



TESIS - SS14 2501

**MODEL REGRESI PROBIT SPASIAL DENGAN PENDEKATAN
*RECURSIVE IMPORTANCE SAMPLING***

(Studi Kasus: Indeks Kesehatan Tahun 2016 di Pulau Papua)

TAUFIQ FAJAR DEWANTO
NRP. 06211650017007

DOSEN PEMBIMBING
Dr. Vita Ratnasari, S.Si., M.Si.
Dr. Purhadi, M.Sc.

**PROGRAM PASCASARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018**



TESIS - SS14 2501

**MODEL REGRESI PROBIT SPASIAL DENGAN PENDEKATAN
*RECURSIVE IMPORTANCE SAMPLING***

(Studi Kasus: Indeks Kesehatan Tahun 2016 di Pulau Papua)

TAUFIQ FAJAR DEWANTO
NRP. 06211650017007

DOSEN PEMBIMBING
Dr. Vita Ratnasari, S.Si., M.Si.
Dr. Purhadi, M.Sc.

**PROGRAM PASCASARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018**



THESIS - SS14 2501

**SPATIAL PROBIT REGRESSION MODEL WITH RECURSIVE
IMPORTANCE SAMPLING APPROACH
(Case Study: Health Index 2016 in Papua Island)**

**TAUFIQ FAJAR DEWANTO
NRP. 06211650017007**

SUPERVISORS

**Dr. Vita Ratnasari, S.Si., M.Si.
Dr. Purhadi, M.Sc.**

**MAGISTER PROGRAM
DEPARTEMEN OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTATION, AND DATA SCIENCE
INSTITUTE OF TECHNOLOGY SEPULUH NOPEMBER
SURABAYA
2018**

**MODEL REGRESI PROBIT SPASIAL DENGAN
PENDEKATAN RECURSIVE IMPORTANCE SAMPLING
(Studi Kasus: Indeks Kesehatan Tahun 2016 di Pulau Papua)**

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh:

**TAUFIQ FAJAR DEWANTO
NRP. 06211650017007**

Tanggal Ujian : 23 Januari 2018
Periode Wisuda : Maret 2018

Disetujui Oleh:



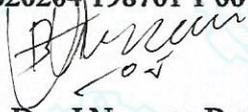
1. Dr. Vita Ratnasari, S.Si., M.Si.
NIP. 19700910 199702 2 001

(Pembimbing I)



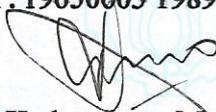
2. Dr. Puhadi, M.Sc.
NIP. 19620204 198701 1 001

(Pembimbing II)



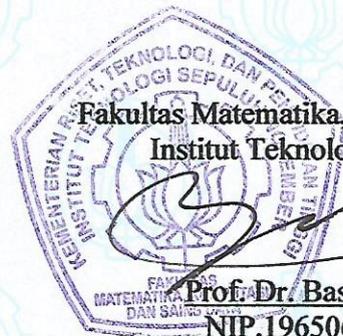
3. Prof. Dr. Drs. I Nyoman Budiantara, M.Si.
NIP. 19650603 198903 1 003

(Penguji)



4. Dr. Kadamanto, M.A.
NIP. 19710707 199312 1 001

(Penguji)



Dekan
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember



Prof. Dr. Basuki Widodo, M.Sc.
NIP. 19650605 198903 1 002

MODEL REGRESI PROBIT SPASIAL DENGAN PENDEKATAN RECURSIVE IMPORTANCE SAMPLING

(Studi Kasus: Indeks Kesehatan Tahun 2016 di Pulau Papua)

Nama Mahasiswa : Taufiq Fajar Dewanto
NRP : 06211650017007
Dosen Pembimbing : Dr. Vita Ratnasari, S.Si., M.Si.
Dr. Purnadi, M.Sc.

ABSTRAK

Model regresi probit merupakan model yang digunakan untuk menganalisis hubungan antara variabel respon yang bersifat kategorik, dengan variabel prediktor yang bersifat numerik, kategorik ataupun gabungan dari keduanya. Pada beberapa kasus, data yang dihasilkan pada variabel respon dalam regresi probit yang bersifat kategorik dapat juga dipengaruhi oleh kebiasaan ataupun opini dari wilayah di sekitarnya yaitu adanya autokorelasi spasial. Apabila pengaruh autokorelasi spasial ini diabaikan dan tetap menggunakan model probit yang standar, maka akan menghasilkan estimasi parameter yang bias dan tidak konsisten. Untuk menangani hal tersebut telah dikembangkan sebuah metode analisis probit yang telah memperhitungkan unsur spasial di dalamnya yaitu regresi probit spasial. Dari beberapa metode estimasi dengan menggunakan regresi probit spasial, metode *recursive importance sampling* (RIS) melakukan yang terbaik atau mengungguli metode lainnya dalam hal akurasi. Metode tersebut kemudian akan diaplikasikan dalam kasus indeks kesehatan di Pulau Papua karena beberapa alasan. Hasil evaluasi dari pencapaian target MDGs pada tahun 2015 menunjukkan bahwa masih terdapat 3 target tujuan MDGs yang masih belum dicapai oleh Indonesia, dimana target tersebut berkaitan tentang bidang kesehatan. Indeks kesehatan merupakan salah satu komponen penyusun IPM yang menunjukkan keberhasilan pembangunan pada dimensi kesehatan. Pulau Papua memiliki disparitas indeks kesehatan yang terbesar yaitu antara 53,08 - 79,85 dan jika dilihat lebih dalam, wilayah kabupaten/kota dengan peringkat 10 terendah, 8 diantaranya berasal dari Pulau Papua yaitu Provinsi Papua dan Papua Barat, serta adanya kemiripan nilai indeks kesehatan antar wilayah yang berdekatan secara geografis diduga akibat adanya autokorelasi spasial. Variabel prediktor yang signifikan dalam membentuk model regresi probit spasial terhadap indeks kesehatan adalah pengeluaran per kapita, persentase rumah tangga yang memiliki sumber air minum layak, persentase penduduk yang memiliki kebiasaan merokok, rata-rata lama sekolah, dan rasio fasilitas kesehatan terhadap jumlah desa. Model tersebut memiliki nilai akurasi sebesar 61,90 persen.

Kata kunci: Regresi Probit, Regresi Probit Spasial, RIS, Indeks Kesehatan.

(Halaman ini sengaja dikosongkan)

**SPATIAL PROBIT REGRESSION MODEL WITH RECURSIVE
IMPORTANCE SAMPLING APPROACH
(Case Study: Health Index 2016 in Papua Island)**

Name of Student : Taufiq Fajar Dewanto
NRP : 06211650017007
Supervisor : Dr. Vita Ratnasari, S.Si., M.Si.
Dr. Purhadi, M.Sc.

ABSTRACT

Probit regression model is a model used to analyze the relation between categorical response variables and numerical predictor variables or the combination of both. In some cases, the result of response variables data in categorical probit regression can be influenced by habits and or public opinion in the certain region which there is spatial autocorrelation. If the effect of spatial autocorrelation ignored and still using the standard probit model, it will produce a bias and inconsistent parameter estimation. To handle the case, a probit analysis method that considers the spatial element in it, which called spatial probit regression, has been developed. From several estimation methods that using spatial probit regression, the recursive importance sampling (RIS) method works optimally or surpass other methods in terms of accuracy. This method will be applied in case of health index on Papua Island for several reasons. The evaluation results of the MDGs target achievement by 2015, indicate that there are 3 MDGs targeted goals have not been achieved by Indonesia, in which the target is concerned on health. The health index is one of the IPM components that shows the succession of development on the health dimension. Papua Island has the largest health disparity index between 53,08 – 79,85. If it's seen more comprehensively, by the region of districts with the lowest ranking of 10, 8 of them belongs to Papua Island. Those are Provinsi Papua and Papua Barat and the similarity of the health index values between geographically adjacent areas allegedly due to spatial autocorrelation. Predictor Variables that is significant in developing spatial probit regression on the health index are; the cost per capita, percentage of household with decent drinking water source, percentage of population with smoking habits, average length of experiencing school , and ratio of health facility to village amount. The model has accuracy value of 61,90 percent.

Keywords: Probit Regression, Spatial Probit Regression, RIS, Health Index.

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kehadirat Allah SWT atas segala limpahan rahmat dan kemurahan-Nya sehingga penulis dapat menyelesaikan tesis dengan judul **“Model Regresi Probit Spasial Dengan Pendekatan Recursive Importance Sampling (Studi Kasus: Indeks Kesehatan Tahun 2016 di Pulau Papua)”**. Tesis ini disusun dalam rangka memenuhi salah satu syarat untuk meraih gelar Magister Sains di Program Pasca Sarjana Statistika Institut Teknologi Sepuluh Nopember (ITS) Surabaya.

Keberhasilan penyusunan tesis ini tidak terlepas dari bantuan, bimbingan, petunjuk, dan dukungan dari berbagai pihak. Sehubungan dengan itu, teriring rasa syukur dan doa, penulis menyampaikan ucapan terima kasih kepada:

1. Badan Pusat Statistik (BPS) yang telah memberi kesempatan serta beasiswa yang diberikan kepada penulis untuk melanjutkan studi program S2 di ITS.
2. Ibu Dr. Vita Ratnasari, S.Si. M.Si. dan Bapak Dr. Purhadi, M.Sc. selaku dosen pembimbing yang telah memberikan arahan, bimbingan dan motivasi dalam penyelesaian tesis ini.
3. Bapak Prof. Dr. Drs. I Nyoman Budiantara, M.Si. dan Bapak Dr. Kadarmanto, M.A. selaku dosen penguji yang telah memberikan saran dan koreksi atas penulisan tesis ini.
4. Bapak Dr. Suhartono, M.Sc selaku Ketua Departemen Statistika dan Bapak Dr.rer.pol. Heri Kuswanto, S.Si, M.Si selaku Ketua Program Studi Pascasarjana Departemen Statistika FMKSD ITS atas arahan dan bantuannya selama penulis menempuh pendidikan di Program Magister Departemen Statistika ITS.
5. Bapak Ibu dosen Statistika ITS yang telah mencurahkan ilmu dan pengalamannya selama proses studi.
6. Seluruh jajaran staf jurusan Statistika ITS atas segala bimbingan dan fasilitas yang telah diberikan.
7. Istri tercinta, Rahmadanie Sapta Irevanie, terima kasih atas segala dukungan, nasehat, kesabaran, doa dan cintanya kepada penulis sehingga dapat menyelesaikan

studi tepat waktu. Terima kasih karena selalu mendukung penulis pada masa-masa sulit dalam menyelesaikan tesis ini.

8. Kedua buah hati, Asheeqa Sheza Ramadhani dan Azzamir Rafka Muhammad atas kesabarannya dalam menghadapi kehidupan yang jauh dari ayahnya. Senyum, tawa, dan canda kalian selalu menjadi menyejukkan hati dan mampu menjadikan semangat dalam menyelesaikan tesis ini.
9. Bapak dan Ibu, Bapak dan Ibu mertua, atas segala doa, nasehat dan dukungan yang tulus kepada penulis.
10. Teman kosan: Sony Triasmoro, Fendy Ardiyanto dan Suprpto yang selalu memberikan bantuan dan dibuat repot oleh penulis. Teman jalan-jalan: Fieldri Sinaga, Ratih Kusumadewi, Prih Wismaningsih dan Mike Fitriandari, terima kasih atas keceriaan, bantuan dan kebersamaan yang takkan terlupa.
11. Teman seperjuangan S2 BPS Batch 10 atas bantuannya dalam menyelesaikan studi. Terima kasih atas kekompakan ini. Bersyukur telah bersama dengan kalian.
12. Teman-teman BPS Kabupaten Jayapura atas dukungannya
13. Teman-teman S2 ITS regular angkatan 2016 dan semua pihak yang telah membantu penyelesaian tesis ini.

Surabaya, Januari 2018

Penulis

DAFTAR ISI

ABSTRAK.....	iii
ABSTRACT.....	v
KATA PENGANTAR	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xv
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian	6
1.4 Batasan Masalah	6
1.5 Manfaat Penelitian	6
BAB 2 TINJAUAN PUSTAKA	7
2.1 Distribusi Bernoulli.....	7
2.2 Model Regresi Probit	8
2.3 Uji Multikolinieritas.....	11
2.4 Model Regresi Spasial	12
2.4.1 Matriks Pembobot Spasial	13
2.4.2 Uji Dependensi Spasial	15
2.5 Model Regresi Probit Spasial.....	16
2.6 Recursive Importance Sampling (RIS)	18
2.7 Pengujian Parameter	21
2.8 Kriteria Keباikan Model	22
2.9 Indeks Kesehatan	24
2.10 Kajian Teori dan Penelitian Terdahulu	25
BAB 3 METODOLOGI PENELITIAN	31
3.1 Sumber Data.....	31
3.2 Variabel Penelitian.....	31
3.3 Tahapan Penelitian.....	34

BAB 4 HASIL DAN PEMBAHASAN	37
4.1 Estimasi Parameter Regresi Probit Spasial.....	37
4.2 Pemodelan Indeks Kesehatan di Pulau Papua dengan Model Regresi Probit Spasial.....	44
4.2.1 Gambaran Indeks Kesehatan Kabupaten/Kota di Pulau Papua...	44
4.2.2 Pendeteksian Multikolinearitas pada Variabel Prediktor	47
4.2.3 Uji Dependensi Spasial Indeks Kesehatan di Pulau Papua	48
4.2.4 Pengujian Parameter secara Serentak dan Parsial Model Regresi Probit Spasial.....	49
4.2.5 Intrepretasi Model Regresi Probit Spasial.....	51
4.2.6 Kriteria Keباikan Model Regresi Probit Spasial	60
BAB 5 KESIMPULAN DAN SARAN.....	63
5.1 Kesimpulan.....	63
5.2 Saran.....	63
DAFTAR PUSTAKA.....	65
LAMPIRAN	69
BIOGRAFI PENULIS.....	91

DAFTAR TABEL

Tabel 2.1 Ilustrasi Tabel <i>Confusion Matrix</i>	23
Tabel 3.1 Variabel Penelitian.....	32
Tabel 3.2 Definisi Operasional Variabel dan Satuan Variabel	33
Tabel 3.3 Struktur Data Penelitian.....	33
Tabel 4.1 Deteksi Multikolinieritas Variabel Prediktor.....	48
Tabel 4.2 Pengujian Signifikansi Parameter Secara Parsial Model Probit Spasial	51
Tabel 4.3 Hasil Pengelompokan Data Aktual dan Prediksi Berdasarkan Kabupaten/Kota di Pulau Papua dengan Model Probit Spasial.....	55
Tabel 4.4 Tabel <i>Confusion Matrix</i> Indeks Kesehatan Di Pulau Papua Tahun 2016	60

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi <i>Contiguity</i> (persinggungan).....	14
Gambar 2.2 Kerangka Pikir Faktor Yang Mempengaruhi Indeks Kesehatan.....	29
Gambar 3.1 Skema Tahapan Penelitian	36
Gambar 4.1 Peta Persebaran Indeks Kesehatan Menurut Kabupaten/Kota di Pulau Papua Tahun 2016	45
Gambar 4.2 <i>Boxplot</i> Data Aktual Indeks Kesehatan Menurut Variabel Prediktor	46
Gambar 4.3 Peta Hasil Prediksi Indeks Kesehatan dari Model Regresi Spasial.	54
Gambar 4.4 Peta Pengelompokan Data Aktual dan Prediksi Indeks Kesehatan	56
Gambar 4.5 <i>Boxplot</i> Data Prediksi Indeks Kesehatan Menurut Variabel Prediktor	57

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1. Data Variabel Penelitian.....	69
Lampiran 2. Matriks Pembobot (<i>Queen Contiguity</i>)	70
Lampiran 3. R code dan Output Uji Multikolinieritas dan Regresi Probit Spasial dengan RIS Simulator	73
Lampiran 4. R code dan Output Uji Dependensi Spasial dan Uji Parameter Secara Simultan dan Parsial.....	75
Lampiran 5. R code Prediksi Probabilitas, <i>Confusion Matrix</i> dan Efek Marginal	77
Lampiran 6. Prediksi Probabilitas Kabupaten/Kota di Pulau Papua dari Model SAR Probit	78
Lampiran 7. Efek Marjinal Variabel Pengeluaran Per Kapita Terhadap Indeks Kesehatan	80
Lampiran 8. Efek Marjinal Variabel Persentase Rumah Tangga dengan Air Minum Layak Terhadap Indeks Kesehatan	82
Lampiran 9. Efek Marjinal Variabel Persentase Penduduk yang Mempunyai Kebiasaan Merokok Terhadap Indeks Kesehatan.....	84
Lampiran 10. Efek Marjinal Variabel Rata-rata Lama Sekolah Terhadap Indeks Kesehatan	86
Lampiran 11. Efek Marjinal Variabel Proporsi Fasilitas Kesehatan Terhadap Indeks Kesehatan.....	88

(Halaman ini sengaja dikosongkan)

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Salah satu metode statistik yang dapat digunakan untuk mengetahui hubungan antara variabel-variabel yang dinyatakan dalam suatu bentuk persamaan matematik adalah analisis regresi. Pada analisis regresi, variabel di dalamnya dibedakan menjadi dua jenis, yaitu variabel respon dan variabel prediktor, dimana analisis ini digunakan untuk menggambarkan seberapa besar pengaruh variabel prediktor terhadap variabel respon. Pada umumnya data yang digunakan dalam variabel respon analisis regresi ini merupakan data numerik (metric), padahal seperti yang kita ketahui selain data numerik, data menurut jenis skala pengukurannya juga terdapat data kategorik (non metric). Gujarati (2004) mengemukakan bahwa apabila variabel respon berjenis kategorik, maka salah satu model yang mampu menyelesaikan adalah regresi probit.

Model regresi probit pertama kali diperkenalkan oleh Chester Bliss pada tahun 1934. Model ini merupakan model non linier yang digunakan untuk menganalisis hubungan antara variabel respon yang bersifat kategorik yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan 0 untuk menyatakan ketidakberadaan sebuah karakteristik, dengan variabel prediktor yang bersifat numerik, kategorik ataupun gabungan dari keduanya. Pada beberapa kasus, data yang dihasilkan pada variabel respon dalam regresi probit yang bersifat kategorik dapat juga dipengaruhi oleh kebiasaan ataupun opini dari wilayah di sekitarnya yaitu adanya autokorelasi spasial. Sebagai contoh, Case (1992) melakukan penelitian di Pulau Jawa tentang keputusan petani untuk memilih teknologi dalam alat pertanian yaitu alat pertanian modern atau tradisional. Dalam paparannya disebutkan bahwa dalam pengambilan keputusan untuk memilih teknologi dalam alat pertanian ternyata dipengaruhi oleh keputusan petani lain yang berada di wilayah sekitarnya. Contoh lain dikemukakan oleh Beron dan Vijberg (2004) bahwa keputusan untuk menaikkan atau tidak pajak suatu kota bergantung kepada keputusan yang diambil oleh wilayah di sekitarnya.

Dari kedua contoh tersebut dapat diambil sebuah kesimpulan, bahwa apabila pengaruh autokorelasi spasial ini diabaikan dan tetap menggunakan model probit yang standar, maka akan menghasilkan estimasi parameter yang bias dan tidak konsisten (McMillen, 1992). Untuk menangani hal tersebut telah dikembangkan sebuah metode analisis probit yang telah memperhitungkan unsur spasial di dalamnya yaitu regresi probit spasial yang pertama kali diteliti oleh McMillen (1992).

Hingga saat ini, terdapat empat metodologi analisis regresi probit spasial yang paling banyak digunakan. McMillen (1992) mengajukan metode expectation-maximization (EM) algorithm yaitu dengan menggantikan variabel laten kontinu dengan nilai harapannya kemudian menerapkannya dengan metode maximum likelihood. Hampir sama dengan McMillen (1992), LeSage (2000) juga mengganti variabel laten kontinu dengan nilai harapannya, namun mengaplikasikan dengan pendekatan Bayesian Gibbs sampler. Beron dan Vijverberg (2004) menyarankan untuk menggunakan recursive importance sampling (RIS) untuk metode maximum likelihood dikarenakan distribusi dari fungsi likelihood adalah multivariat normal. Pinkse dan Slade (1998) mengembangkan sebuah model yang berdasarkan Generalized Method of Moments (GMM).

Calabrese dan Elkind (2014), telah membandingkan hasil yang diperoleh dari keempat metode tersebut. EM algorithm yang diusulkan oleh McMillen (1992) menghasilkan perkiraan yang tidak konsisten dan tidak efisien. Metode GMM yang diterapkan oleh Pinkse dan Slade (1998), estimator mereka kurang efisien daripada estimator maximum likelihood. Estimator Gibbs sampler (LeSage, 2000) melakukan yang terbaik untuk autokorelasi spasial rendah, sedangkan recursive importance sampling (Beron dan Vijverberg, 2004) melakukan yang terbaik untuk autokorelasi spasial yang tinggi dan disimpulkan mengungguli ketiga metode lainnya dalam hal akurasi.

Salah satu isu yang menarik untuk dikaji dan dibahas di Indonesia adalah isu tentang kesehatan. Tujuan Pembangunan Milenium (Millenium Development Goals/MDGs) merupakan kesepakatan dari 189 negara yang tergabung dalam Perserikatan Bangsa-bangsa (PBB) yang dilaksanakan pada September tahun

2000 dan telah berakhir pada tahun 2015. Indonesia sebagai negara yang ikut mendeklarasikan MDGs, memiliki kewajiban untuk melaksanakan upaya untuk mencapai target dan memonitor perkembangan kemajuan pencapaiannya. Hasil evaluasi dari pencapaian target MDGs pada tahun 2015 menunjukkan bahwa masih terdapat 3 target tujuan MDGs yang masih belum dicapai oleh Indonesia yaitu menurunkan angka kematian ibu melahirkan, menurunkan penyebaran virus HIV/AIDS serta mengakses air bersih dan sanitasi dasar, yang ketiganya berkaitan tentang bidang kesehatan (HMPD, 2016)

Salah satu indikator keberhasilan pembangunan pada dimensi kesehatan dapat dilihat dari indikator angka harapan hidup saat lahir. Untuk dapat berumur panjang, diperlukan kesehatan yang lebih baik. Angka harapan hidup merupakan alat untuk mengevaluasi kinerja pemerintah dalam meningkatkan kesejahteraan penduduk pada umumnya, dan meningkatkan derajat kesehatan pada khususnya. Angka harapan hidup merupakan salah satu indikator yang digunakan untuk mengukur capaian pembangunan manusia di suatu wilayah melalui Indeks Pembangunan Manusia (IPM). IPM dibangun melalui pendekatan tiga dimensi dasar. Dimensi tersebut mencakup kesehatan, pendidikan, dan kehidupan layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor. Untuk mengukur dimensi kesehatan, digunakan indikator angka harapan hidup waktu lahir. Selanjutnya untuk mengukur dimensi pendidikan digunakan gabungan indikator harapan lama sekolah dan rata-rata lama sekolah. Adapun untuk mengukur dimensi hidup layak digunakan indikator kemampuan daya beli (BPS, 2015a).

Indikator angka harapan hidup saat lahir disajikan dalam bentuk indeks, yaitu indeks kesehatan. Indeks kesehatan merupakan salah satu komponen penyusun IPM, dengan bobot yang sama untuk masing-masing komponen. Menurut UNDP (2017) untuk mengukur perkembangan pembangunan pada suatu wilayah dapat digunakan IPM sebagai tolak ukurnya. IPM digunakan untuk mengklasifikasikan apakah sebuah wilayah termasuk kategori wilayah maju atau wilayah terbelakang, dimana dalam IPM dapat dikategorikan menjadi rendah, sedang, tinggi, dan sangat tinggi. Dengan demikian indeks kesehatan dapat dikategorikan seperti halnya IPM. Pengkategorian indeks kesehatan bertujuan

untuk mengelompokkan wilayah berdasarkan klasifikasinya, serta untuk mempermudah para pengambil kebijakan di suatu wilayah dalam merencanakan dan mengeluarkan kebijakan yang berkaitan dengan pemberdayaan sumber daya manusia di bidang kesehatan.

Capaian indeks kesehatan pada wilayah kabupaten/kota di Indonesia pada tahun 2015 menunjukkan terjadinya disparitas yang cukup besar, hal ini menunjukkan bahwa tingkat kesehatan di Indonesia belum merata. Apabila dilihat dari pulau-pulau besar yang ada di Indonesia, Pulau Papua memiliki disparitas yang terbesar yaitu antara 53,08 - 79,85. Jika dilihat lebih dalam, wilayah kabupaten/kota dengan peringkat 10 terendah, 8 diantaranya berasal dari Pulau Papua yaitu Provinsi Papua dan Papua Barat. Kemudian diketahui bahwa terdapat kemiripan nilai indeks kesehatan antar wilayah yang berdekatan secara geografis di Pulau Papua, dimana hal ini dapat diduga karena adanya keterkaitan hubungan antar wilayah atau efek spasial. Sehingga berdasarkan alasan-alasan tersebut, studi kasus pada penelitian ini difokuskan terhadap indeks kesehatan di Pulau Papua dan kemudian diterapkan dalam metode regresi probit spasial.

Penelitian menggunakan regresi probit spasial pernah dilakukan oleh Puspita (2013) yang memodelkan regresi probit spasial pada faktor-faktor yang mempengaruhi klasifikasi IPM di Pulau Jawa dengan metode estimasi parameternya MCMC Gibbs Sampler. Fahmi (2016) memodelkan regresi probit spasial pada IPM di Jawa Timur dengan penaksiran parameter menggunakan Algoritma EM (Ekspektasi Maksimalisasi). Beberapa penelitian tentang indeks kesehatan ataupun indikator penyusunnya telah beberapa kali dilakukan. Maully (2014) meneliti faktor-faktor yang mempengaruhi indeks kesehatan kabupaten/kota di Jawa Timur dengan menggunakan regresi logistik. Satomo (2017) memodelkan indeks kesehatan dan indeks pengeluaran kabupaten/kota di Indonesia dengan pendekatan model probit bivariat. Sugiantari (2013) memodelkan Angka Harapan Hidup dengan menggunakan metode regresi semiparametrik spline. Sementara itu, Juliandari (2014) memodelkan Angka Harapan Hidup dan Angka Kematian Bayi di Jawa Timur dengan pendekatan regresi nonparametrik spline birespon. Irevanie (2017) meneliti tentang

perbandingan metode Quantile Regression (QR) dan Geographically Wighted Regression (GWR) pada Angka Harapan Hidup di Indonesia.

Merujuk pada penelitian-penelitian sebelumnya tentang model regresi probit spasial serta indeks kesehatan maupun indikator penyusunan yang telah dilakukan sebelumnya, maka pada penelitian ini ingin mengkaji tentang indeks kesehatan menggunakan regresi probit spasial dengan menggunakan metode estimasi parameter recursive importance sampling (RIS) yang memiliki keunggulan dalam akurasi. Diharapkan melalui penelitian ini dapat memberikan hasil yang spesifik sehingga dapat diketahui faktor-faktor yang mempengaruhi klasifikasi indeks kesehatan di Pulau Papua.

1.2 Rumusan Masalah

Model regresi probit merupakan model yang dapat digunakan pada saat variabel responnya berjenis kategorik. Namun, apabila ada pengaruh autokorelasi spasial, penggunaan model regresi probit akan menghasilkan estimasi parameter yang bias dan tidak konsisten (McMillen, 1992). Sehingga untuk menyelesaikan masalah tersebut digunakan regresi probit spasial yang merupakan gabungan dari regresi probit dan regresi spasial. Dalam melakukan estimasi parameter dengan menggunakan regresi probit spasial, terdapat empat metode yang umum digunakan yaitu MLE dengan pendekatan *EM Algorithm*, MLE dengan pendekatan RIS, Bayesian Gibbs sampler, dan GMM. Dari keempat metode ini, RIS mengungguli metode yang lainnya dalam akurasi yang dihasilkan (Calabrese dan Elkind, 2014). Sehingga pada penelitian ini akan mengkaji tentang bagaimana memperoleh estimasi parameter dari model regresi probit spasial dengan pendekatan *Recursive Importance Sampling* (RIS).

Indeks kesehatan merupakan refleksi dari derajat kesehatan suatu masyarakat, dimana indeks ini digunakan untuk menilai keberhasilan pembangunan suatu wilayah dalam bidang kesehatan. Kegagalan pencapaian 3 tujuan dalam MDGs pada batas waktu terakhir yang ketiganya pada bidang kesehatan, disparitas indeks kesehatan yang lebar dan masih rendahnya pencapaian indeks kesehatan pada sebagian besar kabupaten/kota di Pulau Papua, serta adanya kemiripan indeks kesehatan di wilayah yang berdekatan di Pulau

Papua, maka penggunaan model regresi probit spasial dengan pendekatan *Recursive Importance Sampling* (RIS) akan diaplikasikan pada kasus klasifikasi indeks kesehatan di Pulau Papua, sehingga didapatkan variabel yang berpengaruh dengan menggunakan regresi probit spasial.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mengkaji estimasi parameter model regresi probit spasial.
2. Memodelkan faktor-faktor yang mempengaruhi klasifikasi indeks kesehatan di Pulau Papua sehingga didapatkan variabel yang berpengaruh dengan menggunakan regresi probit spasial.

1.4 Batasan Masalah

Adapun batasan dalam penelitian ini adalah sebagai berikut:

1. Penentuan bobot spasial hanya menggunakan pendekatan area dengan metode *Queen Contiguity*
2. Kategori pada regresi probit hanya menggunakan dua kategori atau biner
3. Model spasial yang digunakan adalah *Spatial Autoregressive* (SAR)

1.5 Manfaat Penelitian

Adapun Manfaat yang ingin dicapai dalam penelitian ini adalah:

1. Mendapatkan gambaran penggunaan metode regresi probit spasial.
2. Bagi pemerintah khususnya di Provinsi Papua dan Papua Barat, faktor-faktor yang secara signifikan berpengaruh dalam model dapat dijadikan sebagai masukan dalam pengambilan kebijakan yang dapat membantu memecahkan permasalahan kesehatan yang ada di Pulau Papua.
3. Bagi Badan Pusat Statistik (BPS), penelitian dengan metode regresi probit spasial ini dapat dijadikan referensi alternatif dalam melakukan analisis pada indikator-indikator makro berbentuk data kategorik yang dihasilkan BPS dengan mempertimbangkan unsur spasial.

BAB 2

TINJAUAN PUSTAKA

Pada bab ini akan dibahas mengenai teori statistika maupun non statistika yang mendasari penelitian tentang pemodelan regresi probit spasial pada klasifikasi indeks kesehatan di Pulau Papua. Teori statistika yang mendasari penelitian ini yaitu distribusi Bernoulli, model regresi probit, uji multikolinearitas, model regresi spasial dan model regresi probit spasial. Teori non statistika yang disajikan dalam bab ini adalah tentang indeks kesehatan serta kajian teori dan penelitian terdahulu.

2.1 Distribusi Bernoulli

Distribusi Bernoulli ditemukan oleh seorang ahli matematika berkebangsaan Swiss bernama Jacob Bernoulli. Distribusi ini bersumber dari Percobaan Bernoulli (*Bernoulli Trial*), yaitu percobaan yang menghasilkan dua kemungkinan hasil, "Sukses" dan "Gagal". Contohnya adalah pelemparan satu buah mata uang logam, dimana terdapat 2 kemungkinan hasil yang bisa diperoleh dari satu kali pelemparan, yaitu Angka dan Gambar. Misalkan munculnya Angka dianggap kejadian yang "Sukses" dimana peluang munculnya adalah p dan munculnya Gambar dianggap kejadian yang "Gagal" dimana peluang munculnya adalah $1-p$. Selanjutnya, variabel random Y terkait percobaan tersebut diberi nilai 1 dengan peluang p jika "Sukses" terjadi dan diberi nilai 0 jika "Gagal" terjadi dengan peluang $1-p$ (Walpole dkk, 2002). Dengan demikian, variabel random Y dikatakan berdistribusi Bernoulli dengan fungsi kepekatkan peluang sebagai berikut:

$$f(y, p) = \begin{cases} p^y(1-p)^{1-y}, & y = 0,1 \\ 0, & \text{lainnya.} \end{cases} \quad (2.1)$$

distribusi bernoulli mempunyai nilai *mean* dan varian sebagai berikut:

$$E(Y) = p$$

$$Var(Y) = p(1-p)$$

2.2 Model Regresi Probit

Menurut Greenberg (1980) analisis probit pertama kali dikemukakan oleh Chester Ittner Bliss pada tahun 1934 yang dalam penelitiannya mengenai pestisida untuk mengendalikan serangga yang hidup pada daun dan buah anggur. Dalam suatu artikel di jurnal Science tersebut, Bliss mengemukakan bahwa istilah probit dalam model regresi probit berasal dari kata *probability unit*, dengan kata lain model regresi probit merupakan suatu model regresi yang berkaitan dengan unit-unit probabilitas.

Model regresi probit adalah salah satu model regresi yang dapat digunakan untuk mengetahui pengaruh variabel prediktor terhadap variabel respon yang berbentuk kategori (kualitatif). Model probit juga disebut model normit karena model probit menggunakan pendekatan CDF distribusi normal dimana pendekatan CDF digunakan untuk mengatasi kelemahan pada *linear probability model*. Kelemahan tersebut adalah kemungkinan dari nilai Y_i keluar dari range variabel respon. Model probit menggunakan asumsi bahwa probabilitas kejadian sukses bergantung pada latent atau variabel yang tidak dapat diobservasi dan terdapat nilai kritisal (*threshold*) pada variabel yang tidak teramati.

Menurut Greene (2008) variabel respon kualitatif Y pada model regresi probit berasal dari variabel respon yang tidak teramati (laten) Y^* , sehingga pemodelan regresi probit diawali dengan melihat model sebagai berikut:

$$y_i^* = \boldsymbol{\beta}^T \mathbf{x}_i + \varepsilon_i \quad (2.2)$$

dimana y_i^* adalah variabel respon yang tidak teramati (laten), $\boldsymbol{\beta}$ adalah vektor koefisien parameter dengan $\boldsymbol{\beta} = [\beta_0 \beta_1 \beta_2 \dots \beta_p]^T$, \mathbf{x}_i adalah vektor variabel prediktor dengan $\mathbf{x}_i = [1 \ x_{1i} \ x_{2i} \dots \ x_{pi}]^T$ dan ε_i adalah error yang diasumsikan berdistribusi normal dengan rata-rata 0 dan varians σ^2 . Jadi fungsi kepadatan probabilitas (PDF) dari variabel y_i^* adalah sebagai berikut (Greene, 2008):

$$f(y_i^*) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i^* - \boldsymbol{\beta}^T \mathbf{x}_i)^2\right) \quad (2.3)$$

dengan y_i^* berdistribusi normal dengan mean $\boldsymbol{\beta}^T \mathbf{x}_i$ dan varians σ^2 .

Variabel y_i^* tidak bisa diamati, sehingga dalam pengkategorian variabel respon y_i menggunakan tingkat ambang batas tertentu (*threshold*), yang kemudian dapat diubah menjadi nol tanpa kehilangan unsur generalitasnya (Verbeek, 2004). Sehingga pengkategorian variabel y_i adalah sebagai berikut:

$$y_i = \begin{cases} 0, & \text{jika } y_i^* < 0 \\ 1, & \text{jika } y_i^* \geq 0. \end{cases}$$

Karena nilai variabel y_i bersifat biner yang hanya bergantung pada tanda y_i^* dan bukan nilai absolutnya, menyebabkan tidak adanya pengaruh oleh jumlah varian error sehingga tidak ada batasan untuk melakukan standarisasi varian error menjadi satu (Arbia, 2014). Oleh karena itu, persamaan dalam (2.2) dapat ditransformasi ke dalam bentuk normal standar dengan mean 0 dan varians 1, sehingga probabilitas untuk $y_i = 0$ atau $P(y_i = 0)$ dan $y_i = 1$ atau $P(y_i = 1)$ adalah:

$$\begin{aligned} P(y_i = 0) &= P(y_i^* < 0) \\ &= P(\boldsymbol{\beta}^T \mathbf{x}_i + \varepsilon_i < 0) \\ &= P(\varepsilon_i < -\boldsymbol{\beta}^T \mathbf{x}_i) \\ &= \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i) \end{aligned} \tag{2.4}$$

$$\begin{aligned} P(y_i = 1) &= P(y_i^* \geq 0) \\ &= 1 - P(y_i^* < 0) \\ &= 1 - P(\boldsymbol{\beta}^T \mathbf{x}_i + \varepsilon_i < 0) \\ &= 1 - \Pr(\varepsilon_i < -\boldsymbol{\beta}^T \mathbf{x}_i) \\ &= 1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i) \end{aligned} \tag{2.5}$$

dimana $\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)$ adalah fungsi distribusi kumulatif normal standar. Pada persamaan (2.5) ini disebut model probit biner.

Untuk menginterpretasikan model probit pada persamaan (2.5) dapat digunakan efek marginal (Greene, 2008). Efek marginal diperoleh dari turunan probabilitas tiap kategorinya. Efek marginal variabel \mathbf{x}_i dari model regresi probit di atas adalah sebagai berikut:

$$\frac{\partial P(y_i = 0)}{\partial \mathbf{x}_i} = \phi(-\boldsymbol{\beta}^T \mathbf{x}_i)(-\boldsymbol{\beta}) \tag{2.6}$$

$$\frac{\partial P(y_i = 1)}{\partial x_i} = 1 - \phi(-\boldsymbol{\beta}^T \mathbf{x}_i)(\boldsymbol{\beta}) \quad (2.7)$$

Persamaan (2.6) dan (2.7) menunjukkan besarnya pengaruh variabel x_i terhadap $P(y_i = 0)$ dan $P(y_i = 1)$.

Salah satu metode estimasi parameter pada regresi probit adalah metode *Maximum Likelihood Estimation* (MLE). Metode MLE adalah metode estimasi dengan cara memaksimalkan fungsi *likelihood*-nya. Tahapan estimasi parameter pada regresi probit univariat dengan menggunakan metode MLE adalah sebagai berikut:

- a. Tentukan n sampel secara random y_i , dengan $i = 1, 2, \dots, n$
- b. Tentukan fungsi *likelihood* dari variabel random Y dalam model. Misalkan dengan model regresi probit biner, dimana variabel random Y berdistribusi Bernoulli $(1, p)$ sehingga fungsi *likelihood* dari Y adalah sebagai berikut:

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n [1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]^{y_i} [\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]^{1-y_i} \quad (2.8)$$

- c. Lakukan transformasi \ln terhadap fungsi *likelihood* pada persamaan (2.8) sehingga diperoleh persamaan sebagai berikut:

$$\begin{aligned} \ln L(\boldsymbol{\beta}) &= \ln \left(\prod_{i=1}^n [1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]^{y_i} [\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]^{1-y_i} \right) \\ &= \sum_{i=1}^n (y_i \ln[1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)] + (1 - y_i) \ln[\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]) \end{aligned} \quad (2.9)$$

- d. Maksimumkan fungsi *likelihood* pada persamaan (2.9) dengan menurunkan fungsi *likelihood*-nya terhadap parameter $\boldsymbol{\beta}$, kemudian menyamakan dengan nol untuk mendapatkan estimasi parameter $\boldsymbol{\beta}$.

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \frac{\partial}{\partial \boldsymbol{\beta}} \left(\sum_{i=1}^n (y_i \ln[1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)] + (1 - y_i) \ln[\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]) \right) \quad (2.10)$$

dari persamaan (2.10) diperoleh turunan pertama dari fungsi *likelihood*-nya adalah:

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n \left(y_i \frac{\partial \ln[1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]}{\partial \boldsymbol{\beta}} + (1 - y_i) \frac{\partial \ln[\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)]}{\partial \boldsymbol{\beta}} \right) = 0$$

$$\begin{aligned}
&= \sum_{i=1}^n \left(\frac{y_i}{1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)} \frac{\partial \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)}{\partial \boldsymbol{\beta}} + \frac{1 - y_i}{\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)} \frac{\partial \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)}{\partial \boldsymbol{\beta}} \right) = 0 \\
&= \sum_{i=1}^n \frac{\partial \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)}{\partial \boldsymbol{\beta}} \left(\frac{y_i}{1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)} + \frac{1 - y_i}{\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)} \right) = 0 \quad (2.11)
\end{aligned}$$

Untuk menyelesaikan persamaan (4.1), Ratnasari (2012) mendeskripsikan beberapa konsep dasar tentang turunan vektor sebagai berikut:

Lemma 2.1

1) Jika diberikan vektor \mathbf{a} yang berukuran $p \times 1$ dan \mathbf{m} berukuran $p \times 1$, maka

$$\frac{\partial(\mathbf{a}^T \mathbf{m})}{\partial \mathbf{a}} = \mathbf{m}.$$

2) Jika $\Phi(\mathbf{a}^T \mathbf{m})$ adalah distribusi kumulatif normal, maka $\frac{\partial \Phi(\mathbf{a}^T \mathbf{m})}{\partial \mathbf{m}} = \mathbf{a} \phi(\mathbf{a}^T \mathbf{m})$, dimana $\phi(\mathbf{a}^T \mathbf{m})$ adalah distribusi normal standar.

3) Jika $\phi(\mathbf{a}^T \mathbf{m})$ adalah distribusi normal standar, maka $\frac{\partial \phi(\mathbf{a}^T \mathbf{m})}{\partial \mathbf{m}} = -\mathbf{a}(\mathbf{a}^T \mathbf{m}) \phi(\mathbf{a}^T \mathbf{m})$.

Berdasarkan Lemma 2.1, maka:

$$\frac{\partial \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)}{\partial \boldsymbol{\beta}} = -\mathbf{x}_i \phi(-\boldsymbol{\beta}^T \mathbf{x}_i)$$

Sehingga turunan *ln-likelihood* terhadap $\boldsymbol{\beta}$ adalah

$$\frac{\partial \ln L(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^n -\mathbf{x}_i \phi(-\boldsymbol{\beta}^T \mathbf{x}_i) \left(\frac{y_i}{1 - \Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)} + \frac{1 - y_i}{\Phi(-\boldsymbol{\beta}^T \mathbf{x}_i)} \right) = 0. \quad (2.12)$$

Persamaan (2.11) menunjukkan bentuk yang tidak *closed form*, sehingga diperlukan suatu metode optimasi numerik untuk penyelesaiannya.

2.3 Uji Multikolinieritas

Salah satu syarat yang harus dipenuhi dalam regresi dengan beberapa variabel prediktor adalah tidak adanya korelasi antara satu variabel prediktor dengan variabel prediktor lainnya. Multikolinieritas adalah suatu kondisi dimana variabel-variabel prediktor berkorelasi tinggi. Menurut Gujarati (2004) salah satu cara mengidentifikasi adanya multikolinieritas yaitu dengan menggunakan *Variance Inflation Factors* (VIF) yang dinyatakan sebagai berikut:

$$VIF_j = \frac{1}{(1 - R_j^2)} \quad (2.13)$$

R_j^2 adalah koefisien determinasi antara X_j dengan variabel prediktor lainnya. Nilai VIF yang lebih besar dari 10 menunjukkan adanya kolinearitas antar variabel prediktor.

2.4 Model Regresi Spasial

Model ini dikembangkan oleh Anselin (1988) dengan menggunakan data spasial *cross section*. Model dari *General Spatial Model* untuk variabel respon yang berbentuk kategori ditunjukkan dengan:

$$\begin{aligned} \mathbf{y}^* &= \rho \mathbf{W} \mathbf{y}^* + \mathbf{X} \boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\varepsilon} \end{aligned} \quad (2.14)$$

dimana $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I}_n)$, \mathbf{y}^* adalah vektor variabel laten yang berukuran $(n \times 1)$, \mathbf{X} adalah matriks variabel prediktor yang berukuran $n \times (k+1)$, $\boldsymbol{\beta}$ adalah vektor koefisien parameter regresi berukuran $(k+1) \times 1$ dan ρ adalah koefisien spasial lag dari variabel respon, sedangkan λ merupakan koefisien spasial lag pada *error* yang bernilai $|\lambda| < 1$. Adapun \mathbf{W} adalah matriks penimbang spasial yang berukuran $(n \times n)$ yang elemen diagonalnya bernilai nol. Matriks penimbang ini biasanya berisi hubungan *contiguity* matriks atau juga fungsi jarak dari suatu daerah/region. \mathbf{u} adalah vektor *error* regresi yang diasumsikan mempunyai efek region random dan juga *error* yang terautokorelasi secara spasial dengan ukuran $(n \times 1)$.

Ada beberapa turunan model yang bisa dibentuk dari *General Spatial Model* pada persamaan (2.14) yaitu:

- a. Apabila $\rho = 0$ dan $\lambda = 0$ maka persamaan menjadi:

$$\mathbf{y}^* = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}. \quad (2.15)$$

Persamaan (2.15) ini disebut model regresi klasik atau lazim dikenal sebagai model regresi *Ordinary Least Square* (OLS), yaitu regresi yang tidak mempunyai efek spasial.

- b. Apabila $\rho \neq 0$ dan $\lambda = 0$ maka persamaan menjadi:

$$\mathbf{y}^* = \rho \mathbf{W} \mathbf{y}^* + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.16)$$

Persamaan (2.16) disebut sebagai regresi *Spatial Lag Model* (SLM). LeSage

(1999) mengistilahkan model ini dengan *Spatial Autoregressive Models* (SAR).

- c. Apabila $\rho = 0$ dan $\lambda \neq 0$ maka persamaan menjadi:

$$\mathbf{y}^* = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \quad (2.17)$$

Persamaan (2.17) disebut juga regresi *Spatial Error Model* (SEM).

- d. Apabila $\rho \neq 0$ dan $\lambda \neq 0$ maka persamaan menjadi:

$$\mathbf{y}^* = \rho\mathbf{W}\mathbf{y}^* + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \text{ dan } \mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \boldsymbol{\varepsilon} \quad (2.18)$$

Persamaan (2.18) disebut *General Spatial Model* (GSM), ada juga Anselin (1988) menamainya sebagai model *Spatial Autoregressive Moving Average* (SARMA).

2.4.1 Matriks Pembobot Spasial

Informasi lokasi dalam analisis spasial dapat menggunakan dua sumber, yaitu *contiguity* dan *distance*. *Contiguity* menggambarkan lokasi relatif dari satu unit spasial ke wilayah lain dalam suatu tempat. Hubungan kebertetanggaan dari unit spasial biasanya dibangun menggunakan peta. Sedangkan *distance* menggunakan garis lintang dan bujur sebagai sumber informasi, sehingga dengan informasi ini bisa diukur jarak anantara satu titik lokasi dengan titik lokasi lainnya di suatu tempat. Harapannya lokasi yang jaraknya dekat mempunyai karakteristik yang mirip.

Untuk memperoleh matriks pembobot spasial (\mathbf{W}), salah satu cara yang bisa digunakan yaitu dengan menggunakan informasi dari ketetanggaan (*neighborhood*), atau kedekatan antara satu region dengan region yang lain. Ada beberapa metode untuk mendefinisikan hubungan persinggungan (*contiguity*) antar wilayah tersebut. Menurut LeSage (1999) metode *contiguity* terdiri dari:

- a. *Linear Contiguity* (Persinggungan tepi);

definisikan $w_{ij} = 1$ untuk region yang berada di tepi kiri maupun kanan region yang menjadi perhatian, $w_{ij} = 0$ untuk region lainnya. Dari Gambar 2.1 terlihat $w_{53} = 1$ (Hubungan *contiguity* antara region 5 dan region 3 adalah 1, sedangkan yang lain = 0).

- b. *Rook Contiguity* (Persinggungan sisi);

definisikan $w_{ij} = 1$ untuk region yang bersisian dengan region yang menjadi perhatian, $w_{ij} = 0$ untuk region lainnya. Bila region 3 menjadi perhatian, maka $w_{34} = 1$, $w_{35} = 1$ dan yang lain sama dengan nol.

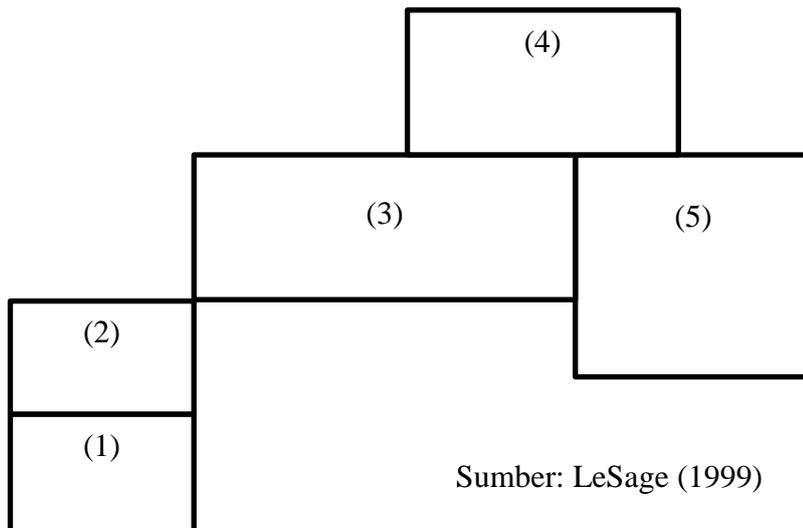
c. *Bhisop Contiguity* (Persinggungan sudut);

definisikan $w_{ij} = 1$ untuk region yang titik sudutnya bertemu dengan sudut region yang menjadi perhatian, $w_{ij} = 0$ untuk region lainnya. Untuk region 2 didapatkan hubungan $w_{23} = 1$ dan yang lain sama dengan nol.

d. *Queen Contiguity* (persinggungan sisi-sudut);

definisikan $w_{ij} = 1$ untuk entity yang bersisian atau titik sudutnya bertemu dengan region yang menjadi perhatian, $w_{ij} = 0$ untuk region lainnya. Untuk region 3, didapatkan $w_{32} = 1$, $w_{34} = 1$, $w_{35} = 1$ dan yang lain sama dengan nol. Sebagai ilustrasi, Gambar 2.1 yang merupakan contoh lima region yang tampak pada suatu peta.

Winarno (2009) pada penelitiannya tentang faktor-faktor yang mempengaruhi angka kematian bayi di Jawa Timur dengan pendekatan model regresi spasial, menjelaskan bahwa pembentukan matriks pembobot spasial yang lebih cocok digunakan pada daerah yang tidak simetris adalah metode *queen contiguity*. Terkait dengan kondisi geografis wilayah kabupaten/kota di Pulau Papua yang juga berbentuk tidak simetris, maka pada penelitian ini menggunakan metode *queen contiguity* sebagai matriks pembobot spasialnya.



Gambar 2.1 Ilustrasi *Contiguity* (persinggungan)

Dari Gambar 2.1, maka diperoleh susunan matriks pembobot spasial *queen contiguity* berukuran 5×5, sebagai berikut:

$$\begin{matrix} & 1 & 2 & 3 & 4 & 5 \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \end{matrix}$$

dimana baris dan kolom menyatakan region yang ada pada peta. Karena matriks pembobot spasial merupakan matriks simetris, dan dengan kaidah bahwa diagonal utama selalu nol. Pada praktiknya pada matriks pembobot spasial tersebut perlu dilakukan standarisasi. Proses standarisasi dilakukan agar diperoleh jumlah baris yang *unity*, yaitu jumlah barisnya sama dengan satu (Lee dan Wong, 2001), sehingga matriksnya menjadi:

$$W_{\text{queen}} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 \\ 0 & 0,3 & 0 & 0,3 & 0,3 \\ 0 & 0 & 0,5 & 0 & 0,5 \\ 0 & 0 & 0,5 & 0,5 & 0 \end{bmatrix}$$

2.4.2 Uji Dependensi Spasial

Pada saat melakukan analisis data cross section, efek autokorelasi spasial atau dependensi spasial dapat muncul dalam dua bentuk. Bentuk pertama adalah *spatial lag*, terjadi ketika informasi di suatu wilayah observasi memiliki korelasi dengan wilayah observasi lainnya yang bersinggungan wilayah. Kedua adalah *spatial error*, yang menunjukkan bahwa korelasi antar wilayah tidak hanya terjadi pada informasi variabel respon, tetapi juga pada variabel prediktor (Anselin, 1988).

Anselin (1988) menyatakan bahwa salah satu metode untuk mengetahui ada tidaknya dependensi lag pada model dapat diuji dengan menggunakan *Lagrange Multiplier Test* dengan pengujian hipotesis sebagai berikut:

$$H_0 : \rho = 0 \text{ (tidak ada ketergantungan lag spasial)}$$

$$H_1 : \rho \neq 0 \text{ (ada ketergantungan lag spasial)}$$

Statistik uji yang digunakan adalah pada persamaan (2.19).

$$LM = \frac{(\mathbf{e}^T \mathbf{W} \mathbf{y} / \hat{\sigma}^2)^2}{\mathbf{D}} \quad (2.19)$$

dimana $\hat{\sigma}^2 = \mathbf{e}^T \mathbf{e} / n$

$$\mathbf{e} = \mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}$$

$$\mathbf{D} = \left[(\mathbf{W} \mathbf{X} \hat{\boldsymbol{\beta}})^T (\mathbf{I} - \mathbf{X} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) (\mathbf{W} \mathbf{X} \hat{\boldsymbol{\beta}}) \right] / \hat{\sigma}^2 + \mathbf{T}$$

$$\mathbf{T} = \text{tr}[(\mathbf{W}^T + \mathbf{W}) \mathbf{W}]$$

Statistik uji ini mengikuti distribusi asimtotik $\chi^2_{(\alpha,1)}$, sehingga H_0 ditolak apabila $LM > \chi^2_{(\alpha,1)}$ yang menunjukkan bahwa pada model terdapat efek spasial lag.

2.5 Model Regresi Probit Spasial

Suatu model probit dengan dependensi spasial pertama kali diteliti oleh McMillen pada tahun 1992 yaitu dengan suatu algoritma EM diperkenalkan untuk menghasilkan suatu dugaan yang konsisten (maksimum likelihood) untuk model tersebut. Model regresi probit spasial merupakan gabungan model regresi probit dan model spasial. Bentuk model regresi probit spasial adalah sebagai berikut :

$$\mathbf{y}^* = \rho \mathbf{W} \mathbf{y}^* + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim N(0, \sigma^2 \mathbf{I}_n) \quad (2.20)$$

dimana \mathbf{y}^* adalah vektor variabel laten berukuran $(n \times 1)$, \mathbf{X} adalah matriks variabel prediktor yang berukuran $n \times (k+1)$, $\boldsymbol{\beta}$ adalah vektor koefisien parameter regresi berukuran $(k+1) \times 1$, \mathbf{W} yaitu matriks pembobot berukuran $(n \times n)$, dan ρ adalah parameter spasial.

Persamaan (2.20) dapat ditulis dalam bentuk direduksi adalah sebagai berikut:

$$\begin{aligned} \mathbf{y}^* &= (\mathbf{I} - \rho \mathbf{W})^{-1} (\mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}) \\ &= (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} + (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon} \\ &= \boldsymbol{\Gamma} \mathbf{X} \boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim \text{MVN}(0, \boldsymbol{\Omega}) \end{aligned} \quad (2.21)$$

dengan $\mathbf{u} = (\mathbf{I} - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon}$, dimana \mathbf{u} merupakan vektor yang berukuran $(n \times 1)$ dan $\boldsymbol{\Gamma} = (\mathbf{I} - \rho \mathbf{W})^{-1}$, dimana $\boldsymbol{\Gamma}$ merupakan matriks berukuran $(n \times n)$.

Variabel laten \mathbf{y}^* mempunyai kategori biner yang didefinisikan sebagai variabel \mathbf{y} melalui persamaan berikut:

$$y_i = \begin{cases} 0, & \text{jika } y_i^* < 0 \\ 1, & \text{jika } y_i^* \geq 0 \end{cases} \quad (2.22)$$

maka probabilitas untuk $P(y_i = 1)$ dan $P(y_i = 0)$ adalah sebagai berikut:

$$\begin{aligned}
P(y_i = 0|X_i) &= P(y_i^* < 0) \\
&= P([\mathbf{\Gamma X \beta}]_i + u_i < 0) \\
&= P(u_i < -[\mathbf{\Gamma X \beta}]_i) \\
&= \Phi(-[\mathbf{\Gamma X \beta}]_i),
\end{aligned} \tag{2.23}$$

$$\begin{aligned}
P(y_i = 1|X_i) &= P(y_i^* \geq 0) \\
&= P([\mathbf{\Gamma X \beta}]_i + u_i \geq 0) \\
&= P(u_i \geq -[\mathbf{\Gamma X \beta}]_i) \\
&= 1 - P(u_i < -[\mathbf{\Gamma X \beta}]_i) \\
&= 1 - \Phi(-[\mathbf{\Gamma X \beta}]_i).
\end{aligned} \tag{2.24}$$

Sehingga didapatkan tingkat ambang batas untuk $y_i = 0$ dan $y_i = 1$ dan persamaan (2.22) dapat disajikan dalam bentuk:

$$y_i = \begin{cases} 0, & \text{jika } u_i \geq -[\mathbf{\Gamma X \beta}]_i \\ 1, & \text{jika } u_i < -[\mathbf{\Gamma X \beta}]_i. \end{cases} \tag{2.25}$$

Karena variabel \mathbf{u} diasumsikan berdistribusi multivariat normal dengan mean 0 dan varians $\mathbf{\Omega}$, maka probabilitas untuk $P(y_i = 1)$ dan $P(y_i = 0)$ adalah sebagai berikut:

$$P(y_i = 0|X) = \Phi\left(\frac{-[\mathbf{\Gamma X \beta}]_i}{\mathbf{\Omega}_{ii}}\right), \tag{2.26}$$

$$P(y_i = 1|X) = 1 - \Phi\left(\frac{-[\mathbf{\Gamma X \beta}]_i}{\mathbf{\Omega}_{ii}}\right), \tag{2.27}$$

dimana $\mathbf{\Omega}_{ii}$ merupakan elemen diagonal utama dari $\mathbf{\Omega}$ yang berbentuk sebagai berikut:

$$\mathbf{\Omega} = [(\mathbf{I} - \rho\mathbf{W})^T(\mathbf{I} - \rho\mathbf{W})]^{-1}. \tag{2.28}$$

Untuk menghitung efek marginal \mathbf{X} dalam estimasi peluang dari hasil $y_i = 1$, adalah sebagai berikut:

$$\begin{aligned}
\frac{\partial P(y_i = 1|X)}{\partial X_i} &= \frac{\partial}{\partial X_i} \left[1 - \Phi\left(\frac{-[\mathbf{\Gamma X \beta}]_i}{\mathbf{\Omega}_{ii}}\right) \right] \\
&= \frac{\partial}{\partial X_i} \Phi\left(\frac{[\mathbf{\Gamma X \beta}]_i}{\mathbf{\Omega}_{ii}}\right) \\
&= \phi\left(\frac{[\mathbf{\Gamma X \beta}]_i}{\mathbf{\Omega}_{ii}}\right) \frac{[\mathbf{\Gamma \beta}]_i}{\mathbf{\Omega}_{ii}}
\end{aligned} \tag{2.29}$$

Estimasi parameter model regresi probit spasial dapat dilakukan dengan menggunakan metode *Maximum Likelihood Estimation* (MLE). Estimasi parameter diperoleh dengan memaksimalkan fungsi *likelihood* dari y_i , yang merupakan variabel random berdistribusi Bernoulli (1, p) yaitu:

$$L(\boldsymbol{\beta}, \rho) = \prod_{i=1}^n \left(\left[1 - \Phi \left(\frac{-[\mathbf{r}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \right]^{y_i} \right) \left(\left[\Phi \left(\frac{-[\mathbf{r}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \right]^{1-y_i} \right). \quad (2.30)$$

Kemudian dilakukan tranformasi ln fungsi *likelihood*, yaitu menjadi:

$$\ln[L(\boldsymbol{\beta}, \rho)] = \sum_{i=1}^n y_i \ln \left[1 - \Phi \left(\frac{-[\mathbf{r}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \right] + \sum_{i=1}^n (1 - y_i) \ln \Phi \left(\frac{-[\mathbf{r}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \quad (2.31)$$

Fungsi ln *likelihood* tersebut akan dimaksimalkan untuk mendapatkan penaksir parameter $\boldsymbol{\beta}$ dan ρ . Persamaan (2.31) pada akhirnya akan menunjukkan bentuk yang tidak *closed form*, untuk menyelesaikan permasalahan tersebut dapat dioptimasi dengan metode *recursive importance sampling* (RIS).

2.6 Recursive Importance Sampling (RIS)

Beron dan Vijverberg (2004) menyarankan penggunaan estimator RIS simulator untuk memperkirakan probabilitas yang sulit dihitung secara analitis dan untuk mengevaluasi secara langsung integral n-dimensi pada model dengan spatial lag dan model dengan spatial error, yaitu dengan menggunakan dekomposisi dari matriks varians kovarians n-dimensi yang menghasilkan matriks segitiga atas (*upper triangular matrix*). Metode ini dapat dijalankan dengan memanfaatkan fakta bahwa observasi yang terakhir saat ini independen dari observasi lain, observasi terakhir kedua hanya bergantung pada yang terakhir, dan seterusnya, sehingga memungkinkan algoritma *recursive sampling*. Simulator RIS ini merupakan pengembangan dari simulator Geweke-Hajivassiliou-Keane (GHK).

Untuk mengestimasi model regresi probit spatial dengan menggunakan metode RIS, kita akan mengemukakan kembali persamaan (2.25) dengan hanya menggunakan ambang batas atasnya saja. Maka didefinisikan:

$$v_i = (1 - 2y_i)u_i$$

untuk $y_i = 0$, maka $v_i = u_i$ sehingga $v_i < -[\mathbf{r}\mathbf{X}\boldsymbol{\beta}]_i$

$$y_i = 1, \text{ maka } v_i = -u_i \text{ sehingga } v_i < [\mathbf{r}\mathbf{X}\boldsymbol{\beta}]_i.$$

Dikarenakan terjadi kesetaraan $u_i = -[\Gamma\mathbf{X}\boldsymbol{\beta}]_i$ pada saat probabilitas 0, maka ketidaksetaraan pada persamaan (2.25) dapat disajikan kembali secara lebih singkat sebagai:

$$v_i < -(1 - 2y_i)[\Gamma\mathbf{X}\boldsymbol{\beta}]_i$$

Kemudian definisikan \mathbf{Z} sebagai matriks berukuran $(n \times n)$ dengan $Z_{ii} = (1 - 2y_i)$ dan \mathbf{Z} merupakan matrik diagonal yang menunjukkan $\mathbf{Z}\mathbf{Z}^T = \mathbf{I}_n$, sebagai matriks identitas berukuran $(n \times n)$. Dengan demikian, kondisi pada v_i dapat dinyatakan dalam bentuk:

$$v_i < -[\mathbf{Z}\Gamma\mathbf{X}\boldsymbol{\beta}]_i \quad (2.32)$$

sehingga fungsi log-likelihood pada persamaan (2.31) dapat ditulis kembali sebagai:

$$\ln[L(\boldsymbol{\beta}, \rho)] = \sum_{i=1}^n \ln \Phi(-[\mathbf{Z}\Gamma\mathbf{X}\boldsymbol{\beta}]_i) \quad (2.33)$$

Untuk mengestimasi parameter pada persamaan (2.33), Beron dan Vijverberg (2004) menggunakan simulator yang dapat memperkirakan n dimensi normal probabilitas dengan presisi yang baik, yaitu RIS simulator, dimana poin utama dari RIS simulator adalah melakukan perhitungan single n dimensi kumulatif normal pada persamaan (2.32):

$$\begin{aligned} p &= \prod_{i=1}^n P(v_i < -[\mathbf{Z}\Gamma\mathbf{X}\boldsymbol{\beta}]_i) \\ &= \prod_{i=1}^n P(v_i < c_i) \end{aligned} \quad (2.34)$$

Distribusi dari \mathbf{v} adalah $N(0, \boldsymbol{\Omega})$ dan estimator RIS pada probit spasial menunjukkan fakta bahwa $\boldsymbol{\Omega}$ sebagai varian kovarian matrik merupakan definit positif, sehingga menurut dekomposisi Cholesky $\boldsymbol{\Omega}^{-1} = \mathbf{A}^T\mathbf{A}$, dengan \mathbf{A} menjadi matrik segitiga atas dan $\boldsymbol{\eta} = \mathbf{A}\mathbf{v}$ memberikan n variabel independen standar normal, dimana $\boldsymbol{\eta}$ merupakan vektor berukuran $(n \times 1)$. Definisikan $\mathbf{B} = \mathbf{A}^{-1}$, dimana \mathbf{B} merupakan matrik segitiga atas dengan $b_{j,j} > 0$ untuk semua j . Kemudian disubsitusikan $\mathbf{v} = \mathbf{A}^{-1}\boldsymbol{\eta} \equiv \mathbf{B}\boldsymbol{\eta}$ ke dalam batas atas kumulatif standar normal pada persamaan (2.32) sehingga:

$$\mathbf{B}\boldsymbol{\eta} < \mathbf{c} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & \cdots & b_{1,n} \\ 0 & b_{2,2} & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{n-1,n-1} & b_{n-1,n} \\ 0 & \cdots & 0 & 0 & b_{n,n} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \vdots \\ \eta_{n-1} \\ \eta_n \end{bmatrix} < \begin{bmatrix} c_1 \\ \vdots \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix} \quad (2.35)$$

Elemen dari vektor $\boldsymbol{\eta}$ bersifat independen, sehingga probabilitas dalam persamaan (2.35) dapat dihitung dengan mengevaluasi fungsi distribusi kumulatif normal pada batas atas, yaitu ditentukan secara rekursif dimulai dengan pengamatan terakhir, kemudian mengalikan probabilitas ini (Franzese dan Hays, 2007). Untuk menentukan probabilitas pada batas atas ini, mulailah dengan memecahkan ketidaksetaraan pada persamaan (2.35) untuk vektor $\boldsymbol{\eta}$:

$$\begin{aligned} P\left(\sum_{i=1}^n b_{1,i}\eta_i < c_1\right) &= P(\eta_1 < b_{1,1}^{-1}\left(c_1 - \sum_{i=2}^n b_{1,i}\eta_i\right)) \equiv p_1 \\ \vdots & \qquad \qquad \qquad \vdots & \qquad \qquad \qquad \vdots \\ P(b_{n-1,n-1}\eta_{n-1} + b_{n-1,n}\eta_n < c_{n-1}) &= P(\eta_{n-1} < b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\eta_n)) \equiv p_{n-1} \\ P(b_{n,n}\eta_n < c_n) &= P(\eta_n < b_{n,n}^{-1}c_n) \equiv p_n \end{aligned} \quad (2.36)$$

Selanjutnya, proses rekursif dimulai dengan melakukan penghitungan batas atas kumulatif standar normal p_n , yaitu menghitung batas atas pada perpotongan distribusi normal dari observasi ke- n , yaitu $b_{n,n}^{-1}c_n$. Kemudian ambil nilai dari perpotongan distribusi normal standar pada $b_{n,n}^{-1}c_n$, yang dinotasikan dengan $\tilde{\eta}_n$, lalu digunakan untuk menghitung batas atas kumulatif standar normal p_{n-1} , yaitu dengan memasukkan $\tilde{\eta}_n$ pada perpotongan distribusi normal untuk observasi ke- $(n-1)$, sehingga menjadi $b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\tilde{\eta}_n)$. Kemudian gunakan dua seri tadi yaitu $\tilde{\eta}_n$ dan $\tilde{\eta}_{n-1}$ untuk menghitung batas atas kumulatif standar normal p_{n-2} yaitu pada observasi ke- $(n-2)$ dan seterusnya hingga n observasi. Sehingga, proses rekursif untuk menghitung batas atas adalah:

$$\begin{aligned} \eta_n &< b_{n,n}^{-1}c_n \equiv \eta_{n,0} \\ \eta_{n-1} &< b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\tilde{\eta}_n) \equiv \eta_{n-1,0} \\ \eta_{n-2} &< b_{n-2,n-1}^{-1}(c_{n-2} - b_{n-1,n-1}\tilde{\eta}_{n-1} - b_{n-2,n}\tilde{\eta}_n) \equiv \eta_{n-2,0} \\ &\vdots \\ \eta_j &< b_{j,j}^{-1}\left(c_j - \sum_{i=j+1}^n b_{j,i}\tilde{\eta}_i\right) \equiv \eta_{j,0} \end{aligned} \quad (2.37)$$

Probabilitas dari observasi dengan pemberian sampel 1 dan 0, dapat ditemukan dengan mengevaluasi fungsi kumulatif distribusi normal univariat pada

masing-masing batas tersebut yaitu p_i , lalu mengalikan probabilitas tersebut, sehingga:

$$\prod_{j=1}^n p_j = \prod_{j=1}^n \Phi(\eta_{j,0}). \quad (2.38)$$

Pengulangan dari seluruh proses sebanyak R kali dan menghitung nilai rata-rata yang dihasilkan dapat memberikan nilai estimasi RIS pada probabilitas gabungan, sehingga didapatkan persamaan sebagai berikut:

$$\hat{p} = \frac{1}{R} \sum_{r=1}^R \left[\prod_{j=1}^n \Phi(\eta_{j,0,r}) \right] \quad (2.39)$$

Kemudian untuk mendapatkan estimasi parameter, maka perlu dilakukan optimasi secara konvensional terhadap $\ln \hat{p}$.

2.7 Pengujian Parameter

Pengujian parameter model regresi probit spasial dilakukan secara serentak dan parsial. Pengujian secara serentak adalah menguji secara simultan apakah terdapat parameter dari variabel prediktor yang berpengaruh secara signifikan terhadap variabel respon, sedangkan pengujian secara parsial adalah menguji apakah masing-masing parameter berpengaruh signifikan terhadap variabel respon.

a. Pengujian parameter model regresi probit spasial secara serentak

Hipotesis dalam pengujian parameter secara serentak adalah sebagai berikut:

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

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

Statistik uji yang digunakan dalam pengujian parameter model regresi probit spasial secara serentak menggunakan *likelihood ratio* (G^2) dengan formula sebagai berikut:

$$G^2 = -2 \ln \left[\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right] = 2 \ln [L(\hat{\Omega}) - L(\hat{\omega})] \quad (2.40)$$

dimana

Ω = himpunan parameter di bawah populasi

ω = himpunan parameter di bawah H_0

$$L(\hat{\Omega}) = \prod_{i=1}^n \left(\left[1 - \Phi \left(\frac{-[\Gamma X \hat{\beta}]_i}{\Omega_{ii}} \right) \right]^{y_i} \right) \left(\left[\Phi \left(\frac{-[\Gamma X \hat{\beta}]_i}{\Omega_{ii}} \right) \right]^{1-y_i} \right)$$

$$L(\hat{\omega}) = \prod_{i=1}^n \left(\left[1 - \Phi \left(\frac{-\Gamma_i \beta_0}{\Omega_{ii}} \right) \right]^{y_i} \right) \left(\left[\Phi \left(\frac{-\Gamma_i \beta_0}{\Omega_{ii}} \right) \right]^{1-y_i} \right)$$

Tolak H_0 jika nilai $G^2 > \chi^2_{(\alpha, df)}$, dengan derajat bebas (df) yaitu banyaknya parameter dibawah populasi dikurangi dengan banyaknya parameter di bawah H_0 .

- b. Pengujian parameter model regresi probit spasial secara parsial

Hipotesis dalam pengujian parameter secara parsial adalah sebagai berikut:

$$H_0 : \beta_j = 0$$

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

Statistik uji yang digunakan dalam pengujian parameter model regresi probit univariat secara parsial adalah sebagai berikut:

$$Z = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (2.41)$$

dimana $\hat{\beta}_j$ merupakan penaksir parameter β_j dan $SE(\hat{\beta}_j)$ merupakan $\sqrt{\widehat{\text{var}}(\hat{\beta}_j)}$,

H_0 ditolak jika $Z_{hitung} < -Z_{\alpha/2}$ atau $Z_{hitung} > Z_{\alpha/2}$, dimana nilai $Z_{\alpha/2}$ dapat diperoleh dari tabel normal standar.

2.8 Kriteria Keباikan Model

Salah satu ukuran untuk pemilihan model terbaik yang dapat digunakan pada pemodelan statistik yang melibatkan variabel respon kategorik adalah menggunakan *confusion matrix* atau ketepatan klasifikasi (Ratnasari, 2012). *Confusion matrix* merupakan evaluasi yang melihat probabilitas kesalahan klasifikasi yang dilakukan oleh suatu fungsi klasifikasi. Nilai ketepatan klasifikasi diperoleh dengan membandingkan nilai prediksi yang benar dari model dengan nilai observasi. Untuk kasus biner (klasifikasi 2 kategori), maka ilustrasi tabel *confusion matrix* dapat dilihat pada Tabel 2.1.

Tabel 2.1 Ilustrasi Tabel *Confusion Matrix*

Hasil Observasi	Hasil Prediksi	
	<i>Positive (P)</i>	<i>Negative (N)</i>
(1)	(2)	(3)
<i>Positive (P)</i>	<i>True Positive (TP)</i>	<i>False Negative (FN)</i>
<i>Negative (N)</i>	<i>False Positive (FP)</i>	<i>True Negative (TN)</i>

True positive (TP) merupakan kasus dimana data hasil obeservasi adalah berkategori positif (*P*) dan diprediksi oleh model sebagai *P*. *False positive (FP)* merupakan kasus dimana data hasil obeservasi adalah berkategori negatif (*N*) dan diprediksi oleh model sebagai *P*. *True negative (TN)* merupakan kasus dimana data hasil observasi adalah berkategori *N* dan diprediksi oleh model sebagai *N*. *False negative (FN)* merupakan kasus dimana data aktual adalah berkategori *P* dan diprediksi oleh model sebagai *N*. Tabel *confusion matrix* digunakan sebagai dasar berbagai penghitungan skor kebaikan model klasifikasi (Fawcett, 2006). Beberapa diantaranya adalah sebagai berikut:

❖ *Accuracy* (tingkat akurasi total)

Ukuran ini menyatakan tingkat keakuratan model secara total (semua kategori). Formula untuk menghitung tingkat akurasi total (ACC) adalah sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2.42)$$

❖ *Sensitivity*

Ukuran ini menyatakan seberapa banyak model dalam mendeteksi kategori (*true positive*). Ukuran ini juga dikenal sebagai *recall*, *true positive rate* (TPR), *hit rate*. Formula untuk menghitungnya adalah sebagai berikut:

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (2.43)$$

❖ *Specivisity*

Ukuran ini menyatakan seberapa banyak model dalam mendeteksi suatu kategori (*true negative*). Ukuran ini juga dikenal sebagai *true negative rate* (TNR). Formula untuk menghitungnya adalah sebagai berikut:

$$Specivisity = \frac{TN}{TN + FP} \times 100\% \quad (2.44)$$

Nilai *accuracy*, *sensivity* dan *specivisity* yang dihitung dengan formula diatas, menunjukkan bahwa semakin tinggi nilai yang dihasilkan, maka model semakin baik.

2.9 Indeks Kesehatan

Indeks kesehatan merupakan cerminan dari derajat kesehatan masyarakat di suatu wilayah. Indeks kesehatan disusun berdasarkan Angka Harapan Hidup saat lahir (AHH). AHH merupakan rata-rata perkiraan banyak tahun yang dapat ditempuh oleh seseorang selama hidup. AHH juga merupakan alat untuk mengevaluasi kinerja pemerintah dalam meningkatkan kesejahteraan penduduk pada umumnya, dan meningkatkan derajat kesehatan pada khususnya. AHH yang rendah di suatu daerah harus diikuti dengan program pembangunan kesehatan, dan program sosial lainnya termasuk kesehatan lingkungan, kecukupan gizi dan kalori termasuk program pemberantasan kemiskinan. Penghitungan AHH melalui pendekatan tak langsung (*indirect estimation*). Jenis data yang digunakan adalah Anak Lahir Hidup (ALH) dan Anak Masih Hidup (AMH). AHH dihitung berdasarkan ALH dan AMH dengan metode *Trussel* dengan model *West*, sesuai dengan histori kependudukan dan kondisi Indonesia dan negara-negara Asia Tenggara umumnya (BPS, 2015).

Cara penghitungan indeks kesehatan adalah sebagai berikut:

$$I_{Kesehatan} = \frac{AHH - AHH_{min}}{AHH_{maks} - AHH_{min}}$$

dengan nilai maksimum dan nilai minimum harapan hidup sesuai standar UNDP, yaitu angka tertinggi sebagai batas atas untuk penghitungan indeks dipakai 85 tahun dan terendah adalah 20 tahun.

2.10 Kajian Teori dan Penelitian Terdahulu

WHO (2017) menyatakan bahwa terdapat beberapa faktor (determinan) yang menentukan tingkat kesehatan, yaitu: status sosial, pendidikan, lingkungan, jaringan pendukung sosial, keturunan, dan pelayanan kesehatan. Semakin tinggi status sosial akan semakin baik tingkat kesehatannya. Semakin tinggi tingkat pendidikan juga berhubungan dengan tingkat kesehatan yang semakin baik. Keadaan lingkungan yaitu kebersihan air, kesehatan tempat kerja, kebersihan rumah berkontribusi terhadap tingkat kesehatan. Jaringan pendukung sosial seperti dukungan dari keluarga dan komunitas juga berhubungan dengan kesehatan yang lebih baik. Akses terhadap pelayanan kesehatan juga mempengaruhi tingkat kesehatan.

Faktor pendidikan berpengaruh pada tingkat kesehatan dan pendapatan seseorang. Pendidikan meningkatkan pengetahuan, kreativitas, dan imajinasi. Sebagai nilai tambah, pendidikan juga akan memperluas pilihan-pilihan lain. Manusia yang berpendidikan akan lebih memperhatikan tingkat kesehatan agar dapat hidup lebih lama. Tidak hanya itu, manusia yang berpendidikan juga akan berpeluang besar mendapatkan pekerjaan dan pendapatan yang lebih layak (BPS, 2015).

Maully (2014) meneliti faktor-faktor yang mempengaruhi indeks kesehatan kabupaten/kota di Jawa Timur dengan menggunakan regresi logistik. Faktor yang berpengaruh signifikan terhadap indeks kesehatan di kabupaten dan kota provinsi Jawa Timur adalah persentase pertolongan pertama kelahiran pada ibu dan persentase bayi diberi imunisasi. Satomo (2017) memodelkan indeks kesehatan dan indeks pengeluaran kabupaten/kota di Indonesia dengan pendekatan model probit bivariat. Dari lima variabel prediktor yang digunakan terdapat empat variabel yang signifikan berpengaruh, yaitu rata-rata lama sekolah, rasio tenaga kesehatan per 10.000 penduduk, rasio fasilitas pelayanan kesehatan per 10.000 penduduk dan tingkat pengangguran terbuka.

Sugiantari (2013) memodelkan Angka Harapan Hidup dengan menggunakan metode regresi semiparametrik spline. Penelitian ini menggunakan 6 faktor yang diduga mempengaruhi Angka Harapan Hidup di Jawa Timur yang meliputi faktor sosial, ekonomi, kesehatan, dan pendidikan. Adapun variabel yang

memberikan pengaruh signifikan adalah angka kematian bayi, persentase bayi berusia 0-11 bulan yang diberi ASI selama 4-6 bulan, dan variabel persentase balita berusia 1-4 tahun yang mendapatkan imunisasi lengkap. Sementara itu, Juliandari (2014) mendapatkan bahwa Angka Harapan Hidup dan Angka Kematian Bayi dipengaruhi oleh presentase rumah tangga menggunakan air minum dari air kemasan, presentase bayi usia 0-11 bulan yang pernah diberi ASI selama 1-3 bulan, presentase persalinan dibantu tenaga medis, presentase persalinan dibantu tenaga non medis, dan laju pertumbuhan ekonomi.

Penelitian lain yang dilakukan oleh Handayani (2014) terhadap derajat kesehatan menyebutkan bahwa faktor ekonomi, faktor pendidikan, faktor lingkungan dan faktor pelayanan kesehatan berpengaruh terhadap derajat kesehatan. Pengaruh faktor ekonomi ditunjukkan dengan keterkaitan antara Pendapatan Domestik Regional Bruto (PRDB) daerah dengan indeks derajat kesehatan masyarakat. Pengaruh faktor pendidikan ditunjukkan dengan keterkaitan antara rata-rata lama sekolah perempuan dan angka melek huruf perempuan terhadap indeks derajat kesehatan masyarakat. Sedangkan pengaruh faktor lingkungan ditunjukkan dengan keterkaitan antara persentase rumah tangga yang memiliki akses terhadap sumber air minum berkualitas dan persentase rumah tangga yang memiliki sanitasi layak pakai sesuai kriteria MDGs dengan indeks derajat kesehatan masyarakat. Sedangkan pengaruh faktor pelayanan kesehatan ditunjukkan dengan keterkaitan antara rasio Posyandu terhadap jumlah desa, persentase bayi yang diimunisasi campak, rasio jumlah Puskesmas terhadap penduduk dan rasio dokter umum di Puskesmas terhadap jumlah Puskesmas terhadap indeks derajat kesehatan masyarakat.

Selain penelitian diatas, pada tahun 1974 Hendrik L. Blum mengeluarkan sebuah teori yang menyebutkan bahwa ada 4 faktor yang mempengaruhi status derajat kesehatan masyarakat atau perorangan (Marisa, 2017). Faktor-faktor tersebut antara lain:

1. Lingkungan

Lingkungan memiliki pengaruh dan peranan terbesar diikuti perilaku, fasilitas kesehatan dan keturunan. Lingkungan sangat bervariasi, umumnya digolongkan menjadi tiga kategori, yaitu yang berhubungan dengan aspek

fisik dan sosial. Lingkungan yang berhubungan dengan aspek fisik contohnya sampah, air, udara, tanah, iklim, perumahan, dan sebagainya. Sedangkan lingkungan sosial merupakan hasil interaksi antar manusia seperti kebudayaan, pendidikan, ekonomi, dan sebagainya.

2. Perilaku

Perilaku merupakan faktor kedua yang mempengaruhi derajat kesehatan masyarakat karena sehat atau tidak sehatnya lingkungan kesehatan individu, keluarga dan masyarakat sangat tergantung pada perilaku manusia itu sendiri. Di samping itu, juga dipengaruhi oleh kebiasaan, adat istiadat, kebiasaan, kepercayaan, pendidikan sosial ekonomi, dan perilaku-perilaku lain yang melekat pada dirinya.

3. Pelayanan Kesehatan

Pelayanan kesehatan merupakan faktor ketiga yang mempengaruhi derajat kesehatan masyarakat karena keberadaan fasilitas kesehatan sangat menentukan dalam pelayanan pemulihan kesehatan, pencegahan terhadap penyakit, pengobatan dan keperawatan serta kelompok dan masyarakat yang memerlukan pelayanan kesehatan. Ketersediaan fasilitas dipengaruhi oleh lokasi, apakah dapat dijangkau atau tidak. Yang kedua adalah tenaga kesehatan pemberi pelayanan, informasi dan motivasi masyarakat untuk mendatangi fasilitas dalam memperoleh pelayanan serta program pelayanan kesehatan itu sendiri apakah sesuai dengan kebutuhan masyarakat yang memerlukan

4. Keturunan

Keturunan (genetik) merupakan faktor yang telah ada dalam diri manusia yang dibawa sejak lahir, misalnya dari golongan penyakit keturunan seperti diabetes melitus dan asma bronchial.

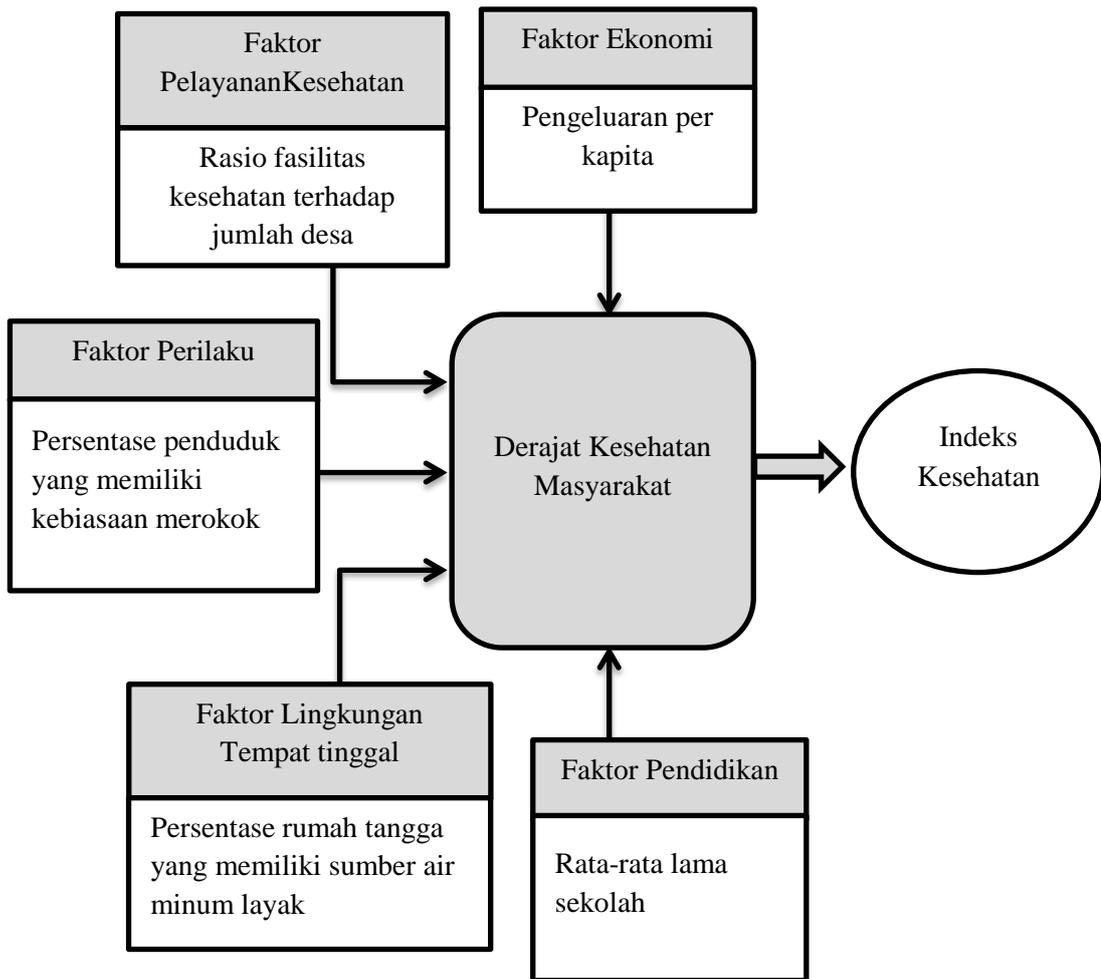
Selain itu menurut OECD (2015) yang merupakan salah satu organisasi multi negara yang didirikan untuk mempererat kerjasama dan pembangunan ekonomi antar negara demi mewujudkan stabilitas perekonomian yang berkelanjutan menyebutkan bahwa pencapaian angka harapan hidup saat lahir dapat ditunjukkan dengan peningkatan standar hidup, perubahan perilaku, pendidikan yang lebih baik, dan kemudahan akses kualitas pelayanan kesehatan.

Pendapatan nasional yang tinggi (diukur dengan pendapatan per kapita) pada umumnya berhubungan dengan angka harapan hidup saat lahir yang tinggi, walaupun hubungan ini kurang dapat dijelaskan untuk level pendapatan nasional yang tertinggi. Hal yang sama juga terlihat pada pengeluaran untuk kesehatan dimana pengeluaran kesehatan tinggi berhubungan dengan tingginya angka harapan hidup saat lahir. Kemudian, AHMAC (2012) menyebutkan bahwa angka harapan hidup dipengaruhi banyak faktor antara lain perilaku kesehatan seperti merokok; faktor sosial seperti pendidikan, pendapatan dan pengangguran; akses ke pelayanan kesehatan; dan faktor lingkungan seperti perumahan yang terlalu padat, kurangnya air minum bersih dan sanitasi yang tidak layak.

Berdasarkan dari literatur-literatur diatas dapat diambil sebuah kesimpulan bahwa indeks kesehatan dipengaruhi oleh lima faktor, yaitu faktor kesehatan, perilaku, lingkungan, ekonomi dan faktor pendidikan yang dijelaskan pada Gambar 2.2.

Dari Gambar 2.2, faktor pelayanan kesehatan dapat dijelaskan melalui rasio fasilitas kesehatan terhadap jumlah desa, yaitu perbandingan jumlah fasilitas kesehatan yang berupa posyandu, puskesmas dan rumah sakit terhadap jumlah desa pada setiap kabupaten/kota di Provinsi Papua. Faktor ekonomi dijelaskan oleh variabel pendapatan per kapita. Dalam penelitian ini variabel yang digunakan adalah pengeluaran per kapita karena datanya sudah dipublikasikan oleh BPS.

Variabel-variabel yang menjelaskan faktor perilaku yang berhubungan dengan kesehatan pada penelitian sebelumnya antara lain persentase penduduk yang memiliki kebiasaan merokok. Faktor lingkungan dijelaskan oleh variable-variabel yang dipublikasikan BPS mengenai perumahan. Variabel tersebut antara lain persentase rumah tangga yang memiliki jenis atap layak, lantai bukan tanah, dinding permanen, sumber air minum layak, fasilitas BAB sendiri, sanitasi layak, dan tempat pembuangan akhir tinja yang layak. Dalam penelitian ini digunakan variabel perumahan yang signifikan pada penelitian-penelitian sebelumnya yaitu sumber air minum layak. Sedangkan untuk faktor pendidikan dijelaskan oleh variabel rata-rata lama sekolah.



Gambar 2.2 Kerangka Pikir Faktor Yang Mempengaruhi Indeks Kesehatan

(Halaman ini sengaja dikosongkan)

BAB 3

METODOLOGI PENELITIAN

Pada bab ini akan dibahas tentang sumber data yang digunakan, kemudian variabel penelitian, definisi operasional dan struktur data pada penelitian ini. Selain itu juga dijelaskan tahapan penelitian dari estimasi parameter sampai mendapatkan kesimpulan.

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS). Data bersumber pada publikasi Indeks Pembangunan Manusia (IPM) 2016 Metode Baru, *raw* data Pendataan Potensi Desa (PODES) 2014, serta publikasi Statistik Kesejahteraan Rakyat 2016. Variabel indeks kesehatan, rata-rata lama sekolah, dan pengeluaran per kapita bersumber dari publikasi IPM 2016. Variabel persentase rumah tangga yang memiliki sumber air minum layak, dan persentase penduduk yang memiliki kebiasaan merokok dari publikasi Statistik Kesejahteraan Rakyat 2016. Variabel rasio fasilitas kesehatan terhadap jumlah desa bersumber dari *raw* data PODES 2014. Penggunaan data PODES dengan tahun 2014 ini berbeda dengan data lainnya yang menggunakan tahun 2016 dikarenakan keterbatasan dan ketiadaan data tentang fasilitas kesehatan pada tahun 2016, sehingga diasumsikan selama kurun waktu 2 tahun berjalan yaitu dari tahun 2014-2016 tidak ada perubahan jumlah fasilitas kesehatan yang signifikan. Unit analisis yang diteliti adalah kabupaten/kota di Pulau Papua dengan jumlah kabupaten/kota sebanyak 42 kabupaten/kota.

3.2 Variabel Penelitian

Berdasarkan hasil penelitian-penelitian pada bab sebelumnya, maka variabel-variabel yang diduga mempunyai korelasi dengan indeks kesehatan yang akan digunakan dalam penelitian ini terdiri dari satu variabel respon dan lima variabel prediktor. Variabel respon (Y) adalah indeks kesehatan, dimana variabel

ini bersifat kategorik. Menurut BPS (2015) indeks kesehatan seperti pada IPM dibagi menjadi empat kategori, yaitu sebagai berikut:

0 = Rendah, jika indeks kesehatan < 60

1 = Sedang, jika $60 \leq$ indeks kesehatan < 70

2 = Tinggi, jika $70 \leq$ indeks kesehatan < 80

3 = Sangat tinggi, jika indeks kesehatan ≥ 80

Pada penelitian ini, jika indeks kesehatan dibagi sesuai dengan ketentuan BPS yaitu empat kategori, maka terdapat beberapa kategori yang kosong. Dengan demikian pada penelitian ini pembagian indeks kesehatan dibatasi menjadi dua kategori yaitu:

0 = Menengah Rendah, merupakan gabungan dari kategori rendah dan sedang yaitu jika indeks kesehatan < 70

1 = Tinggi, merupakan gabungan dari kategori tinggi dan sangat tinggi yaitu jika indeks kesehatan ≥ 70

Variabel penelitian yang akan digunakan dalam penelitian ini secara lengkap seperti pada Tabel 3.1, definisi operasional variabel beserta satuannya terdapat pada Tabel 3.2 dan struktur data pada Tabel 3.3.

Tabel 3.1 Variabel Penelitian

Variabel	Tipe Data	Keterangan
Indeks Kesehatan (Y)	Nominal	0 = Menengah Rendah 1 = Tinggi
Pengeluaran per kapita (X_1)	Rasio	-
Persentase rumah tangga yang memiliki sumber air minum layak (X_2)	Rasio	-
Persentase penduduk yang memiliki kebiasaan merokok (X_3)	Rasio	-
Rata-rata lama sekolah (X_4)	Rasio	-
Rasio fasilitas kesehatan terhadap jumlah desa (X_5)	Rasio	-

Tabel 3.2 Definisi Operasional Variabel dan Satuan Variabel

Variabel	Definisi Operasional Variabel	Satuan
Indeks Kesehatan (Y)	Indeks yang disusun berdasarkan capaian Angka Harapan Hidup sebagaimana dijelaskan pada subbab 2.9	-
Pengeluaran per kapita (X_1)	Rata-rata biaya yang dikeluarkan untuk konsumsi (konsumsi makanan dan non makanan) setiap penduduk di masing-masing kabupaten/kota di Papua dan Papua Barat selama sebulan	Juta Rupiah
Persentase rumah tangga yang memiliki sumber air minum layak (X_2)	Persentase rumah tangga yang memiliki sumber air minum Air kemasan/Leding meteran/SumurBor/ Pompa/Sumur Mata Air Terlindung/Air hujan pada setiap kabupaten/kota di Provinsi Papua dan Papua Barat	Persentase
Persentase penduduk yang memiliki kebiasaan merokok (X_3)	Persentase penduduk 5 tahun ke atas yang memiliki kebiasaan merokok tembakau pada setiap kabupaten/kota di Provinsi Papua dan Papua Barat	Persentase
Rata-rata lama sekolah (X_4)	Jumlah tahun yang digunakan oleh penduduk 25 tahun ke atas dalam menjalani pendidikan formal pada setiap kabupaten/kota di Provinsi Papua	Tahun
Rasio fasilitas kesehatan terhadap jumlah desa (X_5)	Perbandingan jumlah fasilitas kesehatan yang berupa posyandu, puskesmas dan rumah sakit terhadap jumlah desa pada setiap kabupaten/kota di Provinsi Papua dan Papua Barat	-

Tabel 3.3 Struktur Data Penelitian

Kabupaten/ Kota	Y	X_1	X_2	X_3	X_4	X_5
1	Y_1	$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	$X_{1,4}$	$X_{1,5}$
2	Y_2	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	$X_{2,4}$	$X_{2,5}$
3	Y_3	$X_{3,1}$	$X_{3,2}$	$X_{3,3}$	$X_{3,4}$	$X_{3,5}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
42	Y_{42}	$X_{42,1}$	$X_{42,2}$	$X_{42,3}$	$X_{42,4}$	$X_{42,5}$

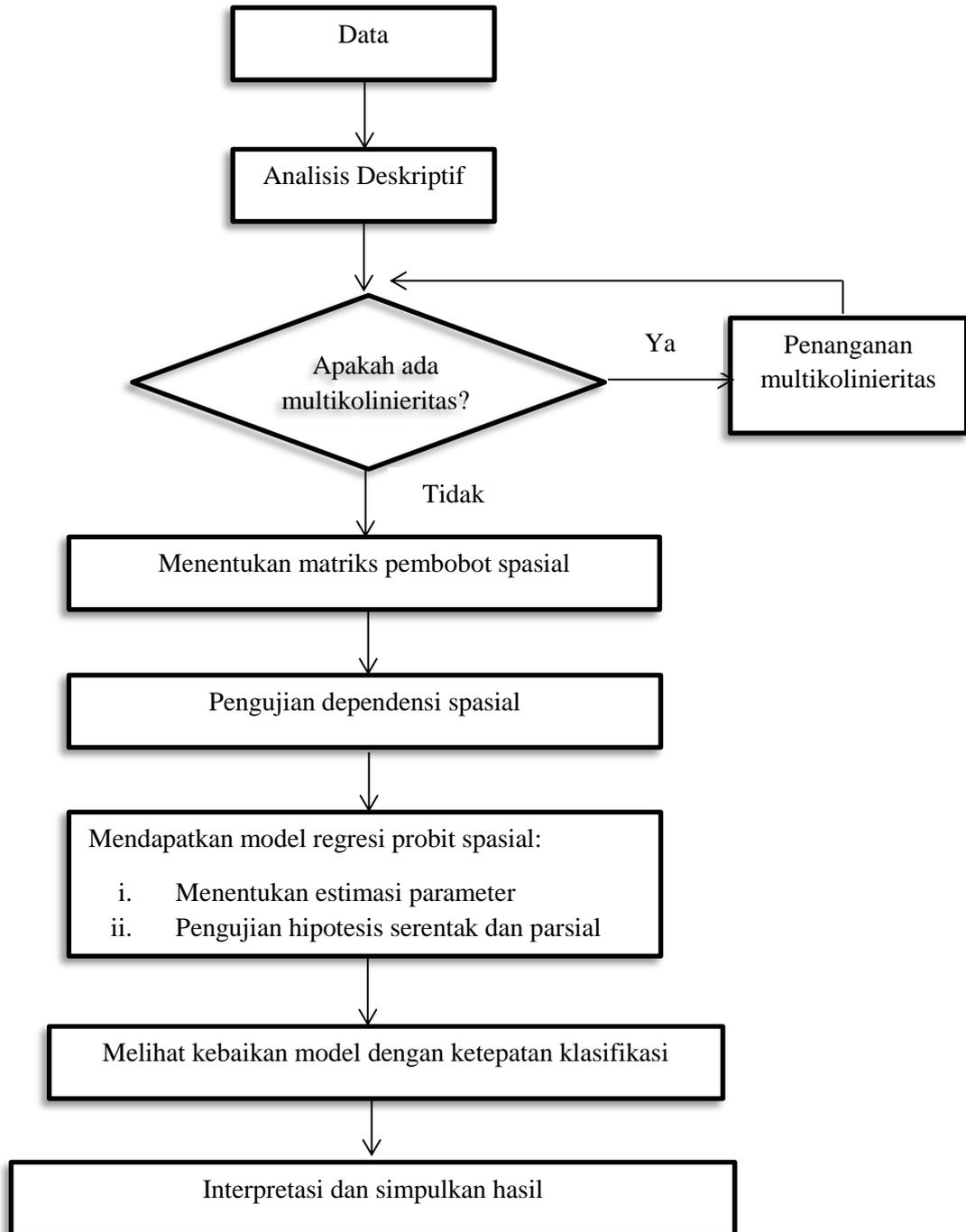
3.3 Tahapan Penelitian

Tahapan penelitian yang digunakan sebagai langkah-langkah untuk mencapai tujuan penelitian dijabarkan sebagai berikut:

1. Mendapatkan bentuk penaksir parameter model regresi probit spasial menggunakan metode RIS simulator, dengan langkah-langkah sebagai berikut:
 - a. Mempersiapkan data yang akan digunakan dalam penelitian
 - b. Menentukan jumlah simulasi (R)
 - c. Melakukan transformasi dengan menggunakan dekomposisi cholesky, sehingga diperoleh persamaan (2.35)
 - d. Menghitung batas atas kumulatif (p_i) pada masing-masing simulasi, dengan prosedur simulasi sebagai berikut:
 - (i) Menghitung batas atas kumulatif standar normal pada observasi ke- n (p_n), yang diperoleh dari persamaan (2.36) sehingga didapatkan nilai dari $b_{n,n}^{-1}c_n$
 - (ii) Mengambil nilai $\tilde{\eta}_n$ yang berasal dari perpotongan distribusi normal standar pada $b_{n,n}^{-1}c_n$
 - (iii) Menghitung batas atas kumulatif standar normal pada observasi $n-1$ (p_{n-1}), yang diperoleh dari persamaan (2.36) dan memasukkan nilai $\tilde{\eta}_n$ yang telah didapat pada langkah (ii) sehingga didapatkan nilai dari $b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\tilde{\eta}_n)$
 - (iv) Mengambil nilai $\tilde{\eta}_{n-1}$ yang berasal dari perpotongan distribusi normal standar pada $b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\tilde{\eta}_n)$
 - (v) Menghitung batas atas kumulatif standar normal pada observasi $n-2$ (p_{n-2}), yang diperoleh dari persamaan (2.36) dan memasukkan nilai $\tilde{\eta}_n$ dan $\tilde{\eta}_{n-1}$ yang telah didapat pada langkah (ii) dan (iv), sehingga didapatkan nilai dari $b_{n-2,n-1}^{-1}(c_{n-2} - b_{n-1,n-1}\tilde{\eta}_{n-1} - b_{n-2,n}\tilde{\eta}_n)$
 - (vi) Mengambil nilai $\tilde{\eta}_{n-2}$ yang berasal dari perpotongan distribusi normal standar pada $b_{n-2,n-1}^{-1}(c_{n-2} - b_{n-1,n-1}\tilde{\eta}_{n-1} - b_{n-2,n}\tilde{\eta}_n)$
 - (vii) Proses terus dilanjutkan untuk semua observasi dan berakhir pada observasi ke-1

- (viii) Mendapatkan *cross product* dari seluruh nilai p_i , sehingga didapatkan persamaan (2.38)
 - e. Mengulangi langkah (i)-(viii) untuk seluruh simulasi, yaitu untuk $r = 1, 2, \dots, R$
 - f. Menghitung probabilitas simulasi (\hat{p}) yang didapatkan dari rata-rata dari seluruh simulasi sesuai dengan persamaan (2.39)
 - g. Melakukan optimasi secara konvensional terhadap $\ln \hat{p}$ untuk mendapatkan estimasi parameter
2. Memodelkan faktor-faktor yang mempengaruhi pola hubungan klasifikasi indeks kesehatan terhadap variabel prediktor, dengan langkah-langkah sebagai berikut:
- a. Melakukan analisis deskriptif pada masing-masing variabel
 - b. Mengidentifikasi masalah multikolinearitas dengan melihat nilai VIF sesuai dengan persamaan (2.13) dan lakukan penanganan apabila terdapat variabel prediktor yang mengalami multikolinearitas.
 - c. Membuat matriks pembobot antar lokasi dengan menggunakan *Queen Contiguity*.
 - d. Melakukan pengujian dependensi spasial sesuai dengan persamaan (2.19)
 - e. Melakukan estimasi parameter dan pembentukan model regresi probit spasial dengan menggunakan variabel-variabel terpilih.
 - f. Melakukan pengujian hipotesis serentak dan parsial sesuai dengan persamaan (2.40) dan (2.41)
 - g. Melakukan pengujian kriteria kebaikan model sesuai dengan persamaan (2.42) sampai dengan persamaan (2.44)
 - h. Menginterpretasikan dan menyimpulkan hasil yang telah diperoleh.

Skema tahapan penelitian di atas dapat dilihat pada Gambar 3.1.



Gambar 3.1 Skema Tahapan Penelitian

BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas mengenai estimasi parameter regresi probit spasial dengan metode *Maximum Likelihood Estimation* (MLE) dengan menggunakan RIS simulator, dan pemodelan indeks kesehatan kabupaten/kota di Pulau Papua dengan menggunakan regresi probit spasial. Pemodelan dilakukan untuk mengetahui faktor-faktor yang mempengaruhi indeks kesehatan, serta untuk memprediksi probabilitas suatu kabupaten/kota memiliki indeks kesehatan menengah rendah atau tinggi.

4.1 Estimasi Parameter Regresi Probit Spasial

Estimasi parameter model regresi probit spasial dalam penelitian ini menggunakan metode *Maximum Likelihood Estimation* (MLE). Metode MLE adalah metode mendapatkan estimasi parameter dengan memaksimumkan fungsi *ln-likelihood*. Variabel random Y pada model regresi probit biner spasial berdistribusi Bernoulli $(1, p)$ sehingga fungsi *likelihood* dari Y sesuai dengan persamaan (2.26) dan (2.27) adalah sebagai berikut:

$$L(\boldsymbol{\beta}, \rho) = \prod_{i=1}^n \left(\left[1 - \Phi \left(\frac{-[\boldsymbol{\Gamma}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \right]^{y_i} \right) \left(\left[\Phi \left(\frac{-[\boldsymbol{\Gamma}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \right]^{1-y_i} \right)$$

dimana $\boldsymbol{\Gamma} = (\mathbf{I}_n - \rho\mathbf{W})^{-1}$, $\boldsymbol{\Gamma}$ merupakan matriks berukuran $(n \times n)$. \mathbf{I}_n adalah matriks identitas berukuran $(n \times n)$. \mathbf{W} yaitu matriks pembobot berukuran $(n \times n)$ dan ρ adalah parameter spasial. \mathbf{X} adalah matriks variabel prediktor yang berukuran $n \times (k+1)$ dan $\boldsymbol{\beta}$ adalah vektor koefisien parameter regresi berukuran $(k+1) \times 1$.

Kemudian dilakukan tranformasi fungsi *ln-likelihood*, yaitu menjadi:

$$\ln[L(\boldsymbol{\beta}, \rho)] = \sum_{i=1}^n y_i \ln \left[1 - \Phi \left(\frac{-[\boldsymbol{\Gamma}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right) \right] + \sum_{i=1}^n (1 - y_i) \ln \Phi \left(\frac{-[\boldsymbol{\Gamma}\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \right)$$

Sebagai syarat perlu untuk memaksimumkan fungsi *ln-likelihood* adalah mencari turunan pertama dari fungsi *ln-likelihood* terhadap masing-masing parameter dan disamakan dengan nol. Turunan fungsi terhadap masing-masing parameter $\boldsymbol{\beta}$ dan ρ , yang kemudian disamakan dengan nol adalah sebagai berikut:

a. Turunan fungsi ln-likelihood terhadap parameter β

$$\begin{aligned}\frac{\partial \ln[L(\beta, \rho)]}{\partial \beta} &= \frac{\partial}{\partial \beta} \left[\sum_{i=1}^n y_i \ln \left[1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \right] + \sum_{i=1}^n (1 - y_i) \ln \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \right] \\ \frac{\partial \ln[L(\beta, \rho)]}{\partial \beta} &= \sum_{i=1}^n y_i \frac{\partial \ln \left[1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \right]}{\partial \beta} + \sum_{i=1}^n (1 - y_i) \frac{\partial \ln \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)}{\partial \beta} \\ &= \sum_{i=1}^n \left[\frac{y_i}{1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} \frac{\partial \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)}{\partial \beta} + \frac{1 - y_i}{\Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} \frac{\partial \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)}{\partial \beta} \right] \\ &= \sum_{i=1}^n \frac{\partial \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)}{\partial \beta} \left[\frac{y_i}{1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} + \frac{1 - y_i}{\Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} \right] \quad (4.1)\end{aligned}$$

Untuk menyelesaikan persamaan (4.1), digunakan Lemma 2.1, maka dihasilkan:

$$\frac{\partial \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)}{\partial \beta} = \frac{-[\Gamma \mathbf{X}]_i}{\Omega_{ii}} \phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)$$

Sehingga turunan *ln-likelihood* terhadap β adalah

$$\frac{\partial \ln[L(\beta, \rho)]}{\partial \beta} = \sum_{i=1}^n \frac{-[\Gamma \mathbf{X}]_i}{\Omega_{ii}} \phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \left[\frac{y_i}{1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} + \frac{1 - y_i}{\Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} \right] \quad (4.2)$$

Untuk mendapatkan penaksir parameter β , persamaan (4.2) disamakan dengan nol.

$$\frac{\partial \ln[L(\beta, \rho)]}{\partial \beta} = 0$$

$$\sum_{i=1}^n \frac{-[\Gamma \mathbf{X}]_i}{\Omega_{ii}} \phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \left[\frac{y_i}{1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} + \frac{1 - y_i}{\Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)} \right] = 0 \quad (4.3)$$

b. Turunan fungsi ln-likelihood terhadap parameter ρ

$$\begin{aligned}\frac{\partial \ln[L(\beta, \rho)]}{\partial \rho} &= \frac{\partial}{\partial \rho} \left[\sum_{i=1}^n y_i \ln \left[1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \right] + \sum_{i=1}^n (1 - y_i) \ln \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \right] \\ &= \sum_{i=1}^n y_i \frac{\partial \ln \left[1 - \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right) \right]}{\partial \rho} + \sum_{i=1}^n (1 - y_i) \frac{\partial \ln \Phi \left(\frac{-[\Gamma \mathbf{X} \beta]_i}{\Omega_{ii}} \right)}{\partial \rho}\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n \left[\frac{y_i}{1 - \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} \frac{\partial \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)}{\partial \rho} + \frac{1 - y_i}{\Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} \frac{\partial \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)}{\partial \rho} \right] \\
&= \sum_{i=1}^n \frac{\partial \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)}{\partial \rho} \left[\frac{y_i}{1 - \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} + \frac{1 - y_i}{\Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} \right] \quad (4.4)
\end{aligned}$$

dimana nilai dari $\boldsymbol{\Gamma} = (\mathbf{I}_n - \rho\mathbf{W})^{-1}$ dan menurut Martinetti dan Geneaux (2017) penggunaan *Taylor expansion* dapat dilakukan terhadap $(\mathbf{I} - \rho\mathbf{W})^{-1}$, sehingga $(\mathbf{I}_n - \rho\mathbf{W})^{-1} = \mathbf{I}_n + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \dots$. Hal ini dapat dilakukan dikarenakan \mathbf{W} merupakan *row-standardised* dan parameter spasial lag ρ bernilai $[-1,1]$.

Berdasarkan Lemma 2.1, maka:

$$\begin{aligned}
\frac{\partial \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)}{\partial \rho} &= \frac{\partial \Phi\left(\frac{-[(\mathbf{I}_n + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \dots)\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)}{\partial \rho} \\
&= \frac{-[(\mathbf{W} + 2\rho\mathbf{W}^2 + 3\rho^2\mathbf{W}^3 + \dots)\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)
\end{aligned}$$

Sehingga turunan *ln-likelihood* terhadap ρ adalah

$$\begin{aligned}
\frac{\partial \ln[L(\boldsymbol{\beta}, \rho)]}{\partial \rho} &= \\
\sum_{i=1}^n \frac{-[(\mathbf{W} + 2\rho\mathbf{W}^2 + 3\rho^2\mathbf{W}^3 + \dots)\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right) &\left[\frac{y_i}{1 - \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} + \frac{1 - y_i}{\Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} \right] \quad (4.5)
\end{aligned}$$

Untuk mendapatkan penaksir parameter ρ , persamaan (4.5) disamakan dengan nol.

$$\begin{aligned}
\frac{\partial \ln[L(\boldsymbol{\beta}, \rho)]}{\partial \rho} &= 0 \\
\sum_{i=1}^n \frac{-[(\mathbf{W} + 2\rho\mathbf{W}^2 + 3\rho^2\mathbf{W}^3 + \dots)\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}} \phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right) &\left[\frac{y_i}{1 - \Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} + \frac{1 - y_i}{\Phi\left(\frac{-[\Gamma\mathbf{X}\boldsymbol{\beta}]_i}{\boldsymbol{\Omega}_{ii}}\right)} \right] = 0 \quad (4.6)
\end{aligned}$$

Karena estimasi parameter dengan metode MLE pada persamaan (4.3) dan persamaan (4.6) menghasilkan bentuk yang tidak *closed form*, maka penyelesaiannya harus menggunakan iterasi numerik untuk mendapatkan nilai

penaksir parameter. Untuk mengatasi kesulitan ini, Beron dan Vijverberg (2004) menyarankan penggunaan RIS simulator yang kemudian hasilnya dioptimasi untuk mendapatkan nilai estimasi parameternya. Tahapan penggunaan RIS simulator adalah sebagai berikut:

- a. Menginput data \mathbf{y} , \mathbf{X} dan \mathbf{W}
- b. Menentukan jumlah simulasi (R)
- c. Mendefinisikan

$$\boldsymbol{\beta}^{(t)} = \theta_i^{(t)}, \text{ dimana } i = 0, 1, 2, \dots, p$$

$$\rho^{(t)} = -1 + 2\Phi(\theta_i^{(t)}), \text{ dimana } i = p + 1$$

- d. Menentukan nilai inisiasi awal dari paramater, yaitu dengan nilai t dimulai dari nol.

$$\theta^{(t)} = \begin{bmatrix} \theta_0^{(t)} \\ \theta_1^{(t)} \\ \vdots \\ \theta_p^{(t)} \\ \theta_{p+1}^{(t)} \end{bmatrix}, \text{ dimulai } t = 0, \text{ sehingga menjadi } \theta^{(0)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

- e. Menghitung nilai matriks diagonal \mathbf{Z} , dengan $Z_{ii} = (1 - 2y_i)$
- f. Menghitung nilai $\boldsymbol{\Gamma} = (\mathbf{I} - \rho^{(t)}\mathbf{W})^{-1}$
- g. Menghitung nilai $\boldsymbol{\Omega} = \mathbf{Z}\boldsymbol{\Gamma}\boldsymbol{\Gamma}^T\mathbf{Z}^T$
- h. Menghitung nilai $\mathbf{v} = -\mathbf{Z}\boldsymbol{\Gamma}\mathbf{X}\boldsymbol{\beta}^{(t)}$
- i. Pada setiap simulasi, hitung nilai $\boldsymbol{\Omega}^{-1}$, kemudian melakukan transformasi dekomposisi cholesky $\boldsymbol{\Omega}^{-1} = \mathbf{A}^T\mathbf{A}$, sehingga didapatkan matriks \mathbf{A} .
- j. Mendefinisikan nilai $\boldsymbol{\eta} = \mathbf{A}\mathbf{v}$, sehingga menjadi:

$$\mathbf{v} = \mathbf{A}^{-1}\boldsymbol{\eta}, \text{ dimisalkan } \mathbf{A}^{-1} = \mathbf{B}$$

$$\mathbf{v} = \mathbf{B}\boldsymbol{\eta}$$

- k. Mensubstitusikan nilai $\mathbf{B}\boldsymbol{\eta}$ terhadap \mathbf{v} pada masing-masing simulasi, sehingga terbentuk:

$$\begin{aligned} P((v_i < -[\mathbf{Z}\boldsymbol{\Gamma}\mathbf{X}\boldsymbol{\beta}]_i)) &= P((v_i < c_i)) \\ &= P(([\mathbf{B}\boldsymbol{\eta}]_i < c_i)) \end{aligned}$$

kemudian didapatkan probabiitas individu ke-j:

$$P\left(\sum_{i=1}^n b_{1,i}\eta_i < c_1\right) = P(\eta_1 < b_{1,1}^{-1}\left(c_1 - \sum_{i=2}^n b_{1,i}\eta_i\right)) \equiv p_1$$

\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots

$$P(b_{n-1,n-1}\eta_{n-1} + b_{n-1,n}\eta_n < c_{n-1}) = P(\eta_{n-1} < b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\eta_n)) \equiv p_{n-1}$$

$$P(b_{n,n}\eta_n < c_n) = P(\eta_n < b_{n,n}^{-1}c_n) \equiv p_n$$

1. Menghitung probabilitas individu ke- j (p_j) pada masing-masing simulasi, dengan prosedur *recursive* sebagai berikut:

- 1) Menghitung p_n :

$$P\left(\eta_n < \frac{c_n}{b_{n,n}}\right) = \Phi\left(\frac{c_n}{b_{n,n}}\right)$$

- 2) Mengambil nilai η_n , yang diberi label $\tilde{\eta}_n$ yang berasal dari perpotongan distribusi normal standar pada $b_{n,n}^{-1}c_n$. Pengambilan dapat dilakukan dengan cara sebagai berikut:

- a) Mengambil dari distribusi standar uniform $\tilde{\mu}_n$

- b) Menghitung nilai $\tilde{\eta}_n = \Phi^{-1}\left(\tilde{\mu}_n \Phi(b_{n,n}^{-1}c_n)\right)$

- 3) Menghitung p_{n-1} :

$$P\left(\eta_{n-1} < \frac{(c_{n-1} - b_{n-1,n}\eta_n)}{b_{n-1,n-1}} \middle| \eta_n = \tilde{\eta}_n\right) = \Phi\left(\frac{(c_{n-1} - b_{n-1,n}\tilde{\eta}_n)}{b_{n-1,n-1}}\right)$$

- 4) Mengambil nilai η_{n-1} , yang diberi label $\tilde{\eta}_{n-1}$ yang berasal dari perpotongan distribusi normal standar pada $b_{n-1,n-1}^{-1}(c_{n-1} - b_{n-1,n}\tilde{\eta}_n)$. Pengambilan dapat dilakukan dengan cara sebagai berikut:

- a) Mengambil dari distribusi standar uniform $\tilde{\mu}_{n-1}$

- b) Menghitung nilai $\tilde{\eta}_{n-1} = \Phi^{-1}\left(\tilde{\mu}_{n-1} \Phi\left(\frac{(c_{n-1} - b_{n-1,n}\tilde{\eta}_n)}{b_{n-1,n-1}}\right)\right)$

- 5) Menghitung p_{n-2} :

$$P\left(\eta_{n-2} < \frac{(c_{n-2} - b_{n-1,n-1}\eta_{n-1} - b_{n-2,n}\eta_n)}{b_{n-2,n-1}} \middle| \eta_n = \tilde{\eta}_n, \eta_{n-1} = \tilde{\eta}_{n-1}\right) = \Phi\left(\frac{(c_{n-2} - b_{n-1,n-1}\tilde{\eta}_{n-1} - b_{n-2,n}\tilde{\eta}_n)}{b_{n-2,n-1}}\right)$$

- 6) Mengambil nilai η_{n-2} , yang diberi label $\tilde{\eta}_{n-2}$, dimana nilai tersebut berasal dari perpotongan distribusi normal standar pada $b_{n-2,n-1}^{-1}(c_{n-2} -$

$b_{n-1,n-1}\tilde{\eta}_{n-1} - b_{n-2,n}\tilde{\eta}_n$). Pengambilan dapat dilakukan dengan cara sebagai berikut:

a) Mengambil dari distribusi standar uniform $\tilde{\mu}_{n-2}$

b) Menghitung nilai $\tilde{\eta}_{n-2} = \Phi^{-1}\left(\tilde{\mu}_{n-2}\Phi\left(\frac{(c_{n-2}-b_{n-1,n-1}\tilde{\eta}_{n-1}-b_{n-2,n}\tilde{\eta}_n)}{b_{n-2,n-1}}\right)\right)$

7) Proses terus dilanjutkan untuk semua observasi dan berakhir pada observasi ke-1

8) Mendapatkan *cross product* dari seluruh nilai p_j , sehingga didapatkan:

$$\hat{p}^r = \prod_{j=1}^n p_j = \prod_{j=1}^n \Phi(\eta_{j,0}).$$

m. Mengulangi langkah 1)-8) untuk seluruh simulasi, yaitu untuk $r = 1, 2, \dots, R$

n. Menghitung probabilitas simulasi (\hat{p}) yang didapatkan dari rata-rata dari seluruh simulasi sehingga didapatkan:

$$\hat{p} = \frac{1}{R} \sum_{r=1}^R \hat{p}^r$$

dimana nilai \hat{p} tersebut digunakan untuk mencari nilai standar *maximum-likelihood estimator* dari $L(\boldsymbol{\beta}, \rho)$, sehingga:

$$\begin{aligned} \ln L(\boldsymbol{\beta}, \rho) &= \ln \left(\frac{1}{R} \sum_{r=1}^R \hat{p}^r \right) \\ &= \ln \frac{1}{R} + \ln \sum_{r=1}^R \hat{p}^r \end{aligned} \quad (4.7)$$

o. Melakukan optimisasi terhadap fungsi $\ln L(\boldsymbol{\beta}, \rho)$ pada persamaan 4.8 secara konvensional yaitu dengan menggunakan iterasi Nelder Mead untuk mendapatkan estimator parameter. Menurut Lagarias dkk. (1998) langkah-langkah optimasinya sebagai berikut:

1) Definiskan fungsi pada persamaan 4.9 yaitu $\ln L(\boldsymbol{\beta}, \rho) = f(\theta)$ dan parameter yang ingin diestimasi sebagai θ

2) Tahap *Order*.

Order nilai $\theta_1, \theta_2, \dots, \theta_{n+1}$ dan hitunglah nilai $f(\theta_1), f(\theta_2), \dots, f(\theta_{n+1})$, kemudian urutkanlah sehingga $f_1 = f(\theta_1) \leq f_2 = f(\theta_2) \leq \dots \leq f_{n+1} = f(\theta_{n+1})$

3) Tahap *Reflection*.

Hitung titik refleksi $\theta_r = \bar{\theta} + \rho_{NM}(\bar{\theta} - \theta_{n+1})$ dengan $\bar{\theta} = \sum_{i=1}^n \theta_i / n$ merupakan centroid dari n nilai terbaik (semua simpul kecuali θ_{n+1}). Evaluasi $f_r = f(\theta_r)$. Jika $f_1 \leq f_r < f_n$, maka terima titik refleksi θ_r dan akhiri iterasi.

4) Tahap *Expansion*.

Jika $f_r < f_1$, maka hitung titik ekspansi $\theta_e = \bar{\theta} + \chi_{NM}(\theta_r - \bar{\theta})$, kemudian evaluasi $f_e = f(\theta_e)$.

Jika $f_e < f_r$, maka terima θ_e kemudian hentikan iterasi; dan jika $f_e \geq f_r$ maka terima θ_r kemudian akhiri proses iterasi.

5) Tahap *Contraction*.

Jika $f_r \geq f_n$, maka lakukan kontraksi antara $\bar{\theta}$ dengan θ_{n+1} atau θ_r .

a) Jika $f_n \leq f_r < f_{n+1}$ menunjukkan kontraksi sisi luar (*out side contraction*), hitung $\theta_c = \bar{\theta} + \gamma_{NM}(\theta_r - \bar{\theta})$, kemudian evaluasi $f_c = f(\theta_c)$.

Jika $f_c \leq f_r$, maka terima θ_c kemudian hentikan iterasi; dan jika $f_c > f_r$ maka lanjut ke tahap 6 (tahap *shrunked*).

b) Jika $f_r > f_{n+1}$ menunjukkan kontraksi sisi dalam (*inside contraction*), hitung $\theta_{cc} = \bar{\theta} + \gamma_{NM}(\bar{\theta} - \theta_{n+1})$, kemudian evaluasi $f_{cc} = f(\theta_{cc})$.

Jika $f_{cc} < f_{n+1}$, maka terima θ_{cc} kemudian hentikan iterasi; dan jika $f_{cc} \geq f_{n+1}$ maka lanjut ke tahap 6 (tahap *shrunked*)

6) Tahap *Shrunked*

Evaluasi f untuk n titik, $\vartheta_i = \theta_1 + \sigma_{NM}(\theta_i - \theta_1)$ dengan $i = 2, 3, \dots, (n + 1)$. Kemudian kembali ke langkah 2 dengan mengganti $\theta_i = \vartheta_i$ dengan $i = 1, 2, 3, \dots, (n + 1)$.

Dalam iterasi Nelder Mead, nilai standar parameter yang digunakan $\rho_{NM}=1$, $\chi_{NM}=2$, $\gamma_{NM}=0.5$ dan $\sigma_{NM}=0.5$.

- p. Nilai iterasi parameter θ pada langkah ke-n inilah yang secara berulang-ulang dimasukkan ke dalam langkah b, sampai didapatkan nilai estimator θ yang optimum dan konvergen, sehingga didapatkan estimator $\hat{\beta}$ dan $\hat{\rho}$.

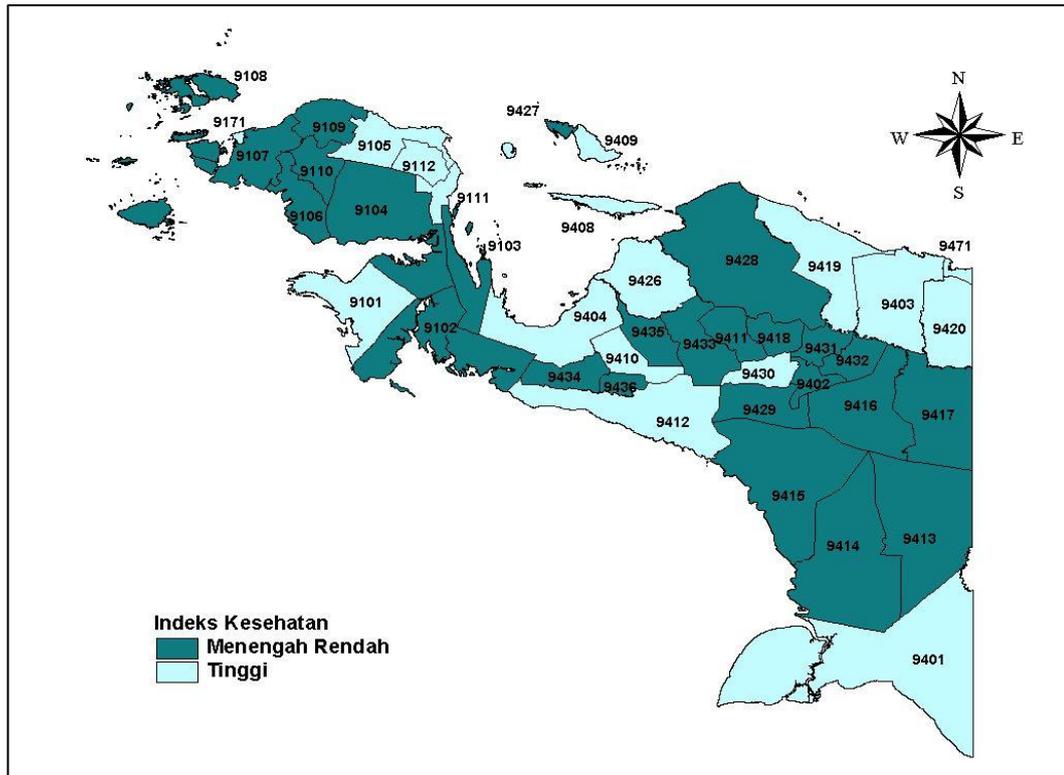
4.2 Pemodelan Indeks Kesehatan di Pulau Papua dengan Model Regresi Probit Spasial

Pemodelan indeks kesehatan menggunakan pendekatan regresi probit spasial diawali dengan melakukan analisis deskriptif terhadap variabel indeks kesehatan terhadap masing-masing variabel prediktor, pengujian multikolinieritas antar variabel prediktor, pengujian parameter baik secara serentak maupun parsial, memilih model terbaik melalui tabel ketepatan klasifikasi (*confusion matrix*), serta interpretasi model terbaik.

4.2.1 Gambaran Indeks Kesehatan Kabupaten/Kota di Pulau Papua

Pulau Papua terdiri dari dua provinsi yaitu Provinsi Papua dan Papua Barat. Provinsi Papua mempunyai 29 kabupaten/kota dan Provinsi Papua Barat mempunyai 13 kabupaten/kota, sehingga jumlah kabupaten/kota di Pulau Papua berjumlah 42 kabupaten/kota. Dalam penelitian ini capaian indeks kesehatan dikategorikan menjadi 2 kategori, yaitu: menengah rendah dan tinggi. Dari kedua kategori tersebut, terdapat 25 kabupaten/kota yang tergolong ke kategori menengah rendah dan terdapat 17 kabupaten/kota yang tergolong ke kategori tinggi, atau secara persentase dapat dilihat bahwa 59,52 persen berkategori menengah rendah dan 40,48 persen berkategori tinggi. Wilayah di Pulau Papua yang memiliki capaian indeks kesehatan terendah terletak di Kabupaten Nduga, Provinsi Papua yaitu sebesar 53,08, sedangkan capaian indeks kesehatan tertinggi terletak di Kabupaten Mimika yaitu sebesar 79,85.

Persebaran indeks kesehatan di Pulau Papua menurut kabupaten/kota pada tahun 2016 disajikan pada Gambar 4.1. Degradasi warna yang terdapat pada gambar menunjukkan bahwa warna gelap menunjukkan indeks kesehatan di wilayah tersebut tergolong menengah rendah, sebaliknya warna terang menunjukkan indeks kesehatan di wilayah tersebut tergolong tinggi. Selain itu, berdasarkan Gambar 4.1 dapat diketahui bahwa beberapa wilayah yang berdekatan memiliki kecenderungan nilai indeks kesehatan yang tidak jauh berbeda, terutama pada antar wilayah pegunungan dan antar wilayah pesisir, hal ini mengindikasikan adanya dependensi spasial pada data indeks kesehatan di Pulau Papua.

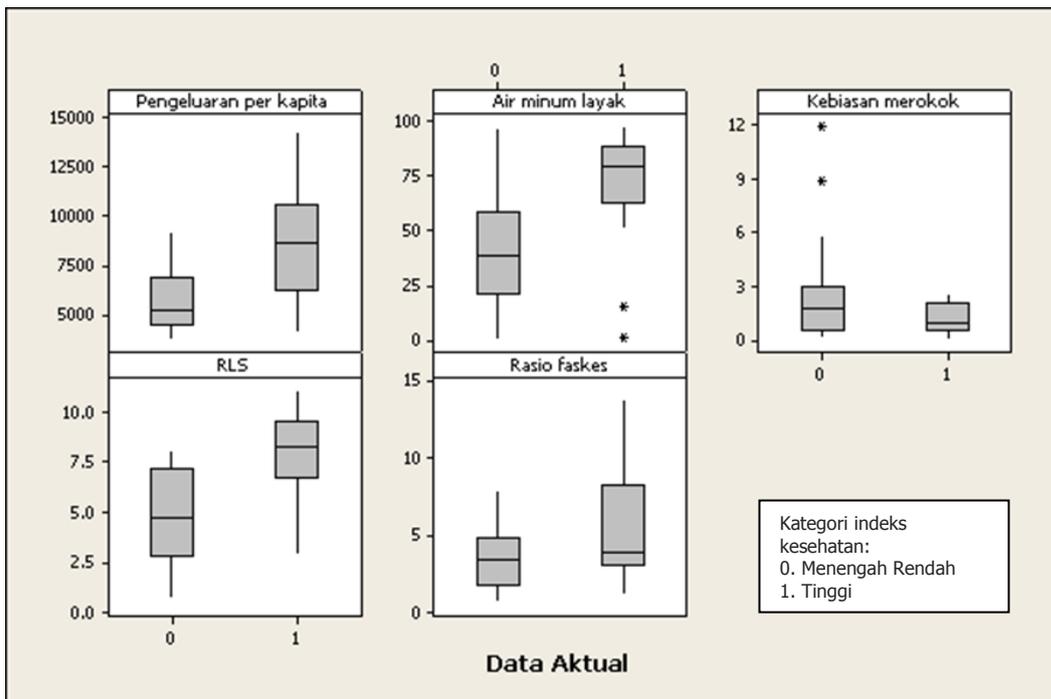


Gambar 4.1 Peta Persebaran Indeks Kesehatan Menurut Kabupaten/Kota di Pulau Papua Tahun 2016

Dalam penelitian ini terdapat 5 variabel prediktor yang diduga mempengaruhi kategori indeks kesehatan yaitu: pengeluaran per kapita, persentase rumah tangga yang memiliki sumber air minum layak, persentase penduduk yang memiliki kebiasaan merokok, rata-rata lama sekolah, serta rasio fasilitas kesehatan terhadap jumlah desa. Karakteristik indeks kesehatan berdasarkan variabel yang diduga mempengaruhinya tersaji pada Gambar 4.2.

Gambar 4.2 memperlihatkan hubungan antara indeks kesehatan di tingkat kabupaten/kota di Pulau Papua dengan masing-masing variabel prediktor. Dari *boxplot* antara indeks kesehatan dan pengeluaran per kapita terlihat bahwa kategori indeks kesehatan tinggi terlihat memiliki nilai median pengeluaran per kapita sebesar 8.671 yang berada di atas median kategori indeks kesehatan menengah rendah yaitu sebesar 5.190. Apabila dilihat dari panjang *whiskers* pada masing-masing kategori yang menunjukkan bahwa range nilai pada variabel pengeluaran per kapita pada kategori menengah rendah tampak lebih sempit dibandingkan dengan kategori tinggi. Sedangkan nilai minimum variabel

pengeluaran pekapita adalah sebesar 3,725 juta rupiah yang terletak di Kabupaten Nduga dan nilai maksimumnya sebesar 14,319 juta rupiah yang terletak di Kota Jayapura. Kemudian dari *boxplot* antara indeks kesehatan dan persentase rumah tangga yang menggunakan air minum yang layak memiliki nilai minimum 0,28 persen yang terletak di Kabupaten Nduga dan nilai maksimumnya sebesar 97,87 persen yang terletak di Kabupaten Mimika. Terlihat bahwa nilai median pada persentase rumah tangga yang menggunakan air minum yang layak dengan kategori tinggi memiliki nilai sebesar 79,23 yang berada di atas dengan kategori menengah rendah dengan median sebesar 38,96. Menariknya, bahwa pada indeks kesehatan kategori tinggi terdapat 2 *outlier*, yang menunjukkan bahwa masih terdapat 2 kabupaten/kota yang tergolong berkategori indeks kesehatan tinggi, namun persentase rumah tangga yang menggunakan air minum layak bernilai rendah yaitu Kabupaten Pegunungan Arfak, Provinsi Papua Barat dan Kabupaten Lanny Jaya, Provinsi Papua.



Gambar 4.2 *Boxplot* Data Aktual Indeks Kesehatan Menurut Variabel Prediktor

Pada *boxplot* antara indeks kesehatan dan persentase penduduk yang memiliki kebiasaan merokok memiliki nilai minimum 0,03 persen yang terletak di Kabupaten Paniai dan nilai maksimumnya sebesar 11,95 persen yang terletak di

Kabupaten Jayawijaya. Terlihat bahwa pada kabupaten/kota dengan indeks kesehatan dengan kategori menengah rendah memiliki nilai median sebesar 1,72 yang berada di atas median kategori tinggi dengan nilai sebesar 0,97. Meskipun nilai mediannya hanya memiliki selisih 0,75 persen, terlihat pada indeks kesehatan menengah rendah terdapat 2 *outlier*, namun masih sejalan dengan teori bahwa semakin tinggi persentase penduduk yang merokok maka semakin rendah indeks kesehatannya, dimana wilayah yang memiliki persentase tertinggi penduduk yang memiliki kebiasaan merokok yaitu Kabupaten Jayawijaya dan Puncak Jaya, Provinsi Papua. Dari *boxplot* antara indeks kesehatan dan rata-rata lama sekolah terlihat bahwa terlihat bahwa nilai median pada indeks kesehatan berkategori tinggi memiliki nilai sebesar 8,26 yang berada di atas indeks kesehatan berkategori menengah rendah dengan nilai sebesar 4,74. Sedangkan nilai minimum variabel rata-rata lama sekolah adalah sebesar 0,70 tahun yang terletak di Kabupaten Nduga dan nilai maksimumnya sebesar 11,95 tahun yang terletak di Kota Jayapura.

Kemudian dari *boxplot* antara indeks kesehatan dan rasio fasilitas kesehatan terhadap jumlah desa memiliki nilai minimum 0,69 yang terletak di Kabupaten Nduga dan nilai maksimumnya sebesar 13,87 yang terletak di Kota Sorong. Terlihat bahwa antar indeks kesehatan kategori menengah rendah dan tinggi memiliki nilai median yang hampir sama yaitu secara berurutan 3,39 dan 3,88. Hal tersebut dimungkinkan bahwa kemudahan akses penduduk ke fasilitas kesehatan sudah relatif seimbang di kabupaten/kota tersebut dan tidak berorientasi pada penambahan jumlah fasilitas untuk meningkatkan derajat kesehatan masyarakatnya tetapi lebih ke peningkatan kualitas pelayanan, seperti peningkatan kualitas tenaga kesehatan, alat kesehatan dan peningkatan kecepatan pelayanan.

4.2.2 Pendeteksian Multikolinearitas pada Variabel Prediktor

Salah satu asumsi dalam analisis regresi adalah tidak diperkenankan terdapat kasus multikolinieritas, yaitu kondisi dengan terdapat hubungan yang erat antar variabel prediktor. Oleh karena itu, sebelum melakukan pemodelan indeks kesehatan dilakukan deteksi multikolinieritas menggunakan nilai *Variance Inflation Factors* (VIF) yang dihitung melalui persamaan (2.13). Berikut adalah

nilai *Variance Inflation Factors* (VIF) dari masing-masing variabel prediktor yang digunakan dalam penelitian ini:

Tabel 4.1 Deteksi Multikolinieritas Variabel Prediktor

Variabel	Nilai VIF
(1)	(2)
X ₁	3,2979
X ₂	2,3740
X ₃	1,0696
X ₄	3,8444
X ₅	1,7705

Sumber: Hasil Olahan Software R Studio

Berdasarkan Tabel 4.1 dapat diketahui bahwa tidak terdapat nilai VIF yang lebih dari 10, hal tersebut menunjukkan bahwa tidak terdapat kasus multikolinieritas pada studi kasus ini. Dengan demikian seluruh variabel prediktor dapat diikutsertakan dalam proses pemodelan selanjutnya.

4.2.3 Uji Dependensi Spasial Indeks Kesehatan di Pulau Papua

Pengujian dependensi spasial pada penelitian ini menggunakan Lagrange Multiplier. Dalam penelitian ini model spasial yang digunakan, dibatasi hanya menggunakan model spasial lag. Sehingga hanya menggunakan Lagrange Multiplier Lag untuk mengetahui dependensi spasial lag pada indeks kesehatan di Pulau Papua. Matriks pembobot spasial (**W**) dapat dilihat pada Lampiran 2.

1) Hipotesis:

$H_0 : \rho = 0$ (tidak ada ketergantungan lag spasial)

$H_1 : \rho \neq 0$ (ada ketergantungan lag spasial)

2) Taraf signifikansi:

$\alpha = 0,05$

3) Statistik uji:

$$LM = \frac{(e^T \mathbf{W} \mathbf{y} / \hat{\sigma}^2)^2}{\mathbf{D}}$$

$$= 5,3539$$

4) Kriteria penolakan:

H_0 ditolak apabila $LM > \chi^2_{(\alpha,1)}$

5) Kesimpulan:

Berdasarkan hasil perhitungan statistik uji yang dapat dilihat nilai outputnya pada Lampiran 4, diperoleh nilai LM-lag adalah sebesar 5,3539 yang lebih besar dari nilai Chi Square dengan derajat bebas satu yaitu 3,8415, sehingga keputusan yang diambil adalah menolak H_0 . Dengan demikian dapat disimpulkan bahwa terdapat dependensi spasial lag dalam memodelkan indeks kesehatan di Pulau Papua pada tahun 2016.

4.2.4 Pengujian Parameter secara Serentak dan Parsial Model Regresi

Probit Spasial

Pengujian parameter bertujuan untuk mengetahui signifikansi dari pengaruh variabel-variabel prediktor terhadap variabel respon. Pengujian parameter pada regresi probit bivariat dilakukan dalam 2 tahap, yaitu pengujian parameter secara serentak dan pengujian parameter secara parsial, dengan model regresi probit spasial yang terbentuk adalah sebagai berikut:

$$[\Gamma X \beta]_i = \beta_0 + \rho \left(\sum_{i=1, j \neq i}^n w_{ij} y_i^* \right) + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i}$$

a. Pengujian parameter secara serentak

Pengujian parameter secara serentak pada regresi probit spasial digunakan untuk menguji peran koefisien β secara keseluruhan.

1) Hipotesis dalam pengujian parameter secara serentak untuk regresi probit spasial adalah sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_5 = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0, \text{ untuk } j = 1, 2, \dots, 5.$$

2) Taraf signifikansi:

$$\alpha = 0,05$$

3) Statistik uji:

$$\begin{aligned} G^2 &= -2 \ln \left[\frac{L(\hat{\omega})}{L(\hat{\Omega})} \right] \\ &= 31,9685 \end{aligned}$$

4) Kriteria penolakan:

H_0 akan ditolak apabila nilai $G^2 > \chi^2_{(\alpha,5)}$

5) Kesimpulan:

Hasil pengujian signifikansi parameter secara serentak menunjukkan bahwa nilai G^2 yang dihasilkan sebesar 31,9685 yang lebih besar dari nilai Chi Square dengan $\alpha = 0,05$ dan derajat bebas 5 yaitu sebesar 11,0705, yang dapat dilihat nilai outputnya pada Lampiran 4, sehingga keputusan yang diambil adalah menolak H_0 . Dengan demikian dapat diartikan bahwa pada regresi probit spasial dengan tingkat kepercayaan 95% minimal ada satu parameter yang signifikan pada model. Oleh karena itu, tahapan pemodelan berikutnya adalah melakukan pengujian parameter secara parsial untuk mengetahui variabel-variabel prediktor yang memiliki pengaruh yang signifikan terhadap variabel respon.

b. Pengujian parameter secara parsial

Pengujian parameter secara parsial dilakukan dengan cara menguji signifikansi dari pengaruh setiap variabel prediktor terhadap variabel respon secara individu.

1) Hipotesis dalam pengujian parameter secara parsial pada pemodelan indeks kesehatan dengan model regresi spasial adalah sebagai berikut:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, \text{ untuk } j = 1, 2, \dots, 5.$$

2) Taraf signifikansi:

$$\alpha = 0,05$$

3) Statistik uji:

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

4) Kriteria penolakan:

H_0 ditolak apabila $Z_{hitung} < -Z_{\alpha/2}$ atau $Z_{hitung} > Z_{\alpha/2}$ atau nilai dari $p\text{-value} < \alpha$

5) Kesimpulan:

Berdasarkan pengujian parameter model secara parsial dengan $\alpha = 0,05$ yang dapat dilihat pada Tabel 4.2, menghasilkan variabel pengeluaran per kapita (X_1), persentase rumah tangga yang memiliki sumber air minum layak (X_2), persentase penduduk yang memiliki kebiasaan merokok (X_3), rata-rata lama sekolah (X_4), serta rasio fasilitas kesehatan terhadap jumlah desa (X_5) berpengaruh signifikan terhadap kategori indeks kesehatan kabupaten/kota.

Tabel 4.2 Pengujian Signifikansi Parameter Secara Parsial Model Probit Spasial

Variabel	Coeffisien	Std. Error	Z-value	P-value	Keputusan
(1)	(2)	(3)	(4)	(5)	(6)
Intercept	0,00256	0,00037	6,83740	0,00000	Tolak Ho
X_1	-0,00015	0,00003	-5,10397	0,00001	Tolak Ho
X_2	0,01866	0,00355	5,25102	0,00001	Tolak Ho
X_3	-0,02808	0,00345	-8,14187	0,00000	Tolak Ho
X_4	0,01605	0,00030	53,17421	0,00000	Tolak Ho
X_5	0,00811	0,00130	6,26322	0,00000	Tolak Ho

Sumber: Hasil Olahan Software R Studio

4.2.5 Interpretasi Model Regresi Probit Spasial

Langkah awal yang harus dilakukan untuk memodelkan klasifikasi indeks kesehatan dengan menggunakan regresi probit spasial adalah melakukan pendugaan nilai parameter model probit spasial. Berikut adalah hasil pendugaan parameter *spatial autoregressive* (SAR) probit yang disajikan pada Tabel 4.2.

Berdasarkan Tabel 4.2, dengan menggunakan $\alpha = 0,05$ dapat diketahui bahwa variabel yang berpengaruh secara signifikan pada model adalah pengeluaran per kapita (X_1), persentase rumah tangga yang memiliki sumber air minum layak (X_2), persentase penduduk yang memiliki kebiasaan merokok (X_3), rata-rata lama sekolah (X_4), dan rasio fasilitas kesehatan terhadap jumlah desa (X_5). Berikut adalah model yang dihasilkan dari pemodelan indeks kesehatan dengan menggunakan SAR probit:

$$[\mathbf{r}\mathbf{x}\hat{\boldsymbol{\beta}}]_i = 0,00256 + 0,6866 \left(\sum_{i=1, j \neq i}^{42} w_{ij} y_i^* \right) - 0,00015X_{1i} + 0,01866X_{2i} + \\ - 0,02808X_{3i} + 0,01605X_{4i} + 0,00811X_{5i}$$

Setelah mendapatkan model, maka langkah selanjutnya adalah menghitung nilai prediksi probabilitasnya dengan menggunakan persamaan (2.27) dan hasil perhitungan nilai prediksi probabilitas secara lengkap terlampir pada Lampiran 6. Dari perhitungan nilai prediksi probabilitas ini, selanjutnya dapat dihitung prediksi dari klasifikasi aktual indeks kesehatan yang selanjutnya dapat digunakan untuk membuat tabel *confusion matrix* atau ketepatan klasifikasi. Kemudian dari tabel ketepatan klasifikasi ini akan didapatkan nilai akurasi, sensitivitas dan spesifitas dari model, yang dapat dilihat pada subbab 4.2.6.

Prediksi probabilitas untuk pengamatan $y_i = 1$ pada model indeks kesehatan untuk wilayah kabupaten/kota di Pulau Papua adalah sebagai berikut.

$$\hat{P}(y_i = 1|X) = \hat{P}(y_i^* \geq 0|X) = 1 - \Phi \left(\frac{-[(\mathbf{I} - 0,6866\mathbf{W})^{-1}\mathbf{x}\hat{\boldsymbol{\beta}}]_i}{\hat{\Omega}_{ii}} \right) \\ = \Phi \left(\frac{[(\mathbf{I} - 0,6866\mathbf{W})^{-1}\mathbf{x}\hat{\boldsymbol{\beta}}]_i}{\hat{\Omega}_{ii}} \right)$$

dimana:

$$\mathbf{\Gamma}_{42 \times 42} = (\mathbf{I} - \hat{\rho}\mathbf{W})^{-1} \\ = \left(\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} - 0,6866 \begin{bmatrix} 0 & 0,50 & 0 & \dots & 0 \\ 0,17 & 0 & 0,17 & \dots & 0 \\ 0 & 0,25 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \right)^{-1} \\ \mathbf{\Gamma}_{42 \times 42} = \left(\begin{bmatrix} 1 & -0,34 & 0 & \dots & 0 \\ 0,11 & 1 & 0,11 & \dots & 0 \\ 0 & 0,17 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 1,10 & 0,48 & 0 & \dots & 0,00 \\ 0,16 & 1,18 & 0,20 & \dots & 0,00 \\ 0,06 & 0,30 & 1,14 & \dots & 0,00 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0,00 & 0,00 & 0,00 & \dots & 1,19 \end{bmatrix} \\ \mathbf{X}_{42 \times (5+1)} = \begin{bmatrix} 1 & 6935 & 81,62 & \dots & 3,98 \\ 1 & 7538 & 69,12 & \dots & 6,63 \\ 1 & 7434 & 20,03 & \dots & 5,71 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 14319 & 96,98 & \dots & 10,26 \end{bmatrix}$$

$$\hat{\beta}_{(5+1) \times 1} = \begin{bmatrix} 0,00256 \\ -0,00015 \\ 0,01866 \\ \vdots \\ 0,00811 \end{bmatrix}$$

$$\hat{\Omega}_{42 \times 42} = [(\mathbf{I} - \hat{\rho}\mathbf{W})^T(\mathbf{I} - \hat{\rho}\mathbf{W})]^{-1}$$

$$= \begin{bmatrix} 1,80 & 0,99 & 0,64 & \cdots & 0,01 \\ 0,99 & 1,71 & 0,82 & \cdots & 0,02 \\ 0,64 & 0,82 & 1,70 & \cdots & 0,01 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0,01 & 0,02 & 0,01 & \cdots & 2,13 \end{bmatrix}$$

dengan diagonal utama dari $\hat{\Omega}_{42 \times 42} = \begin{bmatrix} 1,80 \\ 1,71 \\ 1,70 \\ \vdots \\ 2,13 \end{bmatrix}$ dan nilai $(\mathbf{\Gamma}\mathbf{X}\hat{\beta})_{42 \times 1} = \begin{bmatrix} 0,70 \\ 0,43 \\ -0,43 \\ \vdots \\ 0,004 \end{bmatrix}$

Sebagai contoh, jika dimisalkan ingin dicari nilai prediksi probabilitas Kabupaten Jayawijaya, dimana Kabupaten Jayawijaya berada pada $i = 15$. Maka probabilitas $y_i = 1$ pada model indeks kesehatan untuk wilayah Kabupaten Jayawijaya adalah sebagai berikut:

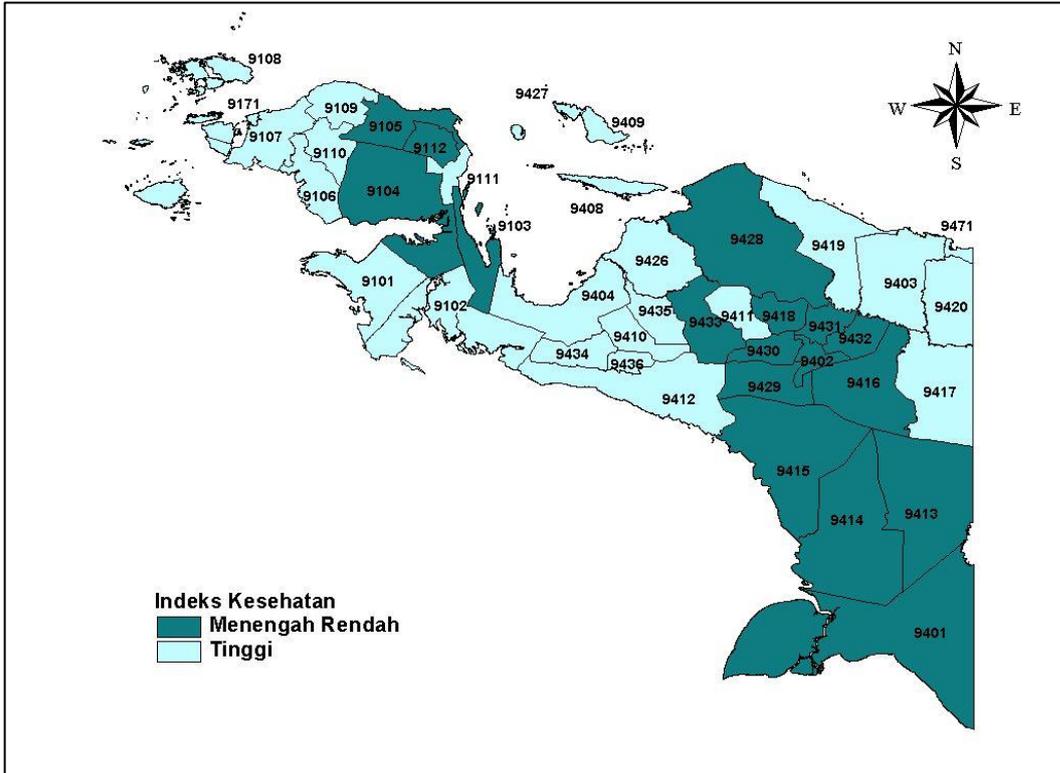
$$\hat{P}(y_{15} = 1 | X_{15}) = \Phi \left(\frac{[(\mathbf{I} - 0,6866\mathbf{W})^{-1}\mathbf{X}\hat{\beta}]_{15}}{\hat{\Omega}_{15,15}} \right)$$

$$= 0,01860$$

Nilai prediksi probabilitas untuk Kabupaten Jayawijaya 0,01860 memiliki makna bahwa peluang Kabupaten Jayawijaya untuk dapat tergolong menjadi kategori indeks kesehatan tinggi adalah sebesar 1,86 persen, sedangkan peluang Kabupaten Jayawijaya untuk dapat tergolong menjadi kategori indeks kesehatan menengah rendah adalah sebesar 98,14 persen Berdasarkan perhitungan nilai probabilitas Kabupaten Jayawijaya tersebut, menunjukkan bahwa besarnya nilai indeks kesehatan di Kabupaten Jayawijaya selain dipengaruhi oleh pengeluaran per kapita, persentase rumah tangga yang memiliki sumber air minum layak, persentase penduduk yang memiliki kebiasaan merokok, rata-rata lama sekolah, dan rasio fasilitas kesehatan terhadap jumlah desa, juga dipengaruhi oleh kabupaten/kota yang berada di sekitar Kabupaten Jayawijaya, antara lain Kabupaten Yahukimo, Kabupaten Tolikara, Kabupaten Nduga, Kabupaten Lanny Jaya, Kabupaten Mamberamo Tengah dan Kabupaten Yalimo, dimana wilayah

tersebut sama-sama memiliki kategori indeks kesehatan menengah rendah menurut prediksi probabilitasnya.

Pada Gambar 4.3 disajikan data prediksi klasifikasi indeks kesehatan yang dihasilkan oleh model SAR probit. Berdasarkan hasil dari model SAR probit, kabupaten/kota yang tergolong dalam kategori tinggi sebesar 25 kabupaten/kota, sedangkan untuk kategori menengah rendah sebesar 17 kabupaten/kota.



Gambar 4.3 Peta Hasil Prediksi Indeks Kesehatan dari Model Regresi Spasial

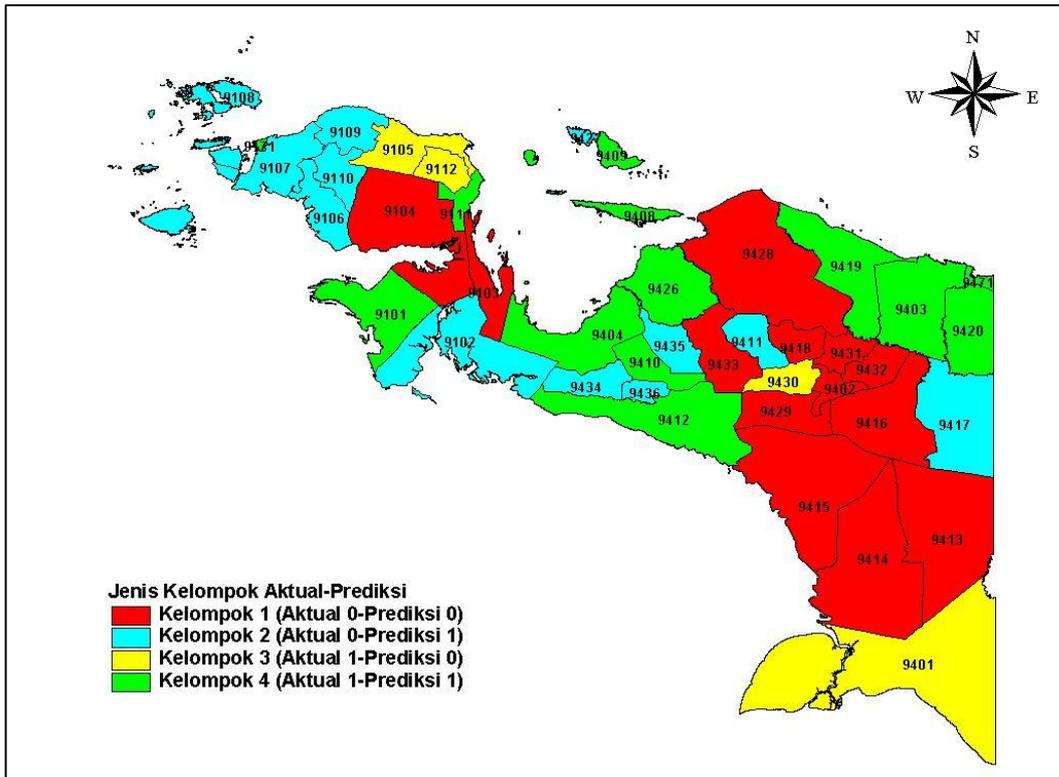
Penjelasan yang lebih rinci untuk kabupaten/kota antara data aktual dan data prediksi akan ditampilkan pada Tabel 4.3. Sebagai contoh, Kabupaten Merauke pada data aktual tergolong dalam indeks kesehatan kategori tinggi, namun ketika dimasukkan ke dalam model SAR probit diprediksi sebagai kabupaten/kota dengan indeks kesehatan kategori menengah rendah, hal yang sama juga berlaku untuk 3 kabupaten/kota lainnya. Contoh lainnya adalah Kabupaten Sorong yang pada data aktual tergolong dalam indeks kesehatan menengah rendah ketika dimasukkan ke dalam model SAR probit diprediksi menjadi kabupaten/kota dengan indeks kesehatan kategori tinggi, hal yang sama juga dialami oleh 11 kabupaten/kota lainnya. Sedangkan sisanya yaitu sebesar 26

kabupaten/kota tidak mengalami perubahan kategori ketika dimasukkan ke dalam model SAR probit.

Tabel 4.3 Hasil Pengelompokan Data Aktual dan Prediksi Berdasarkan Kabupaten/Kota di Pulau Papua dengan Model Probit Spasial

Aktual	Prediksi	Kabupaten/Kota
(1)	(2)	(3)
0	0	Teluk Wondama, Teluk Bintuni, Jayawijaya, Boven Digoel, Mappi, Asmat, Yahukimo, Mamberamo Raya, Nduga, Tolikara, Mamberamo Tengah, Yalimo, Puncak
0	1	Kaimana, Sorong Selatan, Sorong, Raja Empat, Tambrau, Maybrat, Puncak Jaya, Pegunungan Bintang, Supriori, Dogiyai, Intan Jaya, Deiyai
1	0	Manokwari, Pegunungan Arfak, Merauke, Lanny Jaya
1	1	Fak fak, Manokwari Selatan, Kota Sorong, Jayapura, Nabire, Yapen Waropen, Biak Numfor, Paniai, Mimika, Sarmi, Keerom, Waropen, Kota Jayapura

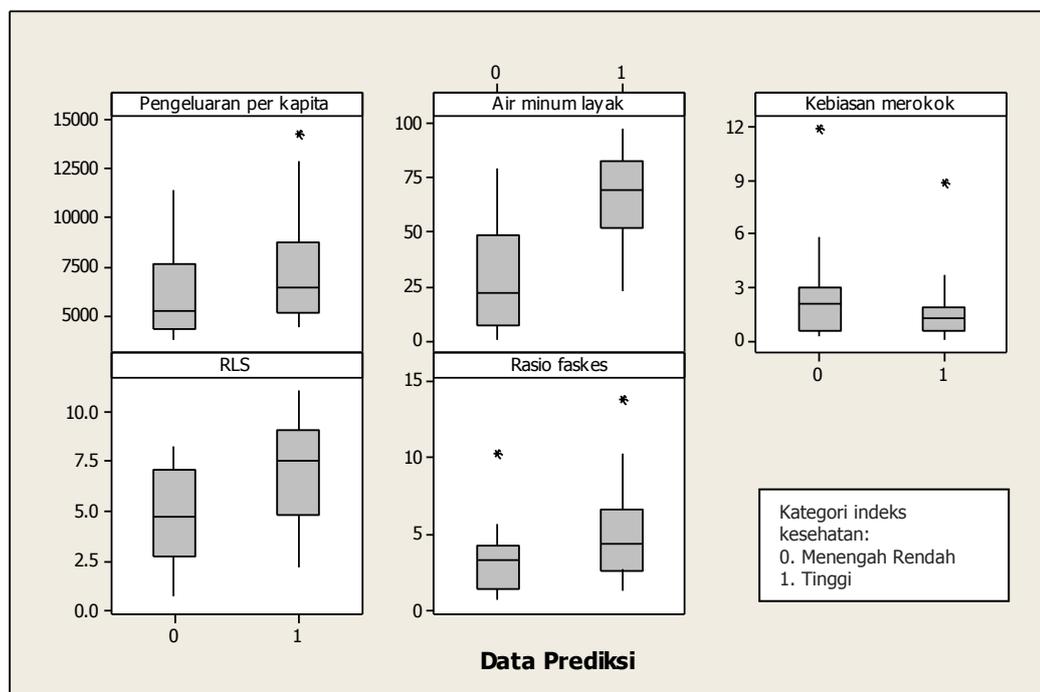
Secara visual, Tabel 4.3 dapat diubah menjadi peta menurut jenis kelompok data aktual dan data prediksi yang ditampilkan pada Gambar 4.4. Apabila dilihat lebih dalam yaitu dari sisi wilayah menurut provinsinya, terlihat bahwa untuk Provinsi Papua sebanyak 72,41 persen telah memiliki ketepatan klasifikasi, dan sisanya yaitu sebesar 27,59 persen tidak tepat dalam klasifikasi. Sebagian besar wilayah pegunungan tengah dan pesisir utara telah tepat dalam melakukan klasifikasi. Berbeda dengan Provinsi Papua, sebaliknya Provinsi Papua Barat ketepatan klasifikasinya hanya sebesar 38,46 persen dan kesalahan klasifikasi sebesar 61,54 persen, yang berarti sebagian besar terjadi ketidaktepatan klasifikasi di Provinsi Papua Barat. Hasil pengukuran ketepatan klasifikasi dari kedua provinsi tersebut, dapat disimpulkan bahwa ketepatan klasifikasi di Provinsi Papua lebih baik dari pada ketepatan klasifikasi di Provinsi Papua Barat.



Gambar 4.4 Peta Pengelompokan Data Aktual dan Prediksi Indeks Kesehatan

Selanjutnya, berdasarkan data hasil prediksi dengan model SAR probit dapat ditampilkan *bloxplot* yang disajikan pada Gambar 4.5, sehingga dapat dibandingkan dengan *bloxplot* data aktual pada Gambar 4.2. Berdasarkan *boxplot* pada Gambar 4.5, hubungan antara indeks kesehatan data prediksi dengan pengeluaran per kapita terlihat bahwa indeks kesehatan kategori tinggi terlihat memiliki nilai median pengeluaran per kapita sebesar 6.417 juta rupiah yang berada di atas median kategori indeks kesehatan menengah rendah yaitu sebesar 5.181 juta rupiah. Jika dibandingkan dengan nilai median pada data aktual, nilai median data prediksi indeks kesehatan kategori tinggi berada di bawah median data aktual dengan selisih 2.254 juta rupiah. Selain itu pada data prediksi ditemukan satu *outlier* pada indeks kesehatan berkategori tinggi yaitu berada pada Kota Jayapura, sedangkan pada data aktual tidak ditemukan *outlier*. Hal ini disebabkan karena perubahan kategori yang terjadi setelah pemodelan dengan menggunakan SAR probit, dimana terdapat 12 kabupaten/kota yang mengalami perubahan kategori dari indeks berkategori tinggi pada data aktual menjadi

berkategori menengah rendah pada data prediksi, sehingga panjang *whisker* pada kategori tinggi mengalami penyusutan.



Gambar 4.5 *Boxplot* Data Prediksi Indeks Kesehatan Menurut Variabel Prediktor

Kemudian dari *boxplot* antara indeks kesehatan dengan persentase rumah tangga yang memiliki sumber air minum layak pada data prediksi terlihat bahwa nilai median pada indeks kesehatan berkategori tinggi memiliki nilai sebesar 69,12 persen yang berada di atas indeks kesehatan berkategori menengah rendah dengan nilai sebesar 22,27 persen. Berbeda dengan data aktual yang mengandung *outlier* pada kategori tinggi, persentase rumah tangga yang memiliki sumber air minum layak pada data prediksi tidak mengandung *outlier* baik pada kategori indeks kesehatan tinggi maupun menengah rendah. Dari *boxplot* tersebut, dapat dilihat hubungan indeks kesehatan pada data prediksi dengan persentase penduduk yang memiliki kebiasaan merokok dimana median pada indeks kesehatan berkategori menengah rendah berada di atas indeks kesehatan berkategori tinggi, dengan nilai secara berurutan sebesar 2,07 persen dan 1,27 persen. Nilai tersebut memiliki sedikit peningkatan dibanding dengan median pada data aktual. Terjadi perpindahan *outlier* dari data aktual ke data prediksi, dimana pada data aktual terdapat dua *outlier* yang berada pada kategori

menengah rendah. Sedangkan pada data prediksi masing-masing kategori memiliki satu *outlier*. Hal ini disebabkan karena perpindahan kategori Kabupaten Puncak Jaya, yang semula berkategori menengah rendah menjadi kategori indeks kesehatan tinggi.

Boxplot pada Gambar 4.5 menunjukkan nilai median pada hubungan antara indeks kesehatan pada data prediksi dengan rata-rata lama sekolah sebesar 7,57 tahun pada kategori tinggi dan 4,74 tahun pada kategori menengah rendah, yang menunjukkan bahwa kategori tinggi berada di atas kategori menengah rendah pada data prediksi. Menariknya, tidak ada perubahan nilai median pada kategori menengah rendah dari aktual ke prediksi dimana memiliki nilai yang sama yaitu sebesar 7,47 tahun. Sedangkan pada kategori indeks kesehatan tinggi, nilai Q1 mengalami perubahan, pada data aktual bernilai 6,78 tahun menjadi 4,80 tahun, sedangkan Q3 tidak mengalami perubahan yang signifikan. Hal ini menunjukkan *box* atau *inner quartile range* mengalami perpanjangan pada kategori indeks kesehatan tinggi.

Selanjutnya dari *boxplot* antara indeks kesehatan dengan rasio fasilitas kesehatan terhadap jumlah desa pada data prediksi terlihat bahwa terlihat bahwa nilai median pada indeks kesehatan berkategori tinggi memiliki nilai sebesar 4,38 yang berada di atas indeks kesehatan berkategori menengah rendah dengan nilai sebesar 3,27. Bila dibandingkan dengan nilai median pada data aktual, median data prediksi pada kategori indeks kesehatan tinggi berada di atas median data aktual. Sedangkan untuk kategori indeks kesehatan menengah rendah, mediannya tidak jauh berbeda. Perbedaan yang mencolok antara data aktual dan data prediksi dapat dilihat dari munculnya *outlier* pada data prediksi, sedangkan pada data aktual tidak ditemukan *outlier*. Kota sorong menjadi *outlier* pada indeks kesehatan berkategori tinggi, sedangkan Kabupaten Merauke menjadi *outlier* pada indeks kesehatan berkategori menengah rendah.

Untuk mengetahui besarnya pengaruh perubahan dari masing-masing variabel prediktor terhadap variabel respon dengan asumsi variabel lainnya konstan dapat dilihat berdasarkan nilai efek marginalnya. Berdasarkan persamaan (2.29) diperoleh hasil perhitungan nilai efek marjinal yang ditunjukkan pada

Lampiran 7 s.d. 11. Sebagai contoh, nilai efek marginal dari pengeluaran per kapita (X_1) untuk Kota Jayapura sebesar -0,00010, hal ini dapat diartikan perubahan pengeluaran perkapita sebesar 1 satuan (1 juta rupiah) akan menurunkan 0,01 persen probabilitas kabupaten tersebut untuk mencapai indeks kesehatan berkategori tinggi. Hubungan antara variabel pengeluaran per kapita terhadap indeks kesehatan menunjukkan hubungan yang negatif, dimana semakin meningkat nilai pengeluaran perkapita, maka semakin menurun nilai indeks kesehatan.

Hal ini dapat dimaklumi dan sesuai dengan pendapat Todaro dan Smith (2012) yang menyatakan bahwa semakin tinggi pendapatan atau pengeluaran belum menjamin menurunnya tingkat morbiditas atau angka kesakitan, terutama pada negara berkembang, dikarenakan banyak keluarga di negara berkembang pengeluaran justru dialihkan dari makanan bergizi kepada makan yang tidak bergizi, seperti *junk food* yang mungkin dianggap modern dan merupakan simbol kesuksesan ekonomi. Namun, dalam mengartikan intepretasi tersebut bukan berarti untuk meningkatkan indeks kesehatan harus mengurangi besaran nilai pengeluaran perkapita, akan tetapi untuk meningkatkan indeks kesehatan yang pertama kali dilakukan adalah dengan merubah pola kebiasaan masyarakat terlebih dahulu, yaitu konsumsi yang dilakukan pada pengeluaran perkapita agar diarahkan untuk meningkatkan aspek kesehatan masyarakat itu sendiri, misalnya kebiasaan konsumsi makanan *junk food* dialihkan menjadi konsumsi makanan yang bergizi atau pengurangan konsumsi minuman beralkohol oleh masyarakat.

Kemudian nilai efek marginal dari persentase rumah tangga yang memiliki sumber air minum layak (X_2) untuk Kota Jayapura sebesar 0,01268, hal ini dapat diartikan perubahan persentase rumah tangga yang memiliki sumber air minum layak sebesar 1 satuan (1 persen) akan meningkatkan 1,27 persen probabilitas kabupaten tersebut untuk mencapai indeks kesehatan berkategori tinggi. Pada variabel persentase penduduk yang memiliki kebiasaan merokok (X_3), nilai efek marginal untuk Kota Jayapura sebesar -0,01908 yang berarti perubahan perubahan persentase penduduk yang memiliki kebiasaan merokok sebesar 1 satuan (1 persen) akan menurunkan 1,91 persen probabilitas kabupaten tersebut untuk

mencapai indeks kesehatan berkategori tinggi. Selanjutnya, nilai efek marjinal dari rata-rata lama sekolah (X_4) untuk Kota Jayapura sebesar 0,01090 yang berarti perubahan perubahan rata-rata lama sekolah sebesar 1 tahun akan meningkatkan 1,09 persen probabilitas kabupaten tersebut untuk mencapai indeks kesehatan berkategori tinggi. Kemudian variabel rasio fasilitas kesehatan terhadap jumlah desa (X_5) untuk Kota Jayapura memiliki nilai efek marjinal sebesar 0,00551 yang berarti perubahan perubahan rasio fasilitas kesehatan terhadap jumlah desa sebesar 1 satuan akan meningkatkan 0,55 persen probabilitas kabupaten tersebut untuk mencapai indeks kesehatan berkategori tinggi.

4.2.6 Kriteria Keباikan Model Regresi Probit Spasial

Tabel Ketepatan klasifikasi atau *confusion matrix* digunakan untuk menggambarkan ukuran ketepatan antara prediksi dan observasi (kenyataan lapangan/hasil penelitian). Melalui tabel ini dapat dihasilkan 3 ukuran yang umum untuk digunakan untuk menilai kebaikan suatu model dalam melakukan pengklasifikasian, yaitu akurasi, sensitivitas dan spesifitas. Tabel *confusion matrix* pada penelitian ini secara lengkap disajikan pada Tabel 4.4.

Tabel 4.4 Tabel *Confusion Matrix* Indeks Kesehatan Di Pulau Papua Tahun 2016

Hasil Aktual	Hasil Prediksi	
	<i>P</i>	<i>N</i>
(1)	(2)	(3)
<i>P</i>	<i>True Positive</i> (13)	<i>False Negative</i> (4)
<i>N</i>	<i>False Positive</i> (12)	<i>True Negative</i> (13)

Sumber: Hasil Olahan Software R Studio

Berdasarkan Tabel 4.4 dapat dihitung nilai akurasi, sensitivitas dan spesifitasnya dengan cara sebagai berikut:

$$\begin{aligned}
 Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \\
 &= \frac{13 + 13}{13 + 13 + 12 + 4} \times 100\% = 61,90\%
 \end{aligned}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%$$

$$= \frac{13}{13 + 4} \times 100\% = 76,47\%$$

$$Specivisity = \frac{TN}{TN + FP} \times 100\%$$

$$= \frac{13}{13 + 12} \times 100\% = 52,00\%$$

Dari hasil tersebut atau output pada Lampiran 5, didapatkan nilai akurasi pada model adalah sebesar 61,90 persen, yang berarti kemampuan model untuk mendeteksi secara benar seluruh objek yang diuji adalah sebesar 61,90 persen. Nilai sensitivitas yang dihasilkan dari model adalah sebesar 76,47 persen, yang berarti proporsi indeks kesehatan yang benar-benar berkategori tinggi yang juga diidentifikasi oleh model sebagai objek yang memiliki indeks kesehatan berkategori tinggi sebesar 76,47 persen. Sedangkan nilai spesifitas yang dihasilkan dari model adalah sebesar 52,00 persen, yang berarti proporsi indeks kesehatan yang benar-benar berkategori menengah rendah yang juga diidentifikasi oleh model sebagai objek yang memiliki indeks kesehatan berkategori menengah rendah sebesar 52,00 persen.

(Halaman ini sengaja dikosongkan)

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah diuraikan sebelumnya, terdapat beberapa kesimpulan yang diperoleh sebagai berikut:

1. Metode estimasi yang digunakan dalam penelitian ini untuk mendapatkan parameter regresi probit spasial adalah menggunakan metode *Maximum Likelihood Estimator* (MLE), dimana dengan menggunakan metode ini dihasilkan estimasi yang tidak *closed form*, Sehingga untuk mengatasi hal tersebut, digunakan metode simulasi *Recursive Importance Sampling* (RIS) untuk memperoleh estimator parameter,
2. Pemodelan regresi probit spasial pada penelitian ini menunjukkan bahwa pengeluaran per kapita (X_1), persentase rumah tangga yang memiliki sumber air minum layak (X_2), persentase penduduk yang memiliki kebiasaan merokok (X_3), rata-rata lama sekolah (X_4), serta rasio fasilitas kesehatan terhadap jumlah desa (X_5) secara signifikan mempengaruhi klasifikasi indeks kesehatan di Pulau Papua pada Tahun 2016 dan tingkat akurasi yang dihasilkan dari pemodelan ini adalah sebesar 61,90 persen,

5.2 Saran

Berdasarkan penelitian yang telah dilakukan, saran yang dapat diberikan adalah sebagai berikut:

1. Pada penelitian ini kategori yang digunakan dibatasi hanya 2 kategori atau biner, pada penelitian selanjutnya dapat menggunakan regresi probit dengan pembagian kategori yang lebih dari dua.
2. Penelitian ini masih menggunakan matriks pembobot *queen contiguity*, sehingga pada penelitian selanjutnya dapat dikembangkan menggunakan matriks pembobot yang lain, misalnya dengan matriks pembobot *customize* yang telah memperhitungkan jalur transportasi darat, udara dan air antar kabupaten/kota

3. Sebagai pengembangan model probit spasial, penelitian selanjutnya dapat menggunakan juga model spasial yang lain selain *spatial autoregressive* model, seperti *spatial error* model.

DAFTAR PUSTAKA

- AHMAC (2012), *Aboriginal and Torres Strait Islander Health Performance Framework 2012*, AHMAC, Canberra.
- Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- Arbia, G. (2014), *A Primer for Spatial Econometrics With Application in R*, Palgrave Macmillan, Hampshire, England.
- Beron, K.J. dan Vijverberg, W.P.M. (2004), "Probit in a Spatial Context: A Monte Carlo Analysis", dalam *Advances in Spatial Econometrics. Methodology, Tools and Applications*, eds. L. Anselin, R. Florax dan J. Sergio, Springer, Berlin, pp. 169-195.
- BPS (2015), *Indeks Pembangunan Manusia 2015*, Badan Pusat Statistik, Jakarta.
- Calabrese, R. dan Elkind, J.A. (2014), "Estimator of Binary Spatial Autoregressive Models: A Monte Carlo Study", *Journal of Regional Science*, Vol. 54, pp. 664-687.
- Case, A. (1992), "Neighborhood Influence and Technological Change", *Regional Science and Urban Economics*, Vol. 22, pp. 491-508.
- Fahmi, E.F.F.E. (2016), *Model Regresi Spasial Pada Indeks Pembangunan Manusia (IPM) di Jawa Timur*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Fawcett, T. (2006), "An introduction to ROC analysis", *Pattern Recognition Letters*, Vol. 27, No. 8, pp. 861–874.
- Franzese, R.J. dan Hays, J.C. (2007), "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data", *Political Analysis*, Vol. 15(2), pp. 140–164.
- Greenberg, B.G. (1980), "Chester I. Bliss, 1899-1979", *International Statistical Review*, Vol. 8, pp. 135-136.
- Greene, W.H. (2008), *Econometrics Analysis, Sixth Edition*, Prentice Hall, New Jersey.
- Gujarati, D.N. (2004), *Basic Econometric, Fourth Edition*, The Macgraw-Hill, New Jersey.

- Handayani, S.T. (2014), Faktor-Faktor Yang Mempengaruhi Derajat Kesehatan Masyarakat (Faktor Ekonomi, Pendidikan Ibu, Lingkungan, dan Pelayanan Kesehatan) Studi Kasus di Aceh, Papua, Bangka Belitung, Sulawesi Utara, Jawa Tengah, dan Daerah Istimewa Yogyakarta, Tesis, Universitas Gajah Mada, Yogyakarta.
- HMPD (2016), MDGs to SDGs. [Online] diakses dari: <http://hmpd.fk.ub.ac.id/mdgs-to-sgds/> [pada 17 Mei 2017].
- Irevanie, R.S. (2017), Perbandingan Metode Quantile Regression (QR) dan Geographically Weighted Regression (GWR) Pada Data Angka Harapan Hidup di Indonesia, Tesis, Insitut Teknologi Sepuluh Nopember, Surabaya.
- Juliandari, N.N.T. (2014), Pemodelan Angka Harapan Hidup dan Angka Kematian Bayi Di Jawa Timur Dengan Pendekatan Regresi Nonparametrik Spline Birespon, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Lagarias, J.C., Reeds, J.A., Wright, M.H. dan Wright, P.E. (1998), "Convergence Properties of The Nelder Mead Simplex Method in Low Dimensions", The Society for Industrial and Applied Mathematics, Vol. Vol. 9, No. 1, pp. 112-147.
- Lee, J. dan Wong, D.W. (2001), Statistical analysis with ArcView GIS, John Wiley & Sons, Inc., New Jersey.
- LeSage, J.P. (1999), The Theory and Practice of Spatial Econometrics, Department of Economics University of Toledo, Toledo.
- LeSage, J.P. (2000), "Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models", Geographical Analysis, Vol. 32(1), pp. 19-35.
- Marisa, I. (2017), Teori Hendrik L. Blum. [Online] diakses dari: <http://info-kesehatan-bidan-irma-marisa.blogspot.co.id/2015/06/teori-hendrik-l-blum.html> [pada 12 Juni 2017].
- Martinetti, D. dan Geniaux, G. (2017), "Approximate likelihood estimation of spatial probit models", Regional Science and Urban Economics, Vol. 64, pp. 30-45.
- Maully, A.V. (2014), Faktor-faktor yang Mempengaruhi Indeks Kesehatan Kabupaten dan Kota di Provinsi Jawa Timur , Tugas Akhir, Institut Teknologi Sepuluh Nopember, Surabaya.

- McMillen, D.P. (1992), "Probit with Spatial Autocorrelation", *Journal of Regional Science*, Vol. 32(3), pp. 335–348.
- OECD (2015), *Health at a Glance 2015: OECD Indicators*, OECD, Paris.
- Pinkse, J. dan Slade, M.E. (1998), "Contracting in Space: An Application of Spatial Statistics to Discrete Choice Models", *Journal of Econometrics*, Vol. 85, pp. 125–154.
- Puspita, F.I. (2013), *Model Probit Spasial Pada Faktor-Faktor yang Mempengaruhi Klasifikasi IPM di Pulau Jawa*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Ratnasari, V. (2012), *Estimasi Parameter dan Uji Signifikansi Model Probit Bivariat*, Disertasi, Institut Teknologi Sepuluh Nopember, Surabaya.
- Satomo, P.D. (2017), *Pemodelan Indeks Kesehatan dan Indeks Pengeluaran Kabupaten/Kota di Indonesia Dengan Pendekatan Model Probit Bivariat*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Sugiantari, A.P. (2013), *Analisis Faktor-faktor yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Spline*, Tesis, Institut Teknologi Sepuluh November, Surabaya.
- Todaro, M.P. dan Smith, S.C. (2012), *Economic Development*, Eleventh Edition ed., Pearson, Boston.
- UNDP (2017), *Sustainable Development Goals: Goals 3 Good Health and Well-being*. [Online] diakses dari: <http://www.undp.org/content/undp/en/home/sustainable-development-goals/goal-3-good-health-and-well-being/targets/> [pada 16 Mei 2017].
- Verbeek, M. (2004), *A Guide to Modern Econometric*, John Wiley & Son, Ltd, London.
- WHO (2017), *The Determinants of Health*. [Online] diakses dari: <http://www.who.int/hia/evidence/doh/en/> [pada 18 Mei 2017].
- Winarno, D. (2009), *Analisis Angka Kematian Bayi di Jawa Timur Dengan Pendekatan Regresi Spasial*, Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Data Variabel Penelitian

No.	Nama Kab/Kota	y_i	X_1	X_2	X_3	X_4	X_5
1	Fak Fak	1	6.935	81,62	1,55	8,22	3,98
2	Kaimana	0	7.538	69,12	3,67	7,83	6,63
3	Teluk Wondama	0	7.434	20,03	0,51	6,57	5,71
4	Teluk Bintuni	0	9.208	58,54	2,37	7,57	2,52
5	Manokwari	1	11.440	72,34	2,53	7,85	3,31
6	Sorong Selatan	0	5.644	58,14	1,66	6,95	4,38
7	Sorong	0	6.563	48,06	1,98	7,57	4,89
8	Raja Ampat	0	7.393	61,34	2,54	7,53	6,69
9	Tambrauw	0	4.561	54,53	1,83	4,7	4,08
10	Maybrat	0	4.692	33,52	1,46	6,33	2,1
11	Manokwari Selatan	1	4.702	51,46	0,62	6,32	2,46
12	Pegunungan Arfak	1	4.594	15,01	0,66	4,9	1,17
13	Kota Sorong	1	12.858	95,92	2,37	10,91	13,87
14	Merauke	1	10.016	79,23	2,07	8,26	10,24
15	Jayawijaya	0	7.282	38,96	11,95	4,74	1,66
16	Jayapura	1	9.653	70,83	2,38	9,53	5,07
17	Nabire	1	8.779	79,35	1,27	9,48	8,43
18	Yapen Waropen	1	7.414	81,94	0,32	8,81	3,88
∴	∴	∴	∴	∴	∴	∴	∴
26	Yahukimo	0	4.248	37,82	2,09	3,99	2,2
27	Pegunungan Bintang	0	5.289	44,73	1,72	2,19	1,26
28	Tolikara	0	4.711	9,86	0,7	3,21	1,16
29	Sarmi	1	6.417	76,53	0,95	8,08	2,82
30	Keerom	1	8.671	77,57	0,88	7,24	8,2
31	Waropen	1	6.270	93,62	0,37	8,66	3,63
32	Supiori	0	5.379	44,66	1,52	8,13	7,89
33	Membramo Raya	0	4.387	16,16	2,83	4,89	4,58
34	Nduga	0	3.725	0,28	5,8	0,7	0,69
35	Lanny Jaya	1	4.106	1,51	2,03	2,92	3,57
36	Mamberamo Tengah	0	4.219	0,75	0,24	2,57	3,39
37	Yalimo	0	4.435	27,37	0,63	2,19	1,01
38	Puncak	0	5.181	22,27	0,26	1,78	2,75
39	Dogiyai	0	5.190	51,93	1,2	4,89	4,81
40	Intan Jaya	0	5.038	22,67	0,28	2,49	1,79
41	Deiyai	0	4.383	67,19	0,08	2,97	4,67
42	Kota Jayapura	1	14.319	96,98	1,09	11,14	10,26

Lampiran 2. Matriks Pembobot (*Queen Contiguity*)

Kode	9101	9102	9103	9104	9105	9106	9107	9108	9109	9110	9111	9112	9171	9401
9101	0	1	0	1	0	0	0	0	0	0	0	0	0	0
9102	1	0	1	1	0	0	0	0	0	0	0	0	0	0
9103	0	1	0	1	0	0	0	0	0	0	1	0	0	0
9104	1	1	1	0	1	1	0	0	0	1	1	1	0	0
9105	0	0	0	1	0	0	0	0	1	1	1	1	0	0
9106	0	0	0	1	0	0	1	0	0	1	0	0	0	0
9107	0	0	0	0	0	1	0	1	1	1	0	0	1	0
9108	0	0	0	0	0	0	1	0	0	0	0	0	0	0
9109	0	0	0	0	1	0	1	0	0	1	0	0	0	0
9110	0	0	0	1	1	1	1	0	1	0	0	0	0	0
9111	0	0	1	1	1	0	0	0	0	0	0	1	0	0
9112	0	0	0	1	1	0	0	0	0	0	1	0	0	0
9171	0	0	0	0	0	0	1	0	0	0	0	0	0	0
9401	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9402	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9403	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9404	0	1	1	0	0	0	0	0	0	0	0	0	0	0
9408	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9409	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9410	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9411	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9412	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9413	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9414	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9415	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9416	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9417	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9418	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9419	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9420	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9426	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9427	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9428	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9429	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9430	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9431	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9432	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9433	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9434	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9435	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9436	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9471	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Lampiran 2. (Lanjutan)

Kode	9402	9403	9404	9408	9409	9410	9411	9412	9413	9414	9415	9416	9417	9418
9101	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9102	0	0	1	0	0	0	0	1	0	0	0	0	0	0
9103	0	0	1	0	0	0	0	0	0	0	0	0	0	0
9104	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9105	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9106	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9107	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9108	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9109	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9110	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9111	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9112	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9171	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9401	0	0	0	0	0	0	0	0	1	1	0	0	0	0
9402	0	0	0	0	0	0	0	0	0	0	0	1	0	0
9403	0	0	0	0	0	0	0	0	0	0	0	1	1	0
9404	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9408	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9409	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9410	0	0	1	0	0	0	0	1	0	0	0	0	0	0
9411	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9412	0	0	0	0	0	1	0	0	0	0	1	0	0	0
9413	0	0	0	0	0	0	0	0	0	1	1	1	1	0
9414	0	0	0	0	0	0	0	0	1	0	1	1	0	0
9415	0	0	0	0	0	0	0	1	1	1	0	1	0	0
9416	1	1	0	0	0	0	0	0	1	1	1	0	1	0
9417	0	1	0	0	0	0	0	0	1	0	0	1	0	0
9418	0	0	0	0	0	0	1	0	0	0	0	0	0	0
9419	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9420	0	1	0	0	0	0	0	0	0	0	0	0	1	0
9426	0	0	1	0	0	0	0	0	0	0	0	0	0	0
9427	0	0	0	0	1	0	0	0	0	0	0	0	0	0
9428	0	0	0	0	0	0	1	0	0	0	0	0	0	1
9429	1	0	0	0	0	0	0	1	0	0	1	1	0	0
9430	1	0	0	0	0	0	1	0	0	0	0	0	0	1
9431	1	0	0	0	0	0	0	0	0	0	0	0	0	1
9432	1	1	0	0	0	0	0	0	0	0	0	1	0	0
9433	0	0	0	0	0	1	1	1	0	0	0	0	0	0
9434	0	0	1	0	0	1	0	1	0	0	0	0	0	0
9435	0	0	1	0	0	1	0	0	0	0	0	0	0	0
9436	0	0	0	0	0	1	0	1	0	0	0	0	0	0
9471	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Lampiran 2. (Lanjutan)

Kode	9419	9420	9426	9427	9428	9429	9430	9431	9432	9433	9434	9435	9436	9471
9101	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9102	0	0	0	0	0	0	0	0	0	0	1	0	0	0
9103	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9104	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9105	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9106	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9107	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9108	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9109	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9110	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9111	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9112	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9171	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9401	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9402	0	0	0	0	0	1	1	1	1	0	0	0	0	0
9403	1	1	0	0	0	0	0	0	1	0	0	0	0	1
9404	0	0	1	0	0	0	0	0	0	0	1	1	0	0
9408	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9409	0	0	0	1	0	0	0	0	0	0	0	0	0	0
9410	0	0	0	0	0	0	0	0	0	1	1	1	1	0
9411	0	0	0	0	1	0	1	0	0	1	0	0	0	0
9412	0	0	0	0	0	1	0	0	0	1	1	0	1	0
9413	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9414	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9415	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9416	0	0	0	0	0	1	0	0	1	0	0	0	0	0
9417	0	1	0	0	0	0	0	0	0	0	0	0	0	0
9418	0	0	0	0	1	0	1	1	0	0	0	0	0	0
9419	0	0	0	0	1	0	0	1	1	0	0	0	0	0
9420	0	0	0	0	0	0	0	0	0	0	0	0	0	1
9426	0	0	0	0	1	0	0	0	0	1	0	1	0	0
9427	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9428	1	0	1	0	0	0	0	1	0	1	0	0	0	0
9429	0	0	0	0	0	0	1	0	0	1	0	0	0	0
9430	0	0	0	0	0	1	0	1	0	1	0	0	0	0
9431	1	0	0	0	1	0	1	0	1	0	0	0	0	0
9432	1	0	0	0	0	0	0	1	0	0	0	0	0	0
9433	0	0	1	0	1	1	1	0	0	0	0	1	0	0
9434	0	0	0	0	0	0	0	0	0	0	0	0	1	0
9435	0	0	1	0	0	0	0	0	0	1	0	0	0	0
9436	0	0	0	0	0	0	0	0	0	0	1	0	0	0
9471	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Lampiran 3. R code dan Output Uji Multikolinieritas dan Regresi Probit Spasial dengan RIS Simulator

```

#PANGGIL DATA
setwd("C:/Users/FlEX/Desktop/olah data")
W=read.csv(choose.files(),header=F)
W=as.matrix(W)
dataku=read.csv(choose.files(),header=T)
y=dataku$y
y=as.matrix(y)
n <- length(y)
X=as.matrix(cbind(rep(1,n),dataku[,-1]))
X=unname(X)

#####

#UJI MULTIKOLINIERITAS
Xm<-dataku[,-1]
ym<-dataku[,1]
library(mctest)
VIF=mctest(Xm, ym, type="i", method="VIF")
VIF[[1]][,1]

> VIF[[1]][,1]
      x1      x2      x3      x4      x5
3.297933 2.374027 1.069554 3.844412 1.770518

#####

#RIS SIMULATOR
library(numDeriv)
ris_estimator <- function(y, X, W, control = list(...), debug = FALSE)
{
  con <- list(
    R = 1000,
    store.density = FALSE,
    optim.max.iterations = 1000,
    optim.reltol = .0025
  )
  con[names(control)] <- control
  with(con, {
    n <- length(y)
    k <- dim(X)[2]
    Z <- diag(as.vector(1-2*y))
    In <- diag(n)

    #Random draw dari importance density function memberikan batas atas
    menggunakan antithetical sampling
    random.draw <- function(upper.bound)
    {
      q <- runif(R/2)
      q <- c(q, 1-q)
      qnorm(q * pnorm(upper.bound))
    }
    iter <- 0
    densities <- NULL
    #Mencari nilai log likelihood dengan menggunakan RIS
    ll <- function(par)
    {
      iter <- iter + 1
      rho <- -1 + 2 * pnorm(par[k+1])
      beta <- par[1:k]
      Ai <- -rho * W
      diag(Ai) <- 1
      A<-solve(Ai)
      omega <- Z %*% A %*% t(A) %*% t(Z)
      V <- -Z %*% A %*% X %*% beta
      B <- solve(chol(solve(omega)))
      Eta0 <- Eta <- matrix(NA, nrow = n, ncol = R)
      Eta0[n,] <- rep(1/B[n,n] * V[n], R)
    }
  }
}

```

Lampiran 3. (Lanjutan)

```
Eta[n,] <- random.draw(Eta0[n,])
for (i in (n-1):1)
  {
    Eta0[i,] <- 1/B[i,i] * (V[i] - t(B[i,(i+1):n]) %% Eta[(i+1):n,])
    if (i > 1)
      Eta[i,] <- random.draw(Eta0[i,])
  }
pnEta0 <- pnorm(Eta0)
a <- 1 / mean(pnEta0)
ll <- log(mean(apply(pnEta0 * a, 2, prod))) - log(1/R) - n * log(a)
ll
}
par <- optim(rep(0,k+1), ll,control=list(fnscale = -1,
  reltol = optim.reltol, maxit = optim.max.iterations))

hess=hessian(x=par$par,func=ll)
se <- sqrt(diag(solve(-hess))[1:k])
z.hit=(par$par[1:k])/se
p.value=c()
for (i in 1:k){
  p.value[i]=2*pt(abs(z.hit[i]),lower.tail = F)}
hasil=data.frame(Estimate=par$par[1:k],Std.error=
se,z.value=z.hit,p.value=p.value)
list(hasil=hasil, rho = -1 + 2 * pnorm(par$par[k+1]), hessian=hess)
})
}
set.seed(666)
estimasi=risk_estimator(y, X, w,control=list(R = 1000, store.density =
FALSE, optim.max.iterations = 1000, optim.reltol = .0025))
round(estimasi$hasil,5)
round(estimasi$rho,5)

> round(estimasi$hasil,5)
  Estimate Std.error  z.value p.value
1  0.00256  0.00037  6.83740  0e+00
2 -0.00015  0.00003 -5.10397  1e-05
3  0.01866  0.00355  5.25102  1e-05
4 -0.02808  0.00345 -8.14187  0e+00
5  0.01605  0.00030 53.17421  0e+00
6  0.00811  0.00130  6.26322  0e+00

> round(estimasi$rho,5)
[1] 0.6866
```

Lampiran 4. R code dan Output Uji Dependensi Spasial dan Uji Parameter Secara Simultan dan Parsial

```

#UJI DEPENDENSI SPASIAL LAG
n <- length(y)
betacap=as.matrix(estimasi$hasil[,1])
E=y-(X %%% betacap)
T=sum(diag((t(W)+W) %%% W))
I=diag(1,n)
D=((t(W %%% X %%% betacap) %%% (I- X %%% solve(t(X) %%% X) %%% t(X)) %%%
(W %%% X %%% betacap))/((t(E) %%% E)/n))+T
LM=((t(E) %%% W %%% y)/((t(E) %%% E)/n))^2/D
#TABEL CHISQUARE
chi_tab=qchisq(0.95,1)
LM_test=data.frame(LM_test=LM,Chi_Tabel=chi_tab)
list(LM_test=LM_test)

> list(LM_test=LM_test)
$LM_test
  LM_test Chi_Tabel
1 5.35393  3.841459

#####

#UJI SIMULTAN
#mencari log dengan var prediktor
logB = 0
for (i in 1:nrow(X)){
  logB1 = (y[i] * log(p1[i]) + (1 - y[i]) * log(p0[i])) + logB
  logB = logB1
}
#mencari log tanpa variabel prediktor
beta0=c(0.00256)
xsig0=cbind(x0)
XB0=xsig0%%beta0
gamma1=solve((I-(rho1*w)))
yhat0=gamma1%%XB0
A0=t(I-(rho1*w))
B0=I-(rho1*w)
omega0=solve(A0%%B0)
omegai0=diag(omega0)
S0=yhat0/omegai0
mean0=mean(S0)
var0=var(S0)
p10=pnorm(S0,mean=mean0,sd=sqrt(var0))
p00=1-p10
logB0 = 0
for (i in 1:nrow(X)){
  logB10 = (y[i] * log(p10[i]) + (1 - y[i]) * log(p00[i])) + logB0
  logB0 = logB10
}
LR=2*(logB-logB0)
chi_tab2=qchisq(0.95,5)
LR_test=data.frame(LR_test=LR,Chi_Tabel=chi_tab2)
list(LR_test=LR_test)

> list(LR_test=LR_test)
$LR_test
  LR_test Chi_Tabel
1 31.96849  11.0705

#####

#UJI PARSIAL
n <- length(y)
k <- dim(X)[2]
WALD=estimasi$hasil[,1]/estimasi$hasil[,2]
p.value=c()
for (i in 1:k)
{

```

Lampiran 4. (Lanjutan)

```
p.value[i]=2*pt(abs(WALD[i]),n-(k+1),lower.tail = F)
}
hasilnya=data.frame(z.value=WALD,p.value=p.value)
list(Uji_wald=hasilnya)

> list(Uji_wald=hasilnya)
$Uji_wald
  z.value      p.value
1  6.837403 6.200138e-08
2 -5.103969 1.171565e-05
3  5.251018 7.504130e-06
4 -8.141870 1.369726e-09
5 53.174209 4.544483e-35
6  6.263216 3.480349e-07
```

Lampiran 5. R code Prediksi Probabilitas, *Confusion Matrix* dan Efek Marginal

```

#data diterjemahkan
x0=x[,1] x1=x[,2] x2=x[,3] x3=x[,4] x4=x[,5] x5=x[,6] y=y[,1]

rho1=estimasi$rho
n <- length(y)
I=diag(1,n)
#masukkan nilai beta yang signifikan
beta1=c(0.00256,-0.00015,0.01866,-0.02808,0.01605,0.00811)
#masukkan xi yang signifikan
Xsig=cbind(x0,x1,x2,x3,x4,x5)
#perhitungan
XB=Xsig%%beta1
gamma1=solve((I-(rho1*w)))
yhat=gamma1%%XB

#PREDIKSI PROBABILITAS
#mencari omega ii
A=t(I-(rho1*w))
B=I-(rho1*w)
omega=solve(A%%B)
omegaii=diag(omega)
S=-yhat/omegaii
mean1=mean(S)
var1=var(S)
p0=pnorm(S,mean=mean1,sd=sqrt(var1))
p1=1-p0
peluang=data.frame(y, prob1=p1, prob0=p0)
list(pred.prob=peluang)

#####

#CONFUSION MATRIX
pcap = matrix(nrow=n, ncol=1)
for (k in 1:n){
  if (p1[k] < 0.5) pcap[k]=0
  else pcap[k] = 1
}
library(crossval)
cm1 = confusionMatrix(y, pcap, negative=0)
cm2 = as.vector(cm1)
names(cm2)=c("FP", "TP", "TN", "FN")
list(confusion.matrix=cm2)
de1 = diagnosticErrors(cm1)
de2=de1[1:3]
list(acc.sens.spec=de2)

> list(confusion.matrix=cm2)
$confusion.matrix
FP TP TN FN
12 13 13 4

> list(acc.sens.spec=de2)
$acc.sens.spec
      acc      sens      spec
0.6190476 0.7647059 0.5200000

#####

#EFEK MARGINAL
gamma2=diag(gamma1)
tur1=gamma2%%t(beta1)
yhat1=dnorm(S,mean=mean1,sd=sqrt(var1))
em=matrix(0,42,6)
for(i in 1:6){
  em[,i]=yhat1 * tur1[,i]/omegaii
}
em1=round(em,6)
list(efek.marjinal=em1[,-1])

```

Lampiran 6. Prediksi Probabilitas Kabupaten/Kota di Pulau Papua dari Model SAR Probit

No.	Nama Kab/Kota	$P(y = 1)$	$P(y = 0)$	\hat{y}_i
1	Fak Fak	0,88595	0,11405	1
2	Kaimana	0,78448	0,21552	1
3	Teluk Wondama	0,22529	0,77471	0
4	Teluk Bintuni	0,40671	0,59329	0
5	Manokwari	0,34007	0,65993	0
6	Sorong Selatan	0,72812	0,27188	1
7	Sorong	0,60680	0,39320	1
8	Raja Ampat	0,64483	0,35517	1
9	Tambrauw	0,73383	0,26617	1
10	Maybrat	0,55246	0,44754	1
11	Manokwari Selatan	0,59699	0,40301	1
12	Pegunungan Arfak	0,26980	0,73020	1
13	Kota Sorong	0,61772	0,38228	1
14	Merauke	0,47174	0,52826	0
15	Jayawijaya	0,01860	0,98140	0
16	Jayapura	0,50642	0,49358	1
17	Nabire	0,85717	0,14283	1
18	Yapen Waropen	0,96382	0,03618	1
19	Biak Namfor	0,67388	0,32612	1
20	Paniai	0,78462	0,21538	1
21	Puncak Jaya	0,78395	0,21605	1
22	Mimika	0,74022	0,25978	1
23	Boven Digoel	0,41589	0,58411	0
24	Mappi	0,25968	0,74032	0
25	Asmat	0,03007	0,96993	0
26	Yahukimo	0,32358	0,67642	0
27	Pegunungan Bintang	0,52267	0,47733	1
28	Tolikara	0,08659	0,91341	0
29	Sarmi	0,71310	0,28690	1
30	Keerom	0,68145	0,31855	1
31	Waropen	0,96012	0,03988	1
32	Supiori	0,66245	0,33755	1
33	Membramo Raya	0,26253	0,73747	0
34	Nduga	0,01694	0,98306	0
35	Lanny Jaya	0,03504	0,96496	0
36	Mamberamo Tengah	0,04911	0,95089	0
37	Yalimo	0,23619	0,76381	0

Lampiran 6. (Lanjutan)

No.	Nama Kab/Kota	$P(y = 1)$	$P(y = 0)$	y^*
38	Puncak	0,26708	0,73292	0
39	Dogiyai	0,87864	0,12136	1
40	Intan Jaya	0,50374	0,49626	1
41	Deiyai	0,95872	0,04128	1
42	Kota Jayapura	0,50828	0,49172	1

Lampiran 7. Efek Marjinal Variabel Pengeluaran Per Kapita Terhadap Indeks Kesehatan

No.	Nama Kab/Kota	Efek Marjinal
1	Fak Fak	-0,00005
2	Kaimana	-0,00009
3	Teluk Wondama	-0,00009
4	Teluk Bintuni	-0,00012
5	Manokwari	-0,00011
6	Sorong Selatan	-0,00009
7	Sorong	-0,00010
8	Raja Ampat	-0,00008
9	Tambrauw	-0,00009
10	Maybrat	-0,00011
11	Manokwari Selatan	-0,00011
12	Pegunungan Arfak	-0,00009
13	Kota Sorong	-0,00008
14	Merauke	-0,00010
15	Jayawijaya	-0,00002
16	Jayapura	-0,00012
17	Nabire	-0,00007
18	Yapen Waropen	-0,00004
19	Biak Namfor	-0,00006
20	Paniai	-0,00009
21	Puncak Jaya	-0,00009
22	Mimika	-0,00011
23	Boven Digoel	-0,00011
24	Mappi	-0,00009
25	Asmat	-0,00002
26	Yahukimo	-0,00012
27	Pegunungan Bintang	-0,00012
28	Tolikara	-0,00005
29	Sarmi	-0,00011
30	Keerom	-0,00010
31	Waropen	-0,00003
32	Supiori	-0,00006
33	Membramo Raya	-0,00010
34	Nduga	-0,00001
35	Lanny Jaya	-0,00003
36	Mamberamo Tengah	-0,00003
37	Yalimo	-0,00010

Lampiran 7. (Lanjutan)

No.	Nama Kab/Kota	Efek Marjinal
38	Puncak	-0,00011
39	Dogiyai	-0,00006
40	Intan Jaya	-0,00012
41	Deiyai	-0,00003
42	Kota Jayapura	-0,00010

Lampiran 8. Efek Marjinal Variabel Persentase Rumah Tangga dengan Air Minum Layak Terhadap Indeks Kesehatan

No.	Nama Kab/Kota	Efek Marjinal
1	Fak Fak	0,00673
2	Kaimana	0,01155
3	Teluk Wondama	0,01145
4	Teluk Bintuni	0,01464
5	Manokwari	0,01309
6	Sorong Selatan	0,01140
7	Sorong	0,01282
8	Raja Ampat	0,00991
9	Tambrauw	0,01107
10	Maybrat	0,01396
11	Manokwari Selatan	0,01372
12	Pegunungan Arfak	0,01120
13	Kota Sorong	0,01015
14	Merauke	0,01281
15	Jayawijaya	0,00181
16	Jayapura	0,01495
17	Nabire	0,00892
18	Yapen Waropen	0,00452
19	Biak Namfor	0,00738
20	Paniai	0,01131
21	Puncak Jaya	0,01102
22	Mimika	0,01320
23	Boven Digoel	0,01421
24	Mappi	0,01148
25	Asmat	0,00267
26	Yahukimo	0,01430
27	Pegunungan Bintang	0,01465
28	Tolikara	0,00586
29	Sarmi	0,01316
30	Keerom	0,01199
31	Waropen	0,00334
32	Supiori	0,00748
33	Membramo Raya	0,01286
34	Nduga	0,00174
35	Lanny Jaya	0,00306
36	Mamberamo Tengah	0,00397
37	Yalimo	0,01203

Lampiran 8. (Lanjutan)

No.	Nama Kab/Kota	Efek Marjinal
38	Puncak	0,01346
39	Dogiyai	0,00767
40	Intan Jaya	0,01517
41	Deiyai	0,00318
42	Kota Jayapura	0,01268

Lampiran 9. Efek Marjinal Variabel Persentase Penduduk yang Mempunyai Kebiasaan Merokok Terhadap Indeks Kesehatan

No.	Nama Kab/Kota	Efek Marjinal
1	Fak Fak	-0,01012
2	Kaimana	-0,01738
3	Teluk Wondama	-0,01724
4	Teluk Bintuni	-0,02202
5	Manokwari	-0,01969
6	Sorong Selatan	-0,01715
7	Sorong	-0,01930
8	Raja Ampat	-0,01492
9	Tambrauw	-0,01666
10	Maybrat	-0,02101
11	Manokwari Selatan	-0,02065
12	Pegunungan Arfak	-0,01686
13	Kota Sorong	-0,01528
14	Merauke	-0,01927
15	Jayawijaya	-0,00272
16	Jayapura	-0,02250
17	Nabire	-0,01342
18	Yapen Waropen	-0,00681
19	Biak Namfor	-0,01110
20	Paniai	-0,01702
21	Puncak Jaya	-0,01658
22	Mimika	-0,01986
23	Boven Digoel	-0,02138
24	Mappi	-0,01728
25	Asmat	-0,00401
26	Yahukimo	-0,02152
27	Pegunungan Bintang	-0,02205
28	Tolikara	-0,00882
29	Sarmi	-0,01980
30	Keerom	-0,01804
31	Waropen	-0,00503
32	Supiori	-0,01126
33	Membramo Raya	-0,01935
34	Nduga	-0,00262
35	Lanny Jaya	-0,00460
36	Mamberamo Tengah	-0,00597
37	Yalimo	-0,01810

Lampiran 9. (Lanjutan)

No.	Nama Kab/Kota	Efek Marjinal
38	Puncak	-0,02025
39	Dogiyai	-0,01154
40	Intan Jaya	-0,02283
41	Deiyai	-0,00479
42	Kota Jayapura	-0,01908

Lampiran 10. Efek Marjinal Variabel Rata-rata Lama Sekolah Terhadap Indeks Kesehatan

No.	Nama Kab/Kota	Efek Marjinal
1	Fak Fak	0,00579
2	Kaimana	0,00993
3	Teluk Wondama	0,00985
4	Teluk Bintuni	0,01259
5	Manokwari	0,01126
6	Sorong Selatan	0,00980
7	Sorong	0,01103
8	Raja Ampat	0,00853
9	Tambrauw	0,00952
10	Maybrat	0,01201
11	Manokwari Selatan	0,01180
12	Pegunungan Arfak	0,00964
13	Kota Sorong	0,00873
14	Merauke	0,01101
15	Jayawijaya	0,00156
16	Jayapura	0,01286
17	Nabire	0,00767
18	Yapen Waropen	0,00389
19	Biak Namfor	0,00635
20	Paniai	0,00973
21	Puncak Jaya	0,00948
22	Mimika	0,01135
23	Boven Digoel	0,01222
24	Mappi	0,00988
25	Asmat	0,00229
26	Yahukimo	0,01230
27	Pegunungan Bintang	0,01261
28	Tolikara	0,00504
29	Sarmi	0,01132
30	Keerom	0,01031
31	Waropen	0,00288
32	Supiori	0,00643
33	Membramo Raya	0,01106
34	Nduga	0,00150
35	Lanny Jaya	0,00263
36	Mamberamo Tengah	0,00341
37	Yalimo	0,01034

Lampiran 10. (Lanjutan)

No.	Nama Kab/Kota	Efek Marjinal
38	Puncak	0,01158
39	Dogiyai	0,00659
40	Intan Jaya	0,01305
41	Deiyai	0,00274
42	Kota Jayapura	0,01090

Lampiran 11. Efek Marjinal Variabel Proporsi Fasilitas Kesehatan Terhadap Indeks Kesehatan

No.	Nama Kab/Kota	Efek Marjinal
1	Fak Fak	0,00292
2	Kaimana	0,00502
3	Teluk Wondama	0,00498
4	Teluk Bintuni	0,00636
5	Manokwari	0,00569
6	Sorong Selatan	0,00495
7	Sorong	0,00557
8	Raja Ampat	0,00431
9	Tambrauw	0,00481
10	Maybrat	0,00607
11	Manokwari Selatan	0,00596
12	Pegunungan Arfak	0,00487
13	Kota Sorong	0,00441
14	Merauke	0,00557
15	Jayawijaya	0,00079
16	Jayapura	0,00650
17	Nabire	0,00388
18	Yapen Waropen	0,00197
19	Biak Namfor	0,00321
20	Paniai	0,00492
21	Puncak Jaya	0,00479
22	Mimika	0,00574
23	Boven Digoel	0,00617
24	Mappi	0,00499
25	Asmat	0,00116
26	Yahukimo	0,00622
27	Pegunungan Bintang	0,00637
28	Tolikara	0,00255
29	Sarmi	0,00572
30	Keerom	0,00521
31	Waropen	0,00145
32	Supiori	0,00325
33	Membramo Raya	0,00559
34	Nduga	0,00076
35	Lanny Jaya	0,00133
36	Mamberamo Tengah	0,00173
37	Yalimo	0,00523

Lampiran 11. (Lanjutan)

No.	Nama Kab/Kota	Efek Marjinal
38	Puncak	0,00585
39	Dogiyai	0,00333
40	Intan Jaya	0,00660
41	Deiyai	0,00138
42	Kota Jayapura	0,00551

(Halaman ini sengaja dikosongkan)

BIOGRAFI PENULIS



Penulis dilahirkan di Magelang, Jawa Tengah pada tanggal 25 Mei 1986, merupakan anak kedua dari tiga bersaudara dan anak dari pasangan Bapak Sudyanto dan Ibu Elyza Dewanti. Saat ini penulis sudah berkeluarga dengan istri bernama Rahmadanie Septa Irevanie dengan dua anak bernama Asheeqa Sheza Ramadhani dan Azzamir Rafka Muhammad. Penulis telah menempuh pendidikan formal di SDN Tanjunganom II Magelang (1992-1998), SLTP Negeri 1 Magelang (1998-2001), SMU Negeri 3 Magelang (2001-2004), dan Sekolah Tinggi Ilmu Statistik (2004-2008) Jurusan Statistik Kependudukan. Setelah menamatkan pendidikan DIV di Sekolah Tinggi Ilmu Statistik, penulis ditugaskan bekerja di Badan Pusat Statistik Kabupaten Jayapura, Provinsi Papua. Pada tahun 2016 penulis memperoleh kesempatan beasiswa dari Badan Pusat Statistik untuk melanjutkan studi S2 di Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data, Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Pembaca yang ingin memberikan kritik, saran dan pertanyaan mengenai penelitian ini, dapat menghubunginya melalui email tf Dewanto@bps.go.id.

(Halaman ini sengaja dikosongkan)