenis Kelamin (X_1)	1.018
Umur (X_2)	1.126
Tingkat Pendidikan (X_3)	1.004
Sektor Pekerjaan (X_4)	1.127

Nilai VIF yang lebih kecil dari 10 menunjukkan bahwa tidak ada kolinieritas antar variabel bebas.

Hasil Pemodelan Analisis Regresi *Zero Inflated Poisson*

Regresi *Zero Inflated Poisson* dipilih untuk mengatasi *excess zero* pada data penelitian. Untuk mengetahui apakah variabel independen secara simultan berpengaruh pada dependen maka dilakukan pengujian serentak *(overall test)* dengan l*ikelihood ratio test* (LRT) sebagai berikut:

H_0 : Semua variabel bebas pada model tidak mempengaruhi frekwensi bepergian baik pada bagian *poisson* maupun *zero inflation (logistic)* ($\beta_1 = \beta_2 = \ ... \ = \beta_k = \alpha_1 = \alpha_2 = \ ... \ = \alpha_k = 0$ dimana k adalah banyaknya variabel bebas)

H_1 : Setidaknya ada satu variabel bebas pada model yang mempengaruhi frekwensi bepergian baik pada bagian *poisson* maupun *zero inflation (logistic)* (setidaknya ada satu $\beta_i \neq 0$ atau $\alpha_i \neq 0$ dimana $i = 1, 2, \ldots, k$).

Hasilnya diperoleh nilai *log-likelihood* sebesar 593,9 dan selanjutnya dibandingkan dengan nilai *chi square* pada taraf uji 5 persen $(x^2 0{,}05;\ 10 = 18{,}307)$. H_0 ditolak, dapat dikatakan bahwa setidaknya ada satu variabel bebas dalam model yang berpengaruh signifikan terhadap frekuensi bepergian pada bagian *poisson* maupun *zero inflation (logistic)*.

Hasil estimasi koefisien parameter model regresi *Zero Inflated Poisson* disajikan sebagai berikut:

Tabel 5. Hasil Estimasi Parameter Model Regresi *Zero Inflated Poisson*.

Koefisien	*Estimasi*	*Simpangan Baku*	*z-value*	*p-values*
Poisson part				
β_0	-1.111905	0.407182	-2.731	0.00632 *
β_1	0.358605	0.209175	1.714	0.08646**
β_2	0.019072	0.006248	3.052	0.00227 *
β_3	0.192826	0.213871	0.902	0.36727
β_4	0.317684	0.213197	1.490	0.13620
zero inflation (logistic) part				
α_0	1.9466075	0.4328288	4.497	6.88e-06 *
α_1	0.3130653	0.2390299	1.310	0.1903
α_2	-0.0006707	0.0068762	-0.098	0.9223
α_3	-0.5463083	0.2479238	-2.204	0.0276 *
α_4	0.0433494	0.2434235	0.178	0.8587

Tahap selanjutnya perlu dilakukan pengujian apakah data poisson terpancung dari proses ZIP tidak mengandung dispersi. Untuk mengetahui keberadaan dispersi dilakukan uji skor.

Uji Skor

Hasil uji skor diperoleh nilai $Z = -12{,}927$ dangan $p - value = 1$ artinya bahwa untuk data pada model poisson terpancung terjadi overdispersi atau underdispersi. Estimasi dari data sampel diperoleh nilai dispersi sebesar 0,331. Berarti data yang kita miliki pada piosson terpancung underdispersi.

Analisis Model *ZINB*

Overdispersi maupun underdispersi dapat diatasi oleh model *Zero Inflated Negative Binomial* (ZINB). Dari regresi diperoleh informasi sebagai berikut:

Tabel 6. Hasil Estimasi Parameter Model Regresi *Zero Inflated Negative Binomial*.

Koefisien	Estimasi	Simpangan Baku	*z-value*	*p-values*
Negative Binomial (log-link) part				
β_0	-1.11196	0.407215	-2.731	0.00632**
β_1	0.358611	0.209189	1.714	0.08647
β_2	0.019072	0.006249	3.052	0.00227**
β_3	0.192826	0.213884	0.902	0.3673
β_4	0.317694	0.213209	1.49	0.13621
zero inflation (logit-link) part				
α_0	1.9465477	0.4328567	4.497	6.89E-06***
α_1	0.3130713	0.2390387	1.31	0.1903
α_2	-0.0006707	0.0068765	-0.098	0.9223
α_3	-0.5463166	0.2479318	-2.203	0.0276*
α_4	0.0433612	0.2434315	0.178	0.8586

Pada bagian binomial negatif dengan $\alpha = 5$ persen hanya intersep dan variabel usia yang berpengaruh signifikan dalam model, untuk $\alpha = 10$ persen didapati bahwa intersep, umur, dan jenis kelamin berpengaruh signifikan pada model frekuensi bepergian. Pada taraf uji 10 persen setiap penambahan 1 tahun umur ART secara rata-rata meningkatkan 0,02 kali frekuensi bepergian, dengan asumsi variabel lain konstan. Sementara itu, dengan rasio odds, laki-laki 1,43 kali lebih mungkin bepergian dibandingkan perempuan. Tingkat pendidikan dan sektor pekerjaan tidak berpengaruh signifikan terhadap peningkatan/penurunan rata-rata frekuensi bepergian penduduk. Hal ini mungkin terjadi karena respon klasifikasi pada kedua variabel tersebut yang relatif homogen menyebabkan independen tidak mampu menjelaskan dependen.

Pada bagian *zero inflation binomial (logistic-link)*, variabel bebas yang berpengaruh pada taraf uji 5 persen adalah tingkat pendidikan. ART dengan pendidikan SLTA ke atas kemungkinannya untuk tidak melakukan perjalanan 0,58 lebih sedikit daripada ART dengan pendidikan SLTP ke bawah pada kondisi variabel lain konstan.

Kesimpulan

1. Model yang tepat untuk frekuensi bepergian yang dilakukan selama 6 bulan terakhir di Kabupaten Tapanuli Selatan, Sumatera Utara, periode Maret 2016 adalah model *Zero Inflated Negative Binomial* (ZINB).
2. Pada *ZINB* bagian regresi Binomial Negatif (log-link) variabel yang berpengaruh signifikan terhadap frekuensi bepergian adalah variabel Usia dan Jenis Kelamin, sedangkan pada *zero inflation (logistic-link) part* variabel yang berpengaruh adalah variabel Tingkat pendidikan.

Referensi

[1] **Agresti**, **Alan**. 2015 Foundation of Liner and Generalized Linier Model,
[2] **Emi Hayati**, 2013. Faktor faktor yang mempengaruhi motivasi wisatawan berkunjung ke candi muara takus Kota Kampar. Universitas Riau. Riau
[3] **Hary Wahyu Pratama**, 2003. Faktor Faktor Yang Mempengaruhi Perilaku Pengunjung Taman Wisata Alam Umbul Sidomukti Kabupaten Semarang. Universitas Semarang. Semarang
[4] **Husaen Hasan**, 2009. Faktor-Faktor Yang Mempengaruhi Keputusan Wisatawan dalam Melakukan Kegiatan Wisata di Kota Tidore Kepulauan. Sekolah Tinggi Manajemen dan Informatika Kota Tidore Kepulauan. Kota Tidore Kepulauan.
[5] **Kemenpar**. 2014. NESPRNAS: Neraca Satelit Pariwisata Nasional. Jakarta: Pusdatin Kemenpar
[6] **Lambert Diane**, 1992. Zero Inflated Poisson Regression, with an Application to Defects in Manufacturing. Technometrics Journal vol 34 pp 1-14
[7] **McIntosh**, 1998. The Tourist: A New Theory of Leisure Class, New York, Schocken
[8] **Ridout**, **Dimitrie**, **Hindie** 1998. Model For Count Data With Many Zeros. International Biometric Conference

THIS PAGE INTENTIONALLY LEFT BLANK