

ANALISIS REGRESI LOGISTIK KELAS LATEN
MINAT MASYARAKAT TERHADAP ASURANSI

Skripsi

Disusun untuk melengkapi syarat-syarat
guna memperoleh gelar Sarjana Sains



IRENA PURWA NINGRIDA

3125136343

PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA

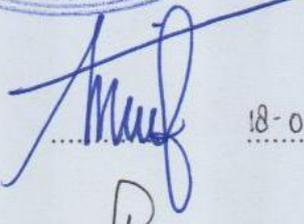
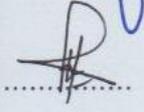
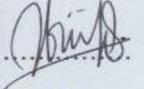
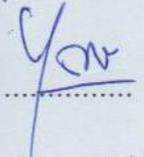
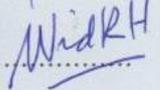
2017

LEMBAR PERSETUJUAN HASIL SIDANG SKRIPSI

ANALISIS REGRESI LOGISTIK KELAS LATEN MINAT MASYARAKAT TERHADAP ASURANSI

Nama : Irena Purwa Ningrida

No. Registrasi : 3125136343

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan	: Prof. Dr. Suyono, M.Si. NIP. 19671218 199303 1 005		18-08-2017
Wakil Penanggung Jawab			
Wakil Dekan I	: Dr. Muktiningsih, M.Si. NIP. 19640511 198903 2 001		18-08-2017
Ketua	: Dr. Lukita Ambarwati, S.Pd, M.Si. NIP. 19721026 200112 2 001		16-08-2017
Sekretaris	: Ibnu Hadi, M.Si. NIP. 19810718 200801 1 017		14-08-2017
Penguji	: Drs. Mulyono, M.Kom. NIP. 19660517 199403 1 003		14-08-2017
Pembimbing I	: Dra. Widianti Rahayu, M.Si. NIP. 19661103 200112 2 001		16-08-2017
Pembimbing II	: Ir. Fariani Hermin, M.T. NIP. 19600211 198703 2 001		16-08-2017

Dinyatakan lulus ujian skripsi tanggal: 11 Agustus 2017

ABSTRACT

IRENA PURWA NINGRIDA, 3125136343. Latent Class Logistic Regression Analysis on Human Interest in Insurance. Thesis. Faculty of Mathematics and Natural Science Jakarta State University. 2017.

Insurance company in Indonesia has been developed and gotten bigger day by day, but human interest in insurance are still low. Human interest in insurance is defined as a desire to have an insurance. The measurement of human interest in insurance can not be observed directly (latent variable), but it can be observed through some indicators that described the human interest in insurance itself. Latent class logistic regression analysis is a statistical method that can describe the latent variable's structure by clustering an object according to the indicators involving covariates or explanatory variables. This kind of analysis assumes the local independence which means among the indicators must be independent on a latent class. This research aim is aim to see the application of latent class logistic regression to cluster the human interest in East Jakarta on insurance. Latent variable used in this research is human interest in insurance, which is measured by some indicators, namely the awareness of safety need, knowledge about insurance, understanding about insurance, and the savor in insurance. Covariates in this research are gender, age, education, and income. The result shows that human interest in insurance can be clustered into 3 latent class, with the lowest AIC 1785.630. The precentage of each of the latent class are 43.05% for the first latent class, 29.30% for the second latent class, dan 27.66% for the third latent class. The three latent class describe the group of human interest in insurance namely good, doubt, and not good.

Keywords : *insurance, interest in insurance, latent class logistic regression analysis, AIC.*

ABSTRAK

IRENA PURWA NINGRIDA, 3125136343. Analisis Regresi Logistik Kelas Laten Minat Masyarakat Terhadap Asuransi. Skripsi. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. 2017.

Perusahaan asuransi yang ada di Indonesia semakin berkembang dan bertambah banyak, namun, minat masyarakat terhadap asuransi masih tergolong rendah. Minat terhadap asuransi didefinisikan sebagai keinginan seseorang untuk memiliki asuransi. Pengukuran terhadap minat masyarakat terhadap asuransi tidak dapat dilakukan secara langsung (variabel laten) melainkan dapat dilakukan melalui indikator-indikator yang menggambarkan minat tersebut. Analisis regresi logistik kelas laten merupakan metode yang dapat menggambarkan struktur variabel laten dengan mengelompokkan objek pengamatan berdasarkan indikator (*symptom*) dengan melibatkan kovariat atau variabel penjelas. Analisis ini memiliki asumsi kebebasan lokal yang berarti antar variabel indikator haruslah saling bebas dalam suatu kelas laten. Penelitian ini bertujuan untuk menerapkan analisis regresi logistik untuk mengelompokkan minat masyarakat di Jakarta Timur terhadap asuransi. Variabel laten yang digunakan adalah minat masyarakat terhadap asuransi, yang diukur melalui variabel indikator berupa kesadaran masyarakat akan perlindungan, pengetahuan masyarakat terhadap asuransi, pemahaman terhadap asuransi, dan ketertarikan masyarakat terhadap asuransi. Kovariat yang digunakan adalah jenis kelamin, usia, tingkat pendidikan, dan pendapatan. Hasil penelitian ini menunjukkan bahwa minat masyarakat terhadap asuransi dapat dikelompokkan menjadi 3 kelas dengan nilai AIC terkecil yaitu 1785.630. Presentase yang dimiliki masing-masing kelas adalah 43.05% untuk kelas pertama, 29.30% untuk kelompok kedua, dan 27.66% untuk kelompok ketiga. Ketiga kelas tersebut menggambarkan minat masyarakat terhadap asuransi baik, ragu-ragu, dan tidak baik.

Kata kunci : asuransi, minat terhadap asuransi, analisis regresi logistik kelas laten, AIC.

PERSEMBAHANKU...

*"Maka sesungguhnya bersama kesulitan ada kemudahan, sesungguhnya
bersama kesulitan ada kemudahan..."*

(QS. Al-Insyirah: 5-6)

"Pure mathematics is, in its way, the poetry of logical ideas"

-Albert Einstein

Skripsi ini kupersembahkan untuk Abah, Mama, dan Adzra.

"Terima kasih atas dukungan, do'a, serta kasih sayang kalian".

KATA PENGANTAR

Puji syukur kepada Allah SWT. atas nikmat pengetahuan dan kemampuan sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Regresi Logistik Kelas Laten" yang merupakan salah satu syarat dalam memperoleh gelar Sarjana Program Studi Matematika Universitas Negeri Jakarta.

Skripsi ini berhasil diselesaikan tidak terlepas dari adanya bantuan dari berbagai pihak. Oleh karena itu, dalam kesempatan ini penulis ingin menyampaikan terima kasih terutama kepada:

1. Ibu Dra. Widyanti Rahayu, M.Si., selaku Dosen Pembimbing I dan Ibu Ir. Fariani Hermin, MT., selaku Dosen Pembimbing II, yang telah meluangkan waktunya dalam memberikan bimbingan, saran, nasehat serta arahan sehingga skripsi ini dapat menjadi lebih baik dan terarah.
2. Ibu Dr. Lukita Ambarwati, S.Pd., M.Si., selaku Koordinator Program Studi Matematika FMIPA UNJ yang telah banyak membantu penulis.
3. Ibu Dian Handayani, M.Si., selaku Pembimbing Akademik atas segala bimbingan, arahan, dan nasihat selama perkuliahan, dan seluruh Bapak/Ibu dosen atas pengajarannya yang telah diberikan, serta karyawan/karyawati FMIPA UNJ yang telah memberikan informasi yang penulis butuhkan dalam menyelesaikan skripsi.
4. Abah dan Mama yang selalu mendo'akan, mendukung, memberi motivasi, menjadi inspirasi, dan setia membantu penulis dengan penuh cinta dan kasih sayang yang tulus.
5. Adik perempuan penulis, Adzra Nur Aqilah yang terus memberi semangat, mendoakan penulis, dan selalu menghibur ketika penulis mengalami

kesulitan dalam penulisan skripsi ini.

6. Sahabat-sahabat alhamdulillahku, Ita, Trias, Atikah, dan Rahida yang selalu mendo'akan, memberi semangat, memotivasi, dan menemani penulis dalam menjalani kerasnya kehidupan kuliah. Semoga persahabatan kita sampai surga-NYA..
7. Ten Brothers, Inggrid, Novita, Ica, Dela, Dian, Oo, Jiwo, Dannior, dan Dion yang selalu setia menghibur, memberi semangat, dan memotivasi penulis.
8. Teman-teman seperjuangan 2017, Hanun, Tias, Defy, Nurul, Syevie, Nisa, Ashadah, Umam, dan Daniel atas diskusi dan pemikirannya dalam membantu penulis menyelesaikan skripsi ini.
9. Teman-teman kosan Plurist, Kak Dinny, Kak Nji, Kak Irma, Indah, Sabil, Cantika, dan Diniar yang menjadi keluarga kedua penulis selama menjalani masa kuliah.
10. Teman-teman Matematika 2013 yang selalu membantu, mendo'akan, menyemangati, dan memotivasi penulis.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Masukan dan kritikan akan sangat berarti. Semoga skripsi ini dapat bermanfaat bagi pembaca sekalian.

Jakarta, Agustus 2017

Irena Purwa Ningrida

DAFTAR ISI

ABSTRACT	i
ABSTRAK	ii
KATA PENGANTAR	iv
DAFTAR ISI	vi
DAFTAR TABEL	viii
DAFTAR GAMBAR	ix
I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Perumusan Masalah	4
1.3 Pembatasan Masalah	4
1.4 Tujuan Penulisan	5
1.5 Manfaat Penulisan	5
II LANDASAN TEORI	6
2.1 Asuransi	6
2.2 Minat terhadap Asuransi	8
2.3 Regresi Logistik Multinomial	10
2.4 Teorema Bayes	14
2.5 Analisis Kelas Laten	16
2.5.1 Model Kelas Laten	16
2.5.2 Estimasi Parameter Model Kelas Laten	21
2.6 Teknik Pengambilan Sampel	25

2.7	Validitas dan Reliabilitas	26
2.8	Pemilihan Model Terbaik	29
2.9	Asumsi Kebebasan Lokal	30
III PEMBAHASAN		31
3.1	Analisis Regresi Logistik Kelas Laten	31
3.1.1	Model Regresi Logistik Kelas Laten	32
3.1.2	Estimasi Parameter	34
3.1.3	Interpretasi Model Regresi Logistik Kelas Laten	36
3.2	Diagram Alir Analisis Regresi Logistik Kelas Laten	37
3.3	Identifikasi Variabel	40
3.4	Sumber Data	41
3.5	Uji Validitas dan Reliabilitas	43
3.6	Hasil Analisis Data	45
3.6.1	Deskripsi Data	45
3.6.2	Estimasi Peluang Awal Kelas Laten	49
3.6.3	Pemilihan Model Terbaik	50
3.6.4	Estimasi Parameter Model Terpilih	50
3.6.5	Model Regresi Logistik Kelas Laten	53
3.6.6	Interpretasi Model	55
IV PENUTUP		57
4.1	Kesimpulan	57
4.2	Saran	58
DAFTAR PUSTAKA		59
LAMPIRAN-LAMPIRAN		61

DAFTAR TABEL

3.1	Variabel Indikator yang Digunakan	40
3.2	Variabel Kovariat yang Digunakan	41
3.3	Hasil Uji Validitas	44
3.4	Peluang Awal Kelas Laten (π_k)	49
3.5	Nilai AIC Untuk Setiap Model	50
3.6	Estimasi Peluang Setiap Kategori pada Variabel Indikator	51
3.7	Estimasi Parameter Model Regresi Logistik Kelas 2 dan 1	54
3.8	Estimasi Parameter Model Regresi Logistik Kelas 3 dan 1	55

DAFTAR GAMBAR

3.1	Diagram Alir Analisis Regresi Logistik Kelas Laten	39
3.2	Karakteristik Responden berdasarkan Jenis Kelamin	46
3.3	Karakteristik Responden berdasarkan Usia	46
3.4	Karakteristik Responden berdasarkan Tingkat Pendidikan . . .	47
3.5	Karakteristik Responden berdasarkan Tingkat Pendapatan . . .	48

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Para era modern saat ini, kehidupan masyarakat telah banyak berkembang, termasuk dalam hal kebutuhan. Kebutuhan masyarakat yang semula hanya terbatas pada kebutuhan fisik seperti pangan, sandang, dan papan, kini telah bertambah menjadi kebutuhan lain yang bersifat non-fisik, seperti kebutuhan akan keselamatan dan rasa aman. Asuransi kemudian muncul dan menjadi salah satu alat pemenuhan akan kebutuhan tersebut. Asuransi merupakan perjanjian atau pertanggung jawaban risiko yang dibuat oleh dua pihak, dimana pihak penanggung bertugas membayar kerugian atas risiko yang dipertanggung-jawankan kepada penanggung, sedangkan tertanggung bertugas membayar sejumlah premi secara periodik kepada penanggung.

Seiring dengan berjalannya waktu, perusahaan asuransi yang ada di Indonesia semakin berkembang dan bertambah banyak. Menurut Otoritas Jasa Keuangan per 31 Desember 2015, terdapat 131 perusahaan asuransi yang terdiri dari 76 perusahaan asuransi umum, 50 perusahaan asuransi jiwa, 3 perusahaan asuransi wajib, 2 perusahaan asuransi sosial, dan 6 perusahaan reasuransi di Indonesia. Namun, banyaknya jumlah perusahaan asuransi yang tersedia ini ternyata tidak diimbangi oleh jumlah pemilik asuransi. Asosiasi Asuransi Jiwa Indonesia (AAJI) mencatat bahwa hanya 7,5 % dari total penduduk Indonesia yang memiliki asuransi. Angka ini tergolong rendah mengingat jumlah penduduk Indonesia yang mencapai lebih dari 255 juta jiwa. Keinginan ma-

syarakat untuk berasuransi menjadi salah satu tantangan yang harus dihadapi oleh banyak perusahaan asuransi.

Sebagai salah satu wilayah di Ibukota, minat masyarakat daerah Jakarta Timur terhadap asuransi diharapkan telah baik. Hal ini dikarenakan daerah ini merupakan daerah dengan risiko bencana alam dan kejahatan tertinggi di wilayah Jakarta. Tercatat bahwa daerah ini menjadi wilayah dengan jumlah kebakaran tertinggi dan jumlah korban kecelakaan tertinggi. Oleh karena itu, kebutuhan akan perlindungan dirasakan perlu dipenuhi oleh penduduk pada wilayah ini. Minat masyarakat Jakarta Timur terhadap asuransi kemudian menjadi hal yang perlu diketahui agar pemilik asuransi di daerah ini terus bertambah.

Minat tidak tumbuh atau timbul tanpa adanya interaksi antara manusia dengan suatu objek. Menurut Crow dalam (Tricahyo, 2012) salah satu faktor yang mendorong tumbuhnya minat seseorang adalah faktor dorongan dari dalam, seperti faktor usia dan jenis kelamin. Hermawati (2013) meneliti pengaruh usia, jenis kelamin, dan tingkat pendidikan terhadap kesadaran berasuransi pada masyarakat Indonesia. Hasil penelitian tersebut menunjukkan bahwa terdapat pengaruh jenis kelamin dan tingkat pendidikan terhadap kesadaran berasuransi pada masyarakat Indonesia. Dalam penelitian kemudian akan digunakan variabel usia, jenis kelamin, tingkat pendidikan, dan pendapatan terhadap pengelompokan minat masyarakat terhadap asuransi.

Pengukuran minat masyarakat terhadap asuransi tidak dapat dilakukan secara langsung, namun dapat dilakukan dengan menggunakan gejala atau indikator lain yang dapat menggambarkan minat tersebut, seperti kesadaran akan tingkat kebutuhan, pengetahuan seseorang terhadap asuransi, ketertarikan terhadap asuransi, serta pemahaman akan asuransi. Variabel seperti minat, yang tidak dapat diukur secara langsung, dikenal sebagai variabel laten (*latent*

variable). Variabel laten dapat bersifat kontinu maupun kategorik, dan dapat dibentuk oleh variabel indikator yang bersifat kontinu maupun kategorik pula (Collins & Lanza, 2010). Analisis kelas laten (*latent class analysis*) merupakan salah satu metode yang dapat digunakan untuk menganalisis variabel laten.

Analisis kelas laten merupakan metode statistik untuk mengelompokkan data bersifat kategorik multivariat. Analisis ini membagi beberapa objek menjadi beberapa kelas dimana objek tersebut bersifat kategorik. Tujuan dari analisis ini adalah mengidentifikasi jumlah kelas yang dibutuhkan dalam rangka menjelaskan hubungan variabel-variabel yang diamati. Analisis kelas laten menggunakan fungsi peluang posterior sebagai basis pengelompokannya. Collins dan Lanza (2010) menyatakan bahwa *latent class model* memiliki asumsi kebebasan lokal (*local independence*) yang artinya setiap variabel teramati yang berada pada satu kelompok/kelas laten haruslah saling bebas.

Analisis regresi logistik kelas laten merupakan pengembangan dari analisis kelas laten atau *latent class analysis*. Jika dalam analisis kelas laten pengelompokan objek dilakukan hanya berdasarkan variabel indikator yang digunakan, dalam analisis regresi logistik kelas laten pengelompokan objek dilakukan berdasarkan variabel indikator (*symptom*) sekaligus melibatkan kovariat atau variabel penjelas. Analisis regresi logistik kelas laten dikenalkan oleh Chung, Flaherty, dan Schafer (2006) dalam penelitian mengenai penggunaan marijuana dan perilaku siswa SMA di Amerika Serikat. Kovariat dalam analisis regresi logistik kelas laten digunakan untuk memprediksi keanggotaan kelas laten. Model regresi logistik yang diterapkan dalam analisis ini sebenarnya sama dengan model regresi logistik biasa, hanya saja *outcome* atau variabel independen yang digunakan adalah laten bukan variabel yang teramati langsung.

Beberapa peneliti yang telah melakukan penelitian menggunakan analisis

ini, antara lain Riswan (2010) yang melakukan pengelompokan prestasi matematika siswa Indonesia berdasarkan hasil survey TIMSS (*Trend in International Mathematics and Science Study*) tahun 2007, dan Kurniawan (2013) yang menggunakan analisis regresi logistik kelas laten dalam menentukan model kelas laten performa studi dari mahasiswa penerima beasiswa. Dalam penelitian ini akan dibahas mengenai analisis regresi logistik kelas laten dan penggunaannya untuk mengetahui minat masyarakat Jakarta Timur terhadap asuransi dengan melibatkan variabel kovariat yaitu jenis kelamin, usia, tingkat pendidikan, dan pendapatan.

1.2 Perumusan Masalah

Perumusan masalah yang akan dikaji adalah sebagai berikut:

1. Bagaimanakah pengelompokan minat responden terhadap asuransi berdasarkan indikatornya menggunakan analisis regresi logistik kelas laten?
2. Bagaimana pengaruh jenis kelamin, usia, tingkat pendidikan, dan pendapatan terhadap pengelompokan minat responden terhadap asuransi menggunakan analisis regresi logistik kelas laten?

1.3 Pembatasan Masalah

Untuk mencegah pembahasan yang terlalu luas, diperlukan pembatasan masalah dalam penelitian ini. Adapun pembatasan masalah tersebut adalah :

1. Variabel laten yang digunakan diasumsikan memiliki sebaran multinomial,

2. Variabel indikator yang digunakan merupakan variabel dengan skala ordinal,
3. Kovariat yang digunakan adalah jenis kelamin, usia, tingkat pendidikan, dan pendapatan,
4. Minat masyarakat terhadap asuransi dalam penelitian ini hanya minat masyarakat kota Jakarta Timur.

1.4 Tujuan Penulisan

Adapun tujuan dari penelitian ini adalah :

1. Mengetahui hasil pengelompokan minat masyarakat Jakarta Timur terhadap asuransi berdasarkan variabel indikatornya,
2. Mengetahui pengaruh jenis kelamin, usia, tingkat pendidikan, dan pendapatan terhadap hasil pengelompokan minat masyarakat Jakarta Timur terhadap asuransi.

1.5 Manfaat Penulisan

Manfaat yang diharapkan dari skripsi ini adalah dapat menambah pengetahuan tentang analisis regresi logistik kelas laten dalam rangka pengelompokan data untuk suatu pengambilan keputusan.

BAB II

LANDASAN TEORI

2.1 Asuransi

Menurut Kamus Besar Bahasa Indonesia, asuransi adalah pertanggung-an atau perjanjian antara dua pihak, dimana pihak yang satu berkewajiban membayar iuran dan pihak yang lain berkewajiban memberikan jaminan se-penuhnya kepada pembayar iuran apabila terjadi sesuatu yang menimpa pi-hak pertama atau barang miliknya sesuai dengan perjanjian yang dibuat. Se-dangkan menurut pasal 246 Kitab Undang-Undang Hukum Dagang (KUHD), asuransi atau pertanggung-an adalah suatu perjanjian, dimana seorang penang-gung mengikatkan diri kepada tertanggung dengan menerima suatu premi, untuk memberikan penggantian kepadanya karena suatu kerugian, kerusakan atau kehilangan keuntungan yang diharapkan, yang mungkin akan dideritanya karena suatu peristiwa yang tidak tertentu. Dari kedua definisi tersebut, da-pat disimpulkan bahwa asuransi merupakan sebuah usaha di bidang jasa yang memberikan jaminan atas risiko yang menimbulkan kerugian.

Menurut Otoritas Jasa Keuangan, asuransi merupakan perjanjian antara dua pihak, yaitu perusahaan asuransi dan pemegang polis, yang menjadi dasar bagi penerimaan premi oleh perusahaan asuransi sebagai imbalan untuk :

1. memberikan penggantian kepada tertanggung atau pemegang polis kare-na kerugian, kerusakan, biaya yang timbul, kehilangan keuntungan, atau tanggung jawab hukum kepada pihak ketiga yang mungkin diderita ter-tanggung atau pemegang polis karena terjadinya suatu peristiwa yang

tidak pasti, atau

2. memberikan pembayaran yang didasarkan pada meninggalnya tertanggung atau pembayaran yang didasarkan pada hidupnya tertanggung dengan manfaat yang besarnya telah ditetapkan dan/atau didasarkan pada hasil pengelolaan dana.

Terdapat beberapa istilah dalam asuransi yang perlu diketahui antara lain polis asuransi dan premi asuransi. Menurut Djojosoedarso dalam Nurnaningsih (2012), polis asuransi merupakan surat yang mengatur segala hak dan kewajiban masing-masing pihak (tertanggung dan penanggung). Sedangkan premi asuransi merupakan pembayaran dari tertanggung kepada penanggung, sebagai imbalan atas jasa atau pengalihan resiko kepada penanggung.

Beberapa manfaat yang dapat diperoleh dari asuransi menurut Djojosoedarso dalam Nurnaningsih (2012) antara lain :

1. Memberikan rasa aman
2. Melindungi keluarga dari perpecahan
3. Menghilangkan ketergantungan
4. Menjamin kehidupan wanita karier
5. Kontribusi terhadap pendidikan
6. Kontribusi terhadap lembaga-lembaga sosial
7. Memberikan manfaat untuk pemupukan kekayaan
8. Stimulasi menabung
9. Menyediakan dana yang dibutuhkan untuk investasi

2.2 Minat terhadap Asuransi

Menurut Kamus Besar Bahasa Indonesia (KBBI), kata minat memiliki arti kecenderungan hati yang tinggi terhadap sesuatu, gairah, dan keinginan. Witherington dalam Tricahyo (2012) mendefinisikan minat sebagai perhatian individu terhadap adanya suatu objek, seseorang, suatu soal atau situasi mengandung sangkut paut dengan dirinya yang dilakukan dengan sadar diikuti dengan perasaan senang.

Minat tidak tumbuh atau timbul tanpa adanya interaksi antara manusia dengan suatu objek. Hal ini berarti bahwa minat terbentuk dalam hubungan dengan suatu objek. Crow and Crow dalam Tricahyo (2012) menyatakan ada tiga faktor yang mendasari timbulnya minat seseorang, yaitu :

1. Faktor dorongan yang berasal dari dalam.

Kebutuhan ini dapat berupa kebutuhan dari dalam jasmani dan kejiwaan.

2. Faktor motif sosial.

Timbulnya minat seseorang dapat tumbuh dari motif sosial yaitu kebutuhan untuk mendapatkan penghargaan dari lingkungan dimana mereka berada.

3. Faktor emosional.

Faktor ini merupakan ukuran intensitas seseorang dalam menaruh perhatian terhadap sesuatu kegiatan atau objek tertentu.

Minat terhadap asuransi kemudian didefinisikan sebagai keinginan seseorang untuk memiliki asuransi. Keadaan ini berkaitan dengan proses dalam melakukan keputusan pembelian. Menurut Kotler dan Armstrong dalam

Nurnaningsih (2012), proses yang dilakukan seseorang untuk mengambil keputusan pembelian terdiri atas lima tahap yaitu :

1. Pengenalan masalah

Pengenalan masalah merupakan tahap pertama dimana seseorang mengenali suatu kebutuhan. Dalam tahap ini seseorang akan termotivasi untuk membeli atau memilih suatu produk untuk mengatasi masalah atau kebutuhan yang dihadapinya.

2. Pencarian informasi

Dalam tahap ini, konsumen yang telah tertarik akan suatu produk mulai mencari informasi lebih mengenai produk tersebut. Informasi tersebut dapat diperoleh melalui sumber mana pun, misalnya melalui teman, keluarga, iklan, media massa, atau melalui orang yang sudah memiliki pengalaman terhadap produk yang dimaksud.

3. Evaluasi berbagai alternatif

Pada tahap evaluasi berbagai alternatif, seseorang akan menggunakan informasi yang diperolehnya untuk mengevaluasi produk-produk alternatif dalam satu susunan pilihan.

4. Keputusan pembelian

Tahap keputusan pembelian merupakan tahap dalam proses pengambilan keputusan sampai konsumen benar-benar membeli produk.

5. Perilaku pasca pembelian

Pada tahap terakhir ini, seseorang akan mengambil tindakan lebih lanjut setelah membeli berdasarkan kepuasan atau ketidakpuasan yang dirasakan.

Minat terhadap asuransi merupakan variabel yang tidak dapat diukur secara langsung atau disebut variabel laten. Pengukuran terhadap variabel ini dilakukan melalui beberapa variabel indikator/*manifest* atau gejala yang dapat diukur. As'at Nasution dalam penelitiannya mengukur minat masyarakat Kabupaten Mandailing Natal terhadap jasa dan fasilitas lembaga asuransi melalui beberapa variabel antara lain kesadaran masyarakat akan kebutuhan terhadap asuransi, tingkat pengetahuan masyarakat terhadap jasa dan fasilitas asuransi, pemahaman terhadap produk asuransi, ketertarikan terhadap produk asuransi, serta kepercayaan terhadap lembaga penyedia jasa asuransi. Variabel-variabel indikator dalam penelitian tersebut disajikan dalam bentuk butir-butir pertanyaan dengan skala likert.

2.3 Regresi Logistik Multinomial

Regresi logistik multinomial merupakan salah satu metode statistik yang digunakan untuk menjelaskan hubungan variabel dependen yang bersifat multinomial (nominal dan ordinal) dengan satu atau lebih variabel independen (Agresti, 2007). Dalam analisis kelas laten, peluang seseorang berada pada suatu kelas laten jika diberikan nilai kovariatnya dibatasi mengikuti model dalam regresi logistik multinomial. Dengan kata lain, model regresi logistik multinomial digunakan untuk mengetahui pengaruh variabel kovariat terhadap hasil pengelompokan variabel laten. Model ini digunakan sebab dalam analisis kelas laten, variabel laten yang digunakan diasumsikan berasal dari sebaran multinomial.

Untuk mempermudah ilustrasi regresi logistik multinomial, dalam penulisan ini akan digunakan variabel independen berskala nominal dengan 3 kategori. Misalkan respon untuk variabel Y dikoding dengan 0, 1, dan 2. Dalam hal ini

dibutuhkan dua fungsi logit, dengan terlebih dahulu ditentukan kategori respon mana yang akan dijadikan acuan atau *baseline*. Pada umumnya, digunakan $Y = 0$ sebagai kategori acuan (*baseline category*) dan untuk membentuk fungsi logit yang membandingkan $Y = 1$ dan $Y = 2$ terhadap $Y = 0$. Misalkan terdapat p kovariat dan sebuah konstanta yang dinotasikan dengan vektor \mathbf{x} dengan panjang $p + 1$ dan $x_0 = 1$. Dua fungsi logit yang terbentuk adalah :

$$\begin{aligned} g_1(\mathbf{x}) &= \ln \left[\frac{P(Y = 1|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] \\ &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1p}x_p \\ &= \mathbf{x}'\beta_1 \end{aligned}$$

dan

$$\begin{aligned} g_2(\mathbf{x}) &= \ln \left[\frac{P(Y = 2|\mathbf{x})}{P(Y = 0|\mathbf{x})} \right] \\ &= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \dots + \beta_{2p}x_p \\ &= \mathbf{x}'\beta_2 \end{aligned}$$

Berdasarkan kedua fungsi logit di atas, maka didapatkan peluang bersyarat dari setiap kategori *outcome* jika diberikan vektor kovariat nya adalah :

$$\begin{aligned} P(Y = 0|\mathbf{x}) &= \pi_0(\mathbf{x}) = \frac{1}{1 + e^{g_1(\mathbf{x})} + e^{g_2(\mathbf{x})}}, \\ P(Y = 1|\mathbf{x}) &= \pi_1(\mathbf{x}) = \frac{e^{g_1(\mathbf{x})}}{1 + e^{g_1(\mathbf{x})} + e^{g_2(\mathbf{x})}}, \end{aligned}$$

dan

$$P(Y = 2|\mathbf{x}) = \pi_2(\mathbf{x}) = \frac{e^{g_2(\mathbf{x})}}{1 + e^{g_1(\mathbf{x})} + e^{g_2(\mathbf{x})}}.$$

Mengikuti model dalam regresi logistik biner, notasikan $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$ untuk $j = 0, 1, 2$.

Secara umum, model multinomial logit dapat dituliskan menjadi bentuk peluang respon sebagai berikut :

$$P(Y = j|\mathbf{x}) = \pi_j(\mathbf{x}) = \frac{\exp(\alpha_j + \beta_j'\mathbf{x})}{\mathbf{1} + \sum_{h=2}^J \exp(\alpha_h + \beta_h'\mathbf{x})} \quad (2.1)$$

dengan

$$\pi_j(\mathbf{x}) = \text{peluang respon kategori } j \text{ jika diberikan nilai kovariat } x,$$

$$\text{dan } \sum_{j=1}^J \pi_j(\mathbf{x}) = 1$$

$$\beta_j = \text{vektor koefisien regresi pada kategori respon ke-}j$$

$$\alpha = \text{nilai konstanta persamaan regresi}$$

$$\mathbf{x} = \text{vektor kovariat}$$

$$J = \text{jumlah kategori respon}$$

Dalam model regresi logistik multinomial, nilai konstanta dan koefisien regresi untuk kategori respon yang menjadi acuan memiliki nilai nol. Dalam persamaan (2.1) di atas, kategori respon acuan yang digunakan merupakan kategori respon pertama, sehingga $\alpha_1 = 0$ dan $\beta_1 = 0$.

Interpretasi Parameter Regresi Logistik Multinomial

Pada pemodelan regresi logistik, interpretasi parameter bertujuan untuk mengetahui arti dari nilai taksiran parameter pada variabel independen. Terdapat dua jenis variabel independen, yaitu variabel yang bersifat kategorik dan variabel yang bersifat kontinu.

Interpretasi parameter regresi logistik multinomial untuk variabel independen yang bersifat kategorik, dapat dilakukan dengan menggunakan nilai *odds* dan *odds ratio* (Hosmer dan Lemeshow, 2000). Misalkan terdapat dua kejadian, yaitu kejadian 1 dan kejadian 2, nilai *odds* dari kejadian satu berkaitan dengan kejadian 2 didefinisikan sebagai peluang dari kejadian 1 dibagi dengan peluang dari kejadian 2. Sebagai contoh, jika hanya terdapat dua kemungkinan kejadian, seperti dalam pelemparan sebuah koin. Misalkan kejadian yang dimaksud merupakan munculnya sisi gambar, maka kejadian 1 merupakan kejadian mun-

culnya sisi gambar dan kejadian 2 merupakan kejadian tidak munculnya sisi gambar. Jika kejadian 1 mempunyai probabilitas $P(\text{Kejadian1})$ untuk muncul, maka nilai *odds* didefinisikan sebagai berikut :

$$\text{odds} = \frac{P(\text{kejadian 1})}{1 - P(\text{kejadian 1})} \quad (2.2)$$

Dalam regresi logistik, nilai *odds* (peluang) diartikan sebagai probabilitas suatu kategori respon dibagi dengan probabilitas kategori respon acuan. Jika diberikan nilai kovariat $x = a$, dan kategori respon acuan J , maka nilai *odds* didefinisikan sebagai :

$$\text{odds} = \frac{\pi_j(x = a)}{1 - \pi_J(x = a)}$$

Sedangkan rasio peluang (*odds ratio*) merupakan sekumpulan peluang (*odds*) yang dibagi dengan peluang lainnya. *odds ratio* didefinisikan sebagai rasio dari nilai *odds* untuk $x = a$ terhadap $x = b$, yaitu :

$$\text{odds ratio} = \frac{\{\pi(a) / [1 - \pi(a)]\}}{\{\pi(b) / [1 - \pi(b)]\}}$$

Nilai *odds ratio* diinterpretasikan sebagai kecendrungan sebuah kategori respon akan terjadi diantara objek dengan kovariat a dibandingkan dengan kovariat b . Dalam regresi logistik multinomial, nilai *odds ratio* untuk kategori respon $Y = j$ dibandingkan dengan kategori respon acuan $Y = J$ untuk kovariat $x = a$ dibandingkan $x = b$ didefinisikan sebagai :

$$\begin{aligned} \text{Odds Ratio}_j &= \frac{P(Y = j|x = a)/P(Y = J|x = a)}{P(Y = j|x = b)/P(Y = J|x = b)} \\ &= \frac{\pi_j(x = a)}{\pi_J(x = b)} \end{aligned}$$

Misalkan diperoleh nilai *odds ratio* sebesar 2, nilai ini diinterpretasikan sebagai peluang kategori respon j dibandingkan J akan terjadi adalah 2 kali lebih besar untuk $x = a$ dibandingkan $x = b$.

Cara yang dapat digunakan untuk menginterpretasi parameter regresi logistik dari variabel kontinu adalah dengan mengasumsikan fungsi logit linier

terhadap variabel independen. Misalkan variabel independen bersifat kontinu dengan fungsi logit $g(x) = \beta_0 + \beta_1 x_1$, interpretasi β_1 memiliki sifat sama dengan parameter pada regresi linier. Setiap kenaikan satu unit x , nilai $g(x)$ naik sebesar β_1 .

2.4 Teorema Bayes

Probabilitas untuk melakukan klasifikasi dalam analisis kelas laten diperoleh dengan menggunakan teorema Bayes. Sebelum mengetahui bunyi teorema Bayes, terlebih dahulu akan diberikan definisi peluang bersyarat, yaitu :

Definisi 2.4.1. Misalkan A dan B adalah kejadian dalam ruang sampel S dengan $P(B) \neq 0$, maka peluang terjadinya kejadian A setelah atau ketika kejadian B terjadi dirumuskan sebagai berikut :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (2.3)$$

Teorema 2.4.1. Jika A sembarang himpunan bagian dari S dan B_1, B_2, \dots, B_m adalah partisi dari S . Maka untuk $i = 1, 2, \dots, m$ dan $j = 1, 2, \dots, m$ berlaku :

$$P(B_j|A) = \frac{P(A|B_j) P(B_j)}{\sum_{i=1}^m P(A|B_i) P(B_i)}$$

Bukti. Misalkan himpunan $B_i \neq \emptyset$ dengan $i = 1, 2, \dots, m$, merupakan partisi dari ruang sampel S , yang bersifat :

1. $B_i \cap B_j \neq \emptyset$ untuk $i \neq j$; $i = 1, 2, 3, \dots, m$; $j = 1, 2, 3, \dots, m$
2. $\bigcup_{i=1}^m B_i = S$
3. $P(B_i) > 0 \forall i$

Jika A merupakan himpunan bagian S , dengan $P(A) \neq 0$, maka A dapat dipandang sebagai $A = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_m) = \bigcup_{i=1}^m (A \cap B_i)$ dimana masing-masing $(A \cap B_i)$ saling lepas secara berpasangan. Dengan demikian peluang A dapat ditulis sebagai :

$$\begin{aligned}
 P(A) &= P[(A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_m)] \\
 &= P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_m) \\
 &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_m)P(B_m) \\
 &= \sum_{i=1}^m P(A|B_i)P(B_i) \tag{2.4}
 \end{aligned}$$

Menurut definisi peluang bersyarat pada persamaan (2.2), diketahui bahwa untuk setiap j berlaku :

$$P(B_j|A) = \frac{P(B_j \cap A)}{P(A)}$$

Peluang $P(B_j \cap A)$ dapat ditulis sebagai berikut :

$$P(B_j \cap A) = P(A \cap B_j) = P(A|B_j)P(B_j)$$

Sehingga,

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} \tag{2.5}$$

Selanjutnya, substitusi persamaan (2.3) pada persamaan (2.4) menghasilkan :

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^m P(A|B_i)P(B_i)} \tag{2.6}$$

□

Himpunan B_1, B_2, \dots, B_m selanjutnya membentuk m hipotesis prior yang digunakan untuk mempertimbangkan peristiwa A . Peluang $P(B_i)$ selanjutnya disebut sebagai peluang *prior*, sedangkan $P(B_j|A)$ disebut sebagai peluang *posterior* untuk hipotesis yang sama.

2.5 Analisis Kelas Laten

Analisis kelas laten (*Latent Class Analysis*) merupakan sebuah metode statistik yang digunakan untuk mengelompokkan beberapa individu atau objek menjadi beberapa kelas dengan menggunakan variabel kategorik. Variabel dalam analisis kelas laten terdiri dari dua jenis, yaitu variabel laten dan variabel indikator. Variabel laten adalah variabel yang tidak dapat diukur secara langsung, sedangkan variabel indikator merupakan variabel manifest (nyata) yang dapat diukur secara langsung, misalnya dari skor respon subjek terhadap instrumen pengukuran (Haerudin, 2013).

Dalam analisis kelas laten, variabel laten yang digunakan bersifat kategorik. Variabel ini terdiri dari sebuah himpunan kelas-kelas laten yang diukur melalui variabel indikator (Collins & Lanza, 2010). Sebagai contoh, Coffman (2007) melakukan pengelompokan terhadap motivasi siswa SMA untuk minum minuman beralkohol. Dalam hal ini, variabel laten yang digunakan merupakan motivasi untuk minum minuman beralkohol, yang terdiri dari 4 kelas laten, yaitu kelompok siswa SMA yang termotivasi karena ingin mencoba alkohol, kelompok siswa SMA yang termotivasi karena ingin mencari jati diri, kelompok siswa SMA yang termotivasi karena ingin *relax*, dan yang terakhir adalah kelompok siswa SMA yang termotivasi karena semua alasan tersebut. Dalam penelitian tersebut, Coffman mengukur motivasi untuk minum minuman beralkohol menggunakan data hasil kuisisioner dari *Monitoring The Future*.

2.5.1 Model Kelas Laten

Model Kelas Laten merupakan model matematika yang menghubungkan peluang respon suatu individu untuk variabel-variabel indikator dengan suatu variabel laten yang bersifat kategorik. Model kelas laten pertama kali

diperkenalkan oleh Lazaferd dan Henry pada tahun 1968 dan kemudian dikembangkan oleh Goodman pada tahun 1974. Model tersebut merupakan bentuk khusus dari model kelas laten dimana variabel indikator yang digunakan bersifat kategorik. Model ini merupakan bagian dari *finite mixture* model yang menggunakan fungsi peluang posterior sebagai basis penggerombolannya (Vermunt dan Magidson, 2002).

Model kelas laten dinyatakan dengan sebaran gabungan dari variabel yang diamati pada data yang memiliki berbagai macam skala pengukuran. Misalkan $j = 1, 2, \dots, J$ merupakan variabel atau variabel indikator yang memiliki kategori respon $r_j = 1, 2, \dots, R_j$, (y_1, y_2, \dots, y_j) dinotasikan sebagai vektor dari j variabel indikator dimana setiap variabel indikator tersebut memiliki sebaran bersyarat dalam keluarga eksponensial seperti Bernoulli, Poisson, multinomial, dan normal. Misalkan y_{ij} adalah nilai respon dari $i = 1, 2, \dots, N$ sampel objek untuk variabel ke- j . Vektor baris $y'_i = (y_{1i}, y_{2i}, \dots, y_{ji})$ merupakan bentuk respon dari i objek. Maka model kelas laten didefinisikan sebagai

$$f(y_i|\theta) = \sum_{k=1}^K \pi_k g(y_{ji}|\theta_{jk}) \quad (2.7)$$

dengan

$f(y_i|\theta)$ = fungsi peluang bersama untuk variabel- variabel pengamatan

$g(y_{ji}|\theta_{jk})$ = sebaran y_j yang diberikan oleh parameter model θ

π_k = peluang awal kelas laten K pada data y , dimana

$$\pi_k = \frac{n_k}{N}$$

K = banyaknya kelas ($k = 1, 2, \dots, K$) dan $\sum_{k=1}^K \pi_k = 1$

Peluang suatu individu atau objek akan berada pada kelas laten k jika diketahui individu tersebut memberikan respon y_i untuk variabel indikator y_{ji}

yang dinotasikan dengan $P(k|Y_i)$, diperoleh melalui formula Bayes, yaitu :

$$P(k|Y_i) = \frac{\pi_k g(y_i|\theta_{jk})}{\sum_{k=1}^K \pi_k g(y_i|\theta_{jk})} \quad (2.8)$$

Sebaran y_{ji} dalam persamaan model kelas laten (2.7) di atas diasumsikan mengikuti skala dari variabel indikator yang akan digunakan. Sehingga, model kelas laten untuk masing-masing variabel indikator dengan skala yang berbeda terdapat dalam uraian berikut.

1. Variabel Biner

Pada kasus dimana variabel y_j berbentuk biner (0 dan 1), sebaran mengikuti distribusi Bernoulli berganda dengan bentuk:

$$g(y_{ji}|\alpha_k) = \prod_{j=1}^B \alpha_{jk}^{y_{ji}} (1 - \alpha_{jk})^{1-y_{ji}}$$

Dengan melakukan substitusi persamaan di atas ke persamaan (2.7), maka model kelas laten untuk variabel biner adalah

$$f(y_i|\alpha_k) = \sum_{k=1}^K \pi_k g(y_{ji}|\alpha_k) = \sum_{k=1}^K \pi_k \prod_{j=1}^B \alpha_{jk}^{y_{ji}} (1 - \alpha_{jk})^{1-y_{ji}} \quad (2.9)$$

dengan

π_k = peluang awal kelas laten K pada data y , dimana

$$\pi_k = \frac{n_k}{N} \text{ dan } \sum_{k=1}^K \pi_k = 1$$

α_{jk} = peluang suatu objek memberikan respon untuk variabel indikator ke- j jika berada pada kelas ke- k dan $\sum_{k=1}^K \alpha_{jk} = 1$

B = banyaknya variabel biner yang digunakan dalam penelitian

2. Variabel Nominal

Pada kasus variabel nominal berbentuk *polytomus*, variabel indikator y_j digantikan oleh sebuah nilai vektor fungsi indikator yang didefinisikan sebagai

$$y_{j(r)} = \begin{cases} 1, & \text{jika respon pada variabel } j \text{ jatuh di } r, r = 1, 2, \dots, R \\ 0, & \text{untuk lainnya} \end{cases}$$

dimana R merupakan jumlah kategori variabel ke j dan $\sum_{r=1}^R y_{j(r)} = 1$

Pola respon dari objek ke- i ditulis sebagai $y'_i = (y_{1i}, y_{2i}, \dots, y_{ji})$ yang merupakan dimensi dari R . Peluang respon tunggal α_{jk} dari kasus biner diganti dengan suatu himpunan dari fungsi $\alpha_{jk(r)} (r = 1, 2, \dots, R)$. Sebaran gabungan untuk variabel nominal ini menggunakan distribusi multinomial dengan bentuk:

$$g(y_{ji}|\alpha_k) = \prod_{j=1}^C \prod_{r=1}^R (\alpha_{jk(r)})^{y_{ji(s)}}$$

maka model kelas laten untuk variabel nominal adalah

$$f(y_i|\alpha_k) = \sum_{k=1}^K \pi_k g(y_{ji}|\alpha_k) = \sum_{k=1}^K \pi_k \prod_{j=1}^C \prod_{r=1}^R (\alpha_{jk(r)})^{y_{ji(s)}} \quad (2.10)$$

dengan

$\alpha_{jk(r)}$ = peluang objek akan memberikan respon ke- r untuk variabel

indikator ke- j jika berada pada kelas ke- k , dan $\alpha_{jk(r)} = \frac{n_{j(r)}}{n_k}$

r = kategori jawaban yang digunakan pada variabel ke- j

$(r = 1, 2, \dots, R)$

K = banyaknya kelas yang terbentuk ($k = 1, 2, \dots, K$)

C = banyaknya variabel nominal yang digunakan dalam penelitian

3. Variabel Ordinal

Pada kasus variabel ordinal, variabel indikator y_j didefinisikan sama dengan kasus variabel nominal. Urutan kategori L memiliki peluang

$\alpha_{k(1)}, \alpha_{k(2)}, \dots, \alpha_{k(L)}$ yang merupakan fungsi dari kelas laten K . Model peluang kumulatif untuk kategori r adalah

$$\gamma_{jk(r)} = \alpha_{jk(1)} + \alpha_{jk(2)} + \dots + \alpha_{jk(r)}$$

dimana $k = 1, 2, \dots, K$ dan $r = 1, 2, \dots, L$.

Sebaran bersyarat untuk variabel indikator y_{ji} mengikuti distribusi multinomial yang diberikan oleh:

$$g(y_{ij}|\alpha_k) = \prod_{r=1}^L \alpha_{jk(r)}^{y_{ji(r)}} = \prod_{j=1}^O \prod_{r=1}^L (\gamma_{jk(r)} - \gamma_{jk(r-1)})^{y_{ji(r)}}$$

Maka model kelas laten untuk variabel ordinal memiliki bentuk:

$$f(y_i|\alpha_k) = \sum_{k=1}^K \pi_k g(y_{ji}|\alpha_k) = \sum_{k=1}^K \pi_k \prod_{j=1}^O \prod_{r=1}^L (\gamma_{jk(r)} - \gamma_{jk(r-1)})^{y_{ji(r)}} \quad (2.11)$$

dengan

$\alpha_{jk(r)}$ = peluang objek akan memberikan respon ke- r untuk variabel

indikator ke- j jika berada di kelas k , dan $\alpha_{jk(r)} = \frac{n_{j(r)}}{n_k}$

r = kategori jawaban yang digunakan pada variabel ke- j

($r = 1, 2, \dots, L$)

K = banyaknya kelas yang terbentuk ($k = 1, 2, \dots, K$)

$\gamma_{jk(r)}$ = peluang kumulatif respon akan berada di kelas ke- k untuk kategori r pada variabel ke- j

O = banyaknya variabel ordinal yang digunakan dalam penelitian

4. Variabel Kontinu

Pada kasus dimana variabel indikator y_j berbentuk kuantitatif, sebaran diasumsikan berbentuk normal dengan bentuk :

$$g(y_{ji}|\mu_k, \sigma_j^2) = (2\pi)^{-\frac{1}{2}} \sigma_j^{-\frac{1}{2}} \exp \left[-\frac{1}{2\sigma_j^2} (y_{ji} - \mu_{jk})^2 \right] \quad (2.12)$$

Fungsi normal ganda dari sebaran peluangnya adalah

$$f(y_i|\mu_k, \sigma_j^2) = \sum_{k=1}^K \pi_k g(y_{ji}|\mu_k, \sigma_j^2) \quad (2.13)$$

Dengan mensubstitusi persamaan (2.11) ke persamaan (2.12), maka model kelas laten untuk variabel kuantitatif adalah

Dengan :

- μ_{jk} = parameter lokasi dari variabel kontinu y_j dalam kelas ke- k
- σ_j^2 = ragam dari variabel ke- j yang diambil secara konstan dari data.

2.5.2 Estimasi Parameter Model Kelas Laten

Metode yang digunakan untuk menduga nilai parameter yang terdapat dalam model kelas laten adalah metode pendugaan *maximum likelihood* (ML) yang dilanjutkan dengan algoritma Ekspektasi Maksimum (EM) dan metode Newton Raphson. Fungsi *log-likelihood* yang disyaratkan pada pendekatan EM dapat diturunkan dari fungsi kepadatan peluang yang mendefinisikan model. Fungsi *likelihood* untuk N individu yang diobservasi dalam model kelas laten adalah

$$L = \prod_{i=1}^N f(y_i|\theta) = \prod_{i=1}^N \sum_{k=1}^K \pi_k g(y_{ji}|\theta_{jk}) \quad (2.14)$$

Sehingga fungsi *log-likelihood* dari model kelas laten tersebut dapat dinyatakan sebagai berikut:

$$\ln L = \ln \prod_{i=1}^N \sum_{k=1}^K \pi_k g(y_{ji}|\theta_{jk}) = \sum_{i=1}^N \ln \sum_{k=1}^K \pi_k g(y_{ji}|\theta_{jk}) \quad (2.15)$$

Setelah mendapatkan fungsi *log-likelihood* untuk model kelas laten, selanjutnya dilakukan pendugaan parameter menggunakan algoritma Ekspektasi Maksimum (EM). Algoritma EM merupakan salah satu algoritma penggerombolan yang termasuk dalam kategori *partitional clustering* (proses pengelom-

pokan yang dilakukan dengan membagi objek berdasarkan kemiripan masing-masing objek) dan berbasiskan model yang menggunakan perhitungan peluang (Utami, 2009). Proses pendugaan pada algoritma EM dimulai dengan iterasi dari EM. Setiap lingkaran proses pada algoritma EM terdiri dari dua tahapan, yaitu tahapan ekspektasi dan tahapan pemaksimuman. Berikut merupakan langkah yang dilakukan untuk pendugaan parameter dengan algoritma EM :

1. Mendefinisikan nilai awal $\theta^{(0)} = (\pi^0, \alpha^0, \alpha_r^0, \text{ dan } (\gamma_r - \gamma_r - 1)^0)$
2. Menghitung nilai dari sebaran peluang bersama, misalkan urutan pengulangan dinotasikan m :

$$f(y_i|\theta) = \sum_{k=1}^K \pi_k^{(m)} g(y_{ji}|\theta_{jk})^{(m)}$$

3. Tahapan E : menghitung $\hat{P}(k, y_i)^{(m)}$ dengan $i = 1, 2, \dots, N; k = 1, 2, \dots, K$, dimana $\hat{P}(k, Y_i)^{(m)}$ merupakan peluang bersyarat yang menyatakan y_i muncul dari kelas k . Adapun rumusan tersebut adalah

$$\hat{P}(k|Y_i)^{(m)} = \frac{\hat{\pi}_k^{(m)} g(y_i|\hat{\theta}_{jk})^{(m)}}{\sum_{k=1}^K \hat{\pi}_k^{(m)} g(y_i|\hat{\theta}_{jk})^{(m)}}$$

4. Tahapan M : menghitung nilai maksimum *likelihood* dengan mengasumsikan parameter sama dengan nilai dugaan dari tahapan ekspektasi, sehingga diperoleh penduga parameter yang baru.

$$\hat{\pi}_k^{(m+1)} = \frac{\sum_{i=1}^N \hat{P}(k|Y_i)^{(m)}}{N}$$

Pendugaan peluang bersyarat yaitu $y_j = 1$ diberikan kelas K untuk variabel biner :

$$\hat{\alpha}_k^{(m+1)} = \frac{\sum_{i=1}^N y_{ij} \hat{P}(k|Y_i)^{(m)}}{n \hat{\pi}_k}$$

Pendugaan peluang bersyarat yaitu $y_j = r$ diberikan kelas K untuk variabel nominal :

$$\hat{\alpha}_{jk(r)}^{(m+1)} = \frac{\sum_{i=1}^N y_{ij(r)} \hat{P}(k|Y_i)^{(m)}}{n\hat{\pi}_k}$$

Pendugaan kumulatif peluang bersyarat dengan $y_j \leq r$ pada kelas ke- K untuk variabel ordinal :

$$\hat{\gamma}_{jk(r)}^{(m+1)} = \hat{\gamma}_{jk(r-1)}^{(m+1)} + \frac{\sum_{i=1}^N y_{ij(r)} \hat{P}(k|Y_i)^{(m)}}{n\hat{\pi}_k^{(m)}}$$

5. Ulangi langkah 2,3,4 sampai konvergen yaitu ketika $|\hat{\pi}_k^{(m+1)} - \hat{\pi}_k^{(m)}|$ dan $|\hat{\alpha}_k^{(m+1)} - \hat{\alpha}_k^{(m)}|$, $k = 1, 2, \dots, K$ mendekati nol.

Ketika algoritma EM telah memiliki solusi yang optimal, program dialihkan ke metode Newton-Raphson. Hal ini dilakukan untuk memanfaatkan keuntungan dari kedua algoritma ini, yaitu stabilitas algoritma EM walaupun ketika solusi jauh dari optimum, dan kecepatan metode Newton-Raphson ketika solusi sudah mendekati optimum. Dalam pengoptimalan nilai melalui program Newton-Raphson ini, parameter-parameter dalam model kelas laten diperbaharui menjadi

$$\hat{\vartheta}^t = \hat{\vartheta}^{t-1} - \varepsilon H^{-1}g$$

dimana

ϑ = himpunan parameter dalam model kelas laten

g = gradien vektor yang terdiri dari turunan pertama *log-posterior* untuk seluruh parameter pada saat $\hat{\vartheta}^{t-1}$

H = matriks Hessian yang terdiri dari turunan kedua dari seluruh parameter

ε = skalar yang menotasikan ukuran tahapan.

Metode Newton-Raphson

Quinn (dalam Ersyad, 2016) menyatakan bahwa metode Newton-Raphson merupakan metode iteratif yang dapat digunakan untuk menghitung estimasi maksimum *likelihood*. Prosedur iteratif dari metode ini dilakukan sampai konvergen, sehingga didapatkan estimasi parameter yang stabil.

Algoritma pokok dari metode Newton-Raphson ini dapat diuraikan sebagai berikut :

1. Misalkan x_n adalah pendekatan awal pada langkah ke- n , $n = 0, 1, 2, \dots$
2. Hitung gradien garis singgung terhadap kurva $y = f(x)$ di titik $(x_n, f(x_n))$, yakni $f'(x_n)$ setelah itu tentukan persamaan garis singgungnya, yaitu $y = f'(x_n)(x - x_n) + f(x_n)$.
3. Pendekatan selanjutnya adalah absis titik potong garis singgung tersebut dengan sumbu- x , yakni

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Apabila variabel atau parameternya berdimensi tinggi, maka fungsi turunan pertamanya berupa vektor T dan turunan keduanya berupa matriks yang disebut matriks Hessian (H). Bentuk multivariat dari metode *Newton-Raphson* adalah sebagai berikut :

$$x_1 = x_0 - H(x_0)^{-1}T(x_0)$$

Misalkan $x = (x_1, x_2, \dots, x_m)$, maka vektor T dan matriks *Hessian* H adalah sebagai berikut :

$$T = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{bmatrix}, \quad H(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \frac{\partial^2 f}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_m^2} \end{bmatrix}$$

2.6 Teknik Pengambilan Sampel

Dalam penelitian kali ini, teknik pengambilan sampel yang dilakukan adalah teknik *probability sampling* yaitu *proportional random sampling*. Teknik ini merupakan teknik pengambilan sampel yang memberikan peluang yang sama bagi setiap anggota populasi yang dipilih untuk menjadi sampel dengan memperhatikan kategori dalam populasi penelitian. Langkah-langkah yang dilakukan dalam teknik *proportional random sampling* adalah

1. Penentuan jumlah sampel keseluruhan menggunakan rumus Slovin (Yamane, 1967) yaitu

$$n = \frac{N}{1 + N(e)^2} \quad (2.16)$$

dengan :

n = ukuran sampel

N = ukuran populasi

e = persen kelonggaran ketidaktelitian karena kesalahan pengambilan sampel yang masih dapat diinginkan, yaitu 10%.

2. Perhitungan jumlah sampel yang diinginkan setiap proporsi dengan menggunakan rumus alokasi proporsional

$$n_i = \frac{N_i}{N} \times n$$

dengan :

n_i = ukuran sampel setiap proporsi

N_i = ukuran populasi setiap proporsi

N = ukuran populasi seluruhnya

2.7 Validitas dan Reliabilitas

Untuk mengukur suatu variabel yang digunakan dalam penelitian diperlukan sebuah alat ukur yang biasa disebut instrumen. Instrumen merupakan suatu alat yang dipergunakan sebagai alat untuk mengukur suatu objek ukur atau mengumpulkan data dari suatu variabel. Suatu instrumen dikatakan baik apabila valid dan reliabel. Instrumen yang tidak valid dan reliabel bila digunakan dalam penelitian akan menghasilkan data yang sulit dipercaya. Untuk itu perlu dilakukan uji validitas dan reliabilitas instrumen penelitian.

Azwar dalam Matondang (2009) menyatakan bahwa validitas berasal dari kata *validity* yang berarti sejauh mana tingkat ketepatan dan kecermatan suatu instrumen pengukur (tes) dalam melakukan fungsi ukurnya. Suatu instrumen dikatakan valid mengandung arti bahwa alat ukur yang digunakan untuk mendapatkan data tersebut, dapat digunakan untuk mengukur apa yang hendak diukur. Suatu penelitian memiliki validitas internal bila terdapat kesesuaian antara bagian-bagian instrumen dengan instrumen secara keseluruhan. Pengujian validitas internal sebuah instrumen dapat dilakukan dengan menggunakan analisis butir, yaitu dengan mengkorelasikan skor pada *item* dengan skor pada total *item*nya. Perhitungan ini dilakukan dengan menggunakan korelasi *product moment* (r). Skor *item* dianggap sebagai nilai X sedangkan skor total dianggap sebagai nilai Y . Apabila skor *item* memiliki skor positif yang signifikan artinya *item* tersebut dapat digunakan sebagai indikator untuk mengukur variabel penelitian.

Hipotesis yang digunakan dalam uji validitas setiap *item* adalah :

$H_0 : \rho = 0$ (skor *item* tidak berkorelasi positif dengan skor total *item* atau *item* pertanyaan tidak valid)

$H_1 : \rho > 0$ (skor *item* berkorelasi positif dengan skor total *item* atau *item*)

pertanyaan valid)

Statistik uji yang digunakan adalah :

$$r = \frac{n \sum_{i=1}^n x_i y_i - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}} \quad (2.17)$$

dengan,

r = koefisien korelasi

X = skor *item*

Y = skor total tiap *item*

n = jumlah pengamatan

$i = 1, 2, \dots, n$

Kriteria pengujian yang dilakukan adalah tolak H_0 jika nilai $r_{hitung} \geq r_{\alpha, n-2}$ atau jika nilai probabilitas $\leq 0,05$. Setelah melakukan uji validitas, *item-item* yang valid akan diteruskan dengan uji reliabilitas, sedangkan *item* yang tidak valid akan dikeluarkan dan tidak digunakan.

Reliabilitas berasal dari kata *reliability* yang berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Suatu hasil pengukuran dikatakan reliabel apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subjek yang sama, diperoleh hasil pengukuran yang relatif sama. Djaali dalam Metondang (2009) menyatakan bahwa reliabilitas dibedakan menjadi dua macam, yaitu reliabilitas konsistensi tanggapan, dan reliabilitas konsistensi gabungan *item*.

Reliabilitas konsistensi tanggapan menitikberatkan pada konsistensi tanggapan responden terhadap instrumen yang dicobakan. Terdapat tiga cara yang dapat dilakukan untuk mengetahui reliabilitas konsistensi tanggapan, yaitu dengan memberikan instrumen yang sama beberapa kali dengan waktu yang

berbeda kepada responden yang sama, mencobakan instrumen dengan dua pertanyaan yang setara (memiliki maksud yang sama namun dengan bahasa berbeda), dan mencobakan instrumen dengan dua pertanyaan setara pada responden yang sama beberapa kali dengan waktu yang berbeda.

Sedangkan reliabilitas konsistensi gabungan berkaitan dengan konsistensi antara butir suatu instrumen. Artinya bahwa terhadap bagian subjek yang sama, hasil ukur butir yang satu tidak kontradiksi dengan hasil ukur butir yang lain. Salah satu teknik yang dapat digunakan untuk melakukan pengujian reliabilitas ini adalah teknik Alpha Cronbach. Adapun rumusan untuk teknik ini adalah :

$$\alpha = \frac{k}{k-1} \left\{ 1 - \frac{\sum_{i=1}^k S_i^2}{S_{total}^2} \right\} \quad (2.18)$$

dengan,

k = banyak *item* kuisisioner dalam atribut

S_i^2 = ragam dari *item* kuisisioner ke- i

S_{total}^2 = total ragam dari keseluruhan *item*

i = 1, 2, ..., k

Menurut Guilford (1956) kereliabilitasan instrumen dikriteriakan menjadi :

1. $\alpha < 0,2$: instrumen tidak reliabel
2. $0,2 \leq \alpha < 0,4$: reliabilitas instrumen kecil
3. $0,4 \leq \alpha < 0,7$: instrumen cukup reliabel
4. $0,7 \leq \alpha < 0,9$: instrumen reliabel
5. $0,9 \leq \alpha < 1,0$: instrumen sangat reliabel
6. $\alpha = 1,0$: reliabilitas instrumen sempurna.

Dari kriteria diatas, dapat disimpulkan bahwa suatu instrumen dikatakan reliabel bila nilai koefisien reliabilitasnya $\leq 0,7$.

2.8 Pemilihan Model Terbaik

Berdasarkan kelas atau model yang dihasilkan dalam analisis, maka akan dipilih model terbaik yang mampu menjelaskan data. Pemilihan kelas atau model terbaik ini dapat dilakukan dengan menggunakan nilai *Akaike's Information Criteria* (AIC). Model terbaik merupakan model yang menampilkan kecocokan dengan data yang optimal dan memenuhi prinsip *parsimony* (Collins, 2010). Prinsip ini merupakan prinsip filosofis yang menyatakan bahwa model yang lebih sederhana lebih baik dibandingkan dengan model yang lebih kompleks (Collins, 2010). Sederhana dalam konteks ini berarti melakukan pendugaan parameter yang lebih sedikit. Berdasarkan prinsip ini, model statistika harus tidak lebih rumit, maksudnya adalah tidak lebih banyak parameter yang diduga dibandingkan dengan yang dibutuhkan untuk merepresentasikan data. Oleh karena itu, model terbaik dipilih berdasarkan nilai AIC terkecil (Vermunt & Magidson, 2001). Nilai AIC terkecil menunjukkan bahwa model yang terbentuk mampu merepresentasikan data secara optimal dan memenuhi prinsip *parsimony*.

Nilai AIC didefinisikan sebagai :

$$AIC = -2[\max \ln(L)] + 2m$$

dimana m merupakan banyaknya parameter dan L adalah nilai *likelihood* yang telah optimal.

2.9 Asumsi Kebebasan Lokal

Analisis kelas laten memiliki asumsi dasar yang harus dipenuhi, yaitu asumsi *local independence*. Asumsi ini mensyaratkan bahwa antar variabel indikator haruslah saling bebas pada suatu kelas laten tertentu. Dengan kata lain, jika dimungkinkan untuk membuat tabel kontingensi yang terpisah untuk variabel indikator yang mengacu pada masing-masing kelas laten, variabel indikator tersebut akan saling bebas (*independent*) dalam masing-masing tabel kontingensi tersebut. Dalam penelitian kali ini, diasumsikan bahwa asumsi kebebasan lokal terpenuhi untuk setiap variabel indikator dalam suatu kelas laten.

BAB III

PEMBAHASAN

3.1 Analisis Regresi Logistik Kelas Laten

Analisis regresi logistik kelas laten merupakan pengembangan dari analisis kelas laten atau *latent class analysis*. Berbeda dengan analisis kelas laten yang melakukan pengelompokan objek hanya berdasarkan variabel indikator yang digunakan, analisis regresi logistik kelas laten mengelompokkan objek pengamatan berdasarkan indikator (*symptom*) sekaligus melibatkan kovariat atau variabel penjelas untuk pengelompokannya (Linzer & Lewis, 2011). Kovariat dalam analisis regresi logistik kelas laten digunakan untuk memprediksi keanggotaan kelas laten. Dalam model analisis kelas laten biasa, diasumsikan bahwa setiap objek/individu memiliki peluang awal kelas laten yang sama, sedangkan dalam model regresi logistik kelas laten peluang awal kelas laten suatu individu diijinkan untuk bervariasi bergantung dari nilai kovariat yang dimilikinya. Model regresi logistik multinomial yang diterapkan dalam analisis ini sebenarnya sama dengan model regresi logistik multinomial biasa, hanya saja variabel independen yang digunakan adalah laten bukan variabel yang teramati langsung. Model ini digunakan karena variabel laten dalam analisis regresi logistik kelas laten diasumsikan memiliki sebaran multinomial. Peluang seseorang berada pada suatu kelas laten jika diketahui nilai kovariatnya diasumsikan mengikuti model pada regresi logistik multinomial. Model tersebut akan dijelaskan pada subbab berikut.

3.1.1 Model Regresi Logistik Kelas Laten

Model analisis kelas laten yang diterapkan dalam Persamaan (2.7) merupakan model kelas laten yang tidak melibatkan kovariat. Dalam analisis regresi logistik kelas laten, pengelompokan objek ke dalam kelas-kelas laten melibatkan kovariat, sehingga model umumnya adalah

$$f(y_i|x_{pi}, \theta) = \sum_{k=1}^K (\pi_k|x_{pi}) \prod_{j=1}^J g(y_{ji}|\theta_{jk}) \quad (3.1)$$

dengan :

$f(y_i|x_{pi}, \theta_{jk})$ = fungsi peluang bersama untuk variabel-variabel pengamatan

$g(y_{ji}|\theta_{jk})$ = sebaran y_j yang diberikan oleh parameter model θ

x_{pi} = nilai kovariat ke- p objek ke i

$(\pi_k|x_i)$ = peluang awal kelas laten K jika diberikan kovariat x_i ,

dimana $\pi_k = \frac{n_k}{N}$ dan $\sum_{k=1}^K \pi_k = 1$

K = banyaknya kelas ($k = 1, 2, \dots, K$)

Kovariat memberikan kemampuan untuk membedakan peubah endogen sebagai indikator dari peubah laten dengan peubah eksogen yang digunakan untuk memprediksi kelas yang dimiliki oleh suatu objek (Vermunt dan Magidson, 2002).

Diketahui bahwa dalam analisis kelas laten, variabel laten diasumsikan memiliki distribusi multinomial, sehingga bentuk peluang seseorang untuk dapat masuk ke kelas laten k jika diberikan nilai kovariat nya dapat diparameterisasi ke dalam model multinomial logit seperti dalam Persamaan (2.1). Parameter β adalah koefisien dalam regresi logistik yang menggunakan kovariat X untuk memprediksi keanggotaan kelas laten. Misalkan terdapat p variabel kovariat, dalam hal ini, parameter $(\pi_k|x_{pi})$ dapat dituliskan sebagai fungsi dari

parameter β yaitu :

$$\pi_k(x_{pi}; \beta) = \frac{\exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)}{\sum_{k=1}^K \exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)} = \frac{\exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)}{1 + \sum_{k=2}^K \exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)} \quad (3.2)$$

Perhatikan bahwa kedua bentuk model dalam persamaan di atas memiliki nilai yang sama. Hal ini dikarenakan asumsi bahwa kelas laten pertama digunakan sebagai kelas acuan. Kelas acuan memiliki nilai β yang sama dengan 0.

Dari Persamaan (3.2) di atas, model fungsi logit didefinisikan sebagai

$$\begin{aligned} \text{logit}[(\pi_k|x_{pi})] &= \log\left(\frac{(\pi_k|x_{pi})}{1 - (\pi_k|x_{pi})}\right) \\ &= \log(\pi_k|x_{pi}) - \log(1 - (\pi_k|x_{pi})) \\ &= \log\left(\frac{\exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)}{1 + \sum_{k=2}^K \exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)}\right) \\ &\quad - \log\left(\frac{1}{1 + \sum_{k=2}^K \exp\left(\sum_{p=1}^P \beta_{kp}x_{pi}\right)}\right) \\ &= \sum_{p=1}^P \beta_{kp}x_{pi} - \sum_{k=2}^K \left(\sum_{p=1}^P \beta_{kp}x_{pi}\right) + \sum_{k=2}^K \left(\sum_{p=1}^P \beta_{kp}x_{pi}\right) \\ &= \sum_{p=1}^P \beta_{kp}x_{pi} \\ &= \beta_{k0}x_{0i} + \beta_{k1}x_{1i} + \dots + \beta_{kP}x_{Pi} \end{aligned} \quad (3.3)$$

Rumus Bayes dari peluang suatu objek akan masuk dalam kelas ke- k dalam regresi logistik kelas laten yaitu :

$$\hat{P}(k|X_{pi}, Y_i) = \frac{\pi_k(X_{pi}, \hat{\beta})g(y_i|\hat{\theta}_{jk})}{\sum_{k=1}^K \pi_k(X_{pi}, \hat{\beta})g(y_i|\hat{\theta}_{jk})} \quad (3.4)$$

Jumlah parameter yang diestimasi dalam model regresi logistik kelas laten adalah $K \sum_{j=1}^J (R_j - 1) + (P + 1)(K - 1)$ dimana K merupakan banyaknya kelas yang terbentuk, R_j merupakan jumlah kategori respon pada variabel indikator ke- j dan P merupakan banyaknya kovariat yang digunakan.

Pendugaan parameter model logit dilakukan secara bersamaan dengan pendugaan parameter model kelas laten dengan menggunakan algoritma EM. Setiap nilai dugaan parameter yang dihasilkan kemudian diuji tingkat signifikansinya dengan menggunakan hipotesis :

$H_0 : \beta_k = 0$ (koefisien logit tidak berpengaruh terhadap model)

$H_1 : \exists \beta_k \neq 0$ (ada koefisien logit yang berpengaruh terhadap model)

Uji hipotesis tersebut dilakukan dengan menggunakan statistik uji *Wald Chi-Square* yang didefinisikan sebagai $W = \left(\frac{\hat{\beta}}{SE(\hat{\beta})} \right)^2$ dengan kriteria uji, tolak H_0 jika nilai $W > \chi_{p,\alpha}^2$ atau $p - value \leq \alpha$.

3.1.2 Estimasi Parameter

Fungsi kepadatan peluang dalam model analisis regresi logistik kelas laten tidak jauh berbeda dengan persamaan (2.14), hanya saja π_k diganti dengan $\pi_k(x_{pi}, \beta)$. Fungsi kepadatan peluang tersebut adalah :

$$f(y_i | x_i, \theta_{jk}) = \sum_{k=1}^K \pi_k(x_{pi}, \beta) g(y_{ji} | \theta_{jk})$$

Fungsi likelihood dari persamaan di atas adalah :

$$L = \prod_{i=1}^N f(y_i | x_{pi}, \theta_{jk}) = \prod_{i=1}^N \sum_{k=1}^K \pi_k(x_{pi}, \beta) g(y_{ji} | \theta_{jk})$$

Sedangkan fungsi log-likelihood nya adalah :

$$\ln L = \ln \prod_{i=1}^N \sum_{k=1}^K \pi_k(x_{pi}, \beta) g(y_{ji} | \theta_{jk}) = \sum_{i=1}^N \ln \sum_{k=1}^K \pi_k(x_{pi}, \beta) g(y_{ji} | \theta_{jk})$$

Pada penelitian ini akan digunakan 12 variabel indikator bertipe ordinal dengan 3 kategori. Pemilihan 3 kategori ini dikarenakan semakin banyak kategori

yang digunakan, maka perbedaan masing-masing kategori tidak akan terlalu terlihat. Sehingga model kelas laten berdasarkan kovariat yang terbentuk adalah :

$$\begin{aligned} f(y_i|x_{pi}, \theta_{jk}) &= \sum_{k=1}^K (\pi_k|x_{pi}) g(y_{ji}|\alpha_k) \\ &= \sum_{k=1}^K (\pi_k|x_{pi}) \prod_{j=1}^{12} \prod_{r=1}^3 (\gamma_{jk(r)} - \gamma_{jk(r-1)})^{y_{ji(r)}} \end{aligned}$$

Proses estimasi parameter model regresi logistik kelas laten tidak jauh berbeda dengan model kelas laten, hanya saja terdapat penambahan kovariat atau variabel penjelas dalam estimasinya. Adapun langkah-langkah estimasi parameter tersebut adalah :

1. Mendefinisikan nilai awal $\theta^{(0)} = ((\gamma_{jk(r)} - \gamma_{jk(r-1)})^0, \beta_k^0)$
2. Menghitung nilai dari sebaran peluang bersama, misalkan urutan pengulangan dinotasikan m :

$$\pi_k(x_i, \beta) = \frac{\exp(\beta_k x_i)}{1 + \sum_{k=1}^K \exp(\beta_k x_i)}$$

$$f(y_i|\alpha_k) = \sum_{k=1}^K \pi_k(x_i, \beta)^{(m)} g(y_{ji}|\alpha_k)^{(m)}$$

3. Tahapan E : menghitung $\hat{P}(k|X_i, Y_i)^{(m)}$ dengan $i = 1, 2, \dots, N; k = 1, 2, \dots, K$, dimana $\hat{P}(k|x_i, y_i)^{(m)}$ merupakan peluang bersyarat yang menyatakan y_i muncul dari kelas k . Adapun rumusan tersebut adalah

$$\hat{P}(k|X_i, Y_i)^{(m)} = \frac{\hat{\pi}_k(x_i, \hat{\beta})^{(m)} g(y_i|\hat{\alpha}_k)^{(m)}}{\sum_{k=1}^K \hat{\pi}_k(x_i, \hat{\beta})^{(m)} g(y_i|\hat{\alpha}_k)^{(m)}}$$

4. Tahapan M : menghitung nilai maksimum *likelihood* dengan mengasumsikan parameter sama dengan nilai dugaan dari tahapan ekspektasi, sehingga diperoleh penduga parameter yang baru. Pendugaan kumulatif

peluang bersyarat dengan $y_j \leq r$ pada kelas ke- K untuk peubah ordinal

:

$$\hat{\gamma}_{jk(r)}^{(m+1)} = \hat{\gamma}_{jk(r-1)}^{(m+1)} + \frac{\sum_{i=1}^N y_{ij(r)} \hat{P}(k|Y_i)^{(m)}}{n \hat{\pi}_k^{(m)}}$$

5. Ulangi langkah 2,3,4 sampai konvergen yaitu ketika $|\hat{\gamma}_{jk(r)}^{(m+1)} - \hat{\gamma}_{jk(r)}^{(m)}|$ dan $|\hat{\beta}_k^{(m+1)} - \hat{\beta}_k^{(m)}|$, $k = 1, 2, \dots, K$ mendekati nol.

Ketika algoritma EM telah memiliki solusi yang optimal, program dialihkan ke metode Newton-Raphson. Dalam pengoptimalan nilai melalui program Newton-Raphson ini, parameter-parameter dalam model kelas laten diperbaharui menjadi

$$\hat{\beta}_k^{(m+1)} = \hat{\beta}_k^{(m)} - H_{\beta}^{-1} g$$

dimana

β_k = himpunan parameter dalam model kelas laten

g = gradien vektor yang terdiri dari turunan pertama *log-posterior* untuk seluruh parameter pada saat $\hat{\beta}_k^{t-1}$

H = matriks Hessian yang terdiri dari turunan kedua dari seluruh parameter

ε = skalar yang menotasikan ukuran tahapan.

3.1.3 Interpretasi Model Regresi Logistik Kelas Laten

Interpretasi model regresi logistik kelas laten dilakukan dengan memperhatikan nilai dari *odds* dan *odds ratio* seperti melakukan interpretasi pada analisis regresi logistik. Nilai intersep (β_0) dalam regresi logistik dapat ditransformasi kedalam nilai *odds*, sedangkan koefisien regresi yang lain (β_1, β_2 , dst) dapat ditransformasi kedalam nilai *odds ratio* dengan melakukan eksponensial terhadap koefisien-koefisien tersebut (e^{β_0}, e^{β_1} , dst)

Untuk melakukan interpretasi dalam analisis regresi logistik kelas laten, nilai π_k dalam Persamaan (3.2) terlebih dahulu diubah kedalam bentuk nilai *odds*. Misalkan kelas laten ke- K adalah kategori referensinya, maka nilai *odds* kelas laten ke- k dalam hubungannya dengan kelas laten K dapat didefinisikan sebagai berikut :

$$\frac{\pi_k}{\pi_K} \quad (3.5)$$

Dengan nilai log dari *odds* yang disebut sebagai logit adalah :

$$\text{logit} = \log \left(\frac{\pi_k}{\pi_K} \right) \quad (3.6)$$

Hasil fungsi logit tersebut terdapat dalam Persamaan (3.3).

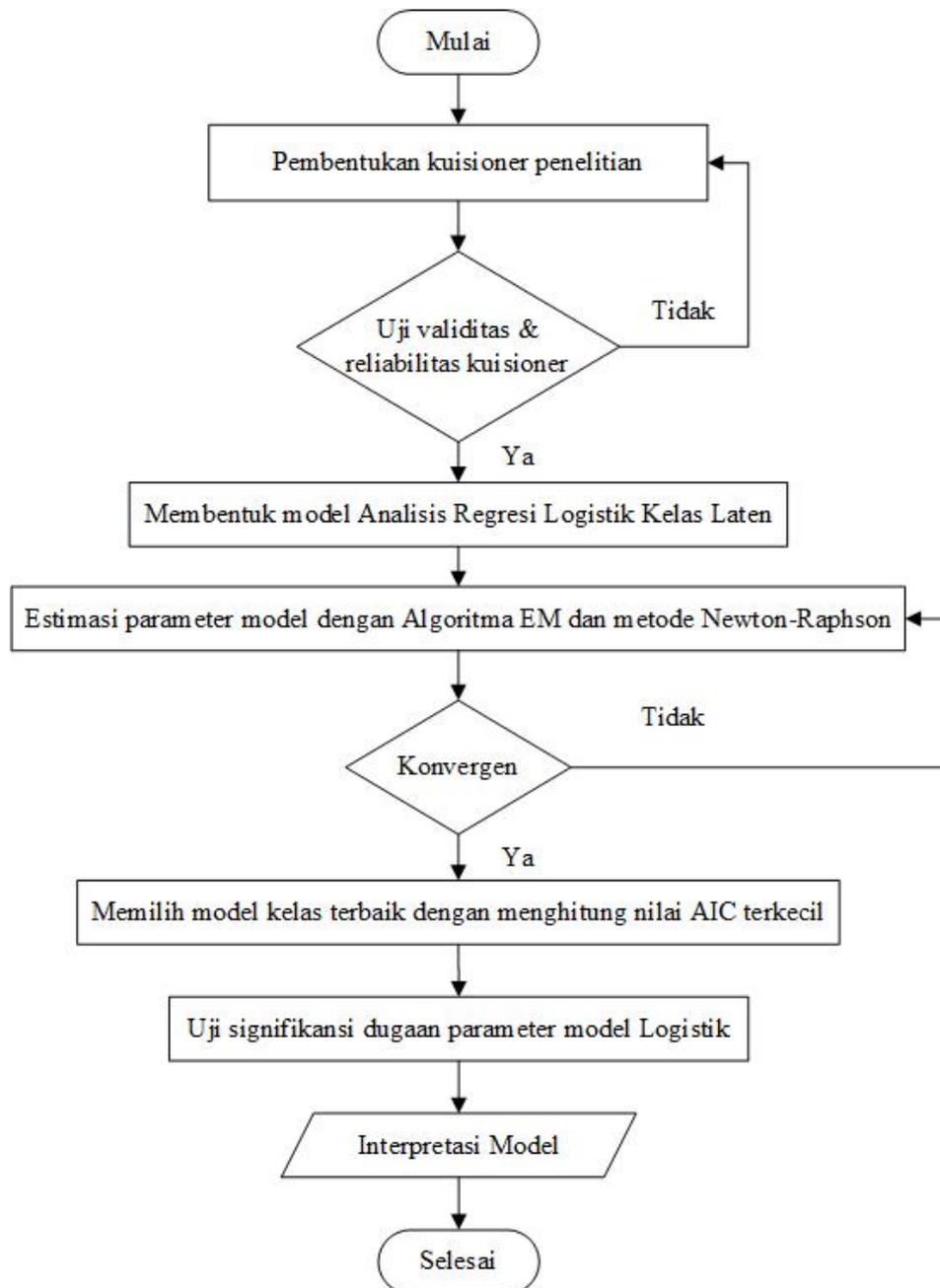
Untuk mempermudah dalam melakukan interpretasi dalam analisis logistik kelas laten, diberikan contoh berikut. Misalkan dalam pengelompokan kebiasaan sehat masyarakat terdapat dua kelas, yaitu kelas laten sehat dan kelas laten tidak sehat. Misalkan digunakan kovariat jenis kelamin, dengan $x = 1$ untuk laki-laki, dan $x = 0$ untuk perempuan, dengan nilai *odds* sama dengan 2 dan digunakan kelas laten tidak sehat sebagai referensi, interpretasi yang dapat dilakukan adalah kecenderungan perempuan masuk ke dalam kelas laten sehat, 2 kali lebih besar dibandingkan masuk ke kelas laten tidak sehat.

3.2 Diagram Alir Analisis Regresi Logistik Kelas Laten

Diagram alir dari analisis regresi logistik kelas laten terdapat dalam Gambar 3.1. Adapun langkah-langkah dalam melakukan analisis regresi logistik kelas laten adalah sebagai berikut :

1. Pembentukan kuisioner penelitian berdasarkan variabel indikator dan skala yang digunakan

2. Melakukan uji validitas dan reliabilitas terhadap instrumen kuisioner. Butir pertanyaan yang valid diteruskan pada waktu uji reliabilitas, sedangkan butir yang tidak valid dihilangkan dan tidak digunakan
3. Membentuk model regresi logistik kelas laten berdasarkan nilai peluang dari semua sampel. Model regresi logistik kelas laten yang terbentuk dicobakan untuk beberapa alternatif kelas, dimulai dengan model dengan 2 kelas laten, 3 kelas laten, dan seterusnya.
4. Melakukan estimasi parameter untuk kelas laten dan parameter model logit menggunakan algoritma EM. Dalam model analisis regresi logistik kelas laten, nilai parameter kelas laten merupakan nilai $(\pi_k|x_{pi})$, sedangkan parameter model logit merupakan nilai β_k
5. Ketika Algoritma EM telah memiliki solusi optimal, tahapan dilanjutkan ke metode Newton-Raphson
6. Memilih model kelas terbaik dengan menggunakan nilai AIC terkecil. Setelah diperoleh nilai AIC untuk model dengan 2 kelas dan 3 kelas, dipilih nilai AIC terkecil untuk memperoleh model kelas terbaik. Nilai ini menunjukkan bahwa model atau kelas yang terbentuk mampu merepresentasikan data secara optimal.
7. Menguji signifikansi nilai dugaan parameter regresi logistik kelas laten dengan *Wald Chi-Square Statistic*. Jika nilai probabilitas untuk masing-masing parameter regresi logistik kurang dari $\alpha = 0.05$ maka kovariat yang dimaksud memiliki pengaruh terhadap pengelompokan.
8. Interpretasi hasil analisis logistik kelas laten dengan menggunakan *odds ratio*.



Gambar 3.1: Diagram Alir Analisis Regresi Logistik Kelas Laten

3.3 Identifikasi Variabel

Variabel yang digunakan dalam penelitian ini terdiri dari variabel laten, kelompok variabel kovariat (penjelas) dan kelompok variabel indikator. Variabel laten yang digunakan dalam penelitian ini adalah minat masyarakat terhadap asuransi. Adapun rincian kelompok variabel indikator dan variabel kovariat yang digunakan terdapat dalam Tabel 3.1 dan 3.2 :

Tabel 3.1: Variabel Indikator yang Digunakan

Variabel Indikator	Keterangan	Tipe	Respon
Kesadaran tingkat kebutuhan	Adanya peristiwa tidak pasti yang menimbulkan kerugian	Ordinal	1 : Tidak Setuju 2 : Ragu-Ragu 3 : Setuju
	Keinginan memiliki dana untukantisipasi kerugian	Ordinal	1 : Tidak Setuju 2 : Ragu-Ragu 3 : Setuju
Pengetahuan tentang asuransi	Manfaat asuransi sebagai pengalih resiko	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
	Istilah pihak tertanggung & penanggung	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
	Istilah polis dalam asuransi	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
	Istilah premi dalam asuransi	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
Pemahaman tentang asuransi	Prosedur melakukan klaim	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
	Produk asuransi beragam	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
	Produk asuransi dapat dibeli di Bank	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
	Pemegang polis tidak selalu pihak tertanggung	Ordinal	1 : Salah 2 : Ragu-Ragu 3 : Benar
Ketertarikan akan asuransi	Persepsi bahwa asuransi lebih menguntungkan dibandingkan menabung/investasi	Ordinal	1 : Tidak Setuju 2 : Ragu-Ragu 3 : Setuju
	Persepsi bahwa asuransi memiliki banyak manfaat dibanding menabung/investasi	Ordinal	1 : Tidak Setuju 2 : Ragu-Ragu 3 : Setuju
	Kesediaan menyisihkan sebagian penghasilan untuk membayar premi dibanding menabung	Ordinal	1 : Tidak Setuju 2 : Ragu-Ragu 3 : Setuju

Tabel 3.2: Variabel Kovariat yang Digunakan

Variabel Kovariat	Keterangan	Respon	Tipe Data
X_1	Jenis Kelamin	1: Laki-Laki	Biner
		2: Perempuan	
X_2	Usia	-	Rasio
X_3	Tingkat Pendidikan	1: SD/Sederajat	Ordinal
		2: SMP/Sederajat	
		3: SMA/Sederajat	
		4: Diploma/S1	
		5: Pasca Sarjana	
X_4	Pendapatan	1: Di Bawah UMR	Ordinal
		2: UMR - Rp. 7.000.000	
		3: >Rp 7.000.000	

3.4 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data primer yang diperoleh melalui kuisisioner mengenai minat masyarakat terhadap asuransi. Pendapat masyarakat inilah yang kemudian akan dijadikan variabel indikator untuk dasar penelitian menggunakan analisis regresi logistik kelas laten. Kuisisioner ini diberikan kepada masyarakat di wilayah Jakarta Timur yang berusia di atas 18 tahun dan telah memiliki penghasilan tetap. Adapun rincian kuisisioner yang disebarkan terdapat pada lampiran 1.

Berdasarkan data Badan Pusat Statistik Provinsi DKI Jakarta, jumlah penduduk yang berada di wilayah Jakarta Timur tahun 2015 sebanyak 2.923.502 jiwa, yang terdiri dari 1.481.483 penduduk laki-laki dan 1.442.019 penduduk perempuan. Pengambilan sampel pada penelitian ini menggunakan teknik *proportional random sampling*. Kategori yang digunakan dalam pengambilan sampel pada penelitian ini merupakan jenis kelamin. Adapun tahapan yang dilakukan dalam pengambilan sampel adalah sebagai berikut. Tahap pertama dalam teknik ini adalah menentukan jumlah sampel keseluruhan menggunakan rumus Slovin (Yamane, 1967) seperti dalam Persamaan (2.16). Dengan presentase kelonggaran kesalahan 10%, diperoleh jumlah keseluruhan sampel minimal yang harus diambil adalah :

$$\begin{aligned}
 n &= \frac{N}{1 + N(e)^2} \\
 &= \frac{2.923.502}{1 + 2.923.502(0,10)^2} \\
 &= \frac{2.923.502}{29.236,02} \\
 &= 99,997 \\
 &\approx 100
 \end{aligned}$$

Tahap kedua adalah perhitungan jumlah sampel yang diinginkan setiap strata. Strata yang digunakan dalam penelitian ini merupakan jenis kelamin, sehingga jumlah masing-masing sampel untuk laki-laki dan perempuan adalah sebagai berikut.

Sampel penduduk laki-laki :

$$\begin{aligned}
 n_1 &= \frac{N_1}{N} \times n \\
 &= \frac{1.481.483}{2.923.502} \times 100 \\
 &= 50,67 \\
 &\approx 51
 \end{aligned}$$

Sampel penduduk perempuan :

$$\begin{aligned}
 n_2 &= \frac{N_2}{N} \times n \\
 &= \frac{1.442.019}{2.923.502} \times 100 \\
 &= 49,33 \\
 &\approx 49
 \end{aligned}$$

Dari perhitungan di atas, jumlah sampel minimal yang dibutuhkan adalah 100 penduduk. Jumlah ini terdiri dari sampel laki-laki sebanyak 51 penduduk dan sampel perempuan sebanyak 49 penduduk. Berdasarkan data yang

diperoleh dari proses pengisian kuisioner, dalam penelitian ini akan digunakan 100 sampel yang merupakan penduduk di wilayah Jakarta Timur. Data hasil pengisian kuisioner ini terdapat dalam lampiran 2.

3.5 Uji Validitas dan Reliabilitas

Setelah melakukan identifikasi variabel indikator dan variabel kovariat yang akan digunakan, tahap selanjutnya adalah melakukan uji validitas dan reliabilitas instrumen. Pengujian ini dilakukan terhadap item-item pertanyaan dalam kuisioner penelitian. Dalam penelitian ini, uji validitas dan reliabilitas dilakukan terhadap item-item pertanyaan yang mengukur minat masyarakat terhadap asuransi, atau dengan kata lain merupakan variabel indikator yang digunakan. Sedangkan variabel kovariat yang digunakan dalam penelitian ini tidak perlu diuji validitas dan reliabilitasnya, hal ini dikarenakan variabel-variabel tersebut dapat diketahui secara langsung tanpa instrumen.

Pengujian validitas dan reliabilitas ini dilakukan terhadap 30 responden sebelum kuisioner diedarkan secara resmi. Tujuannya adalah untuk menguji apakah item-item pertanyaan yang digunakan untuk mengukur variabel minat masyarakat terhadap asuransi sudah valid dan reliabel atau belum. Analisis ini dimulai dengan melakukan uji validitas terlebih dahulu, baru kemudian dilanjutkan dengan uji reliabilitas.

Hipotesis :

H_0 : item pertanyaan tidak valid

H_1 : item pertanyaan valid

Statistik Uji : Persamaan (2.16)

Kriteria Uji : Tolak H_0 jika $r_{hitung} \geq r_{\alpha, n-2}$ atau jika nilai probabilitas ≤ 0.05 .

Digunakan nilai $\alpha = 5\%$

Dengan sampel 30 responden dan $\alpha = 5\%$, maka derajat bebas yang digunakan adalah $n - 2 = 28$ sehingga nilai $r_{\alpha, n-2} = r_{0.05, 28} = 0.3610$. Hasil uji validitas dan reliabilitas dengan menggunakan SPSS 16 dalam penelitian ini dapat dilihat dalam lampiran 3. Adapun rangkuman hasil uji validitas terhadap item-item pertanyaan dalam kuisioner yang akan digunakan adalah sebagai berikut.

Tabel 3.3: Hasil Uji Validitas

Variabel Indikator	r_{hitung}	Probabilitas	Keterangan
Y_1	0.528	0.003	Valid
Y_2	0.389	0.034	Valid
Y_3	0.756	0.000	Valid
Y_4	0.420	0.210	Valid
Y_5	0.576	0.001	Valid
Y_6	0.412	0.024	Valid
Y_7	0.701	0.000	Valid
Y_8	0.414	0.023	Valid
Y_9	0.287	0.124	Tidak Valid
Y_{10}	0.640	0.000	Valid
Y_{11}	0.584	0.001	Valid
Y_{12}	0.624	0.000	Valid
Y_{13}	0.704	0.000	Valid

Dari tabel di atas, terlihat bahwa tidak semua item pertanyaan dalam kuisioner tersebut valid. Item pertanyaan yang berkaitan dengan variabel indikator ke-9 yaitu, pemahaman bahwa produk asuransi dapat dibeli di Bank tidak valid. Untuk itu variabel ini akan digugurkan dalam kuisioner.

Setelah melakukan pengujian validitas terhadap item-item pertanyaan atau variabel indikator yang digunakan dalam kuisioner, tahap selanjutnya adalah melakukan uji reliabilitas terhadap item-item pertanyaan yang valid. Dari hasil uji validitas di atas, maka item-item atau variabel indikator yang akan dilakukan uji reliabilitasnya adalah seluruh variabel indikator, kecuali variabel indikator ke-9. Dengan menggunakan persamaan (2.17) diperoleh

nilai $\alpha_{Cronbach} = 0.810$, maka menurut aturan Guilford, disimpulkan bahwa item-item pertanyaan yang dibentuk dalam kuisisioner sudah reliabel atau dapat dikatakan bahwa kuisisioner yang dibagikan dapat memberikan data yang konsisten.

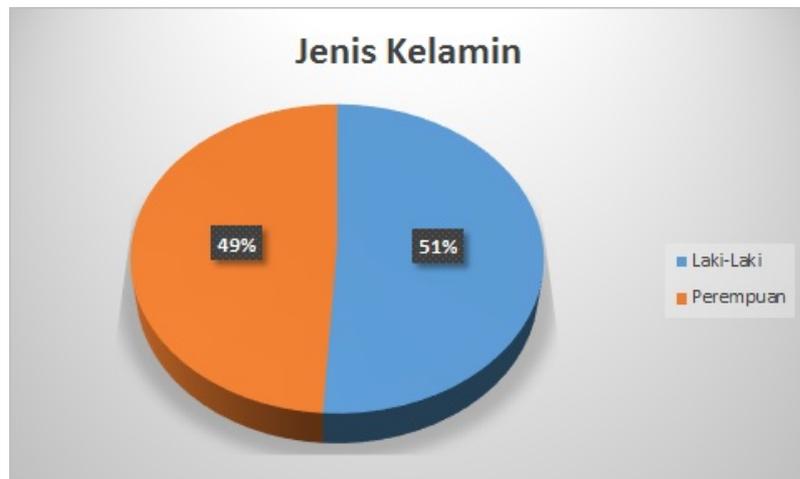
Dari uraian di atas dapat disimpulkan bahwa kuisisioner yang dibentuk sudah valid dan reliabel dengan menggugurkan variabel indikator pemahaman bahwa produk asuransi dapat dibeli di Bank. Selanjutnya, penelitian ini akan dilakukan dengan menghilangkan variabel pemahaman bahwa produk asuransi dapat dibeli di Bank.

3.6 Hasil Analisis Data

3.6.1 Deskripsi Data

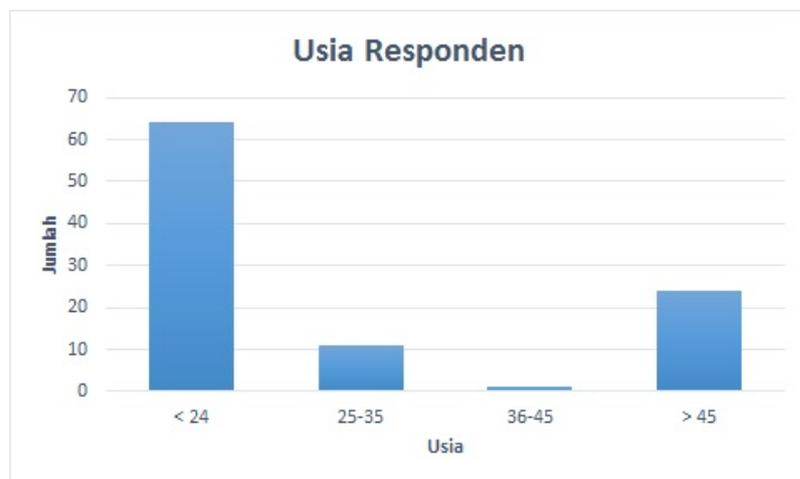
Setelah melakukan uji validitas dan reliabilitas terhadap kuisisioner, selanjutnya kuisisioner tersebut diberikan kepada penduduk Jakarta Timur yang berusia lebih dari 18 tahun dan telah memiliki penghasilan tetap. Adapun data hasil pengisian kuisisioner yang dilakukan oleh 100 responden dapat dilihat dalam lampiran 2.

Karakteristik yang dimiliki oleh 100 responden dalam penelitian ini terdapat dalam gambar diagram-diagram di bawah. Gambar 3.2 menjelaskan karakteristik responden berdasarkan jenis kelamin yang dimiliki. Dalam diagram di atas terlihat bahwa responden dalam penelitian ini responden paling banyak merupakan responden yang diwakili oleh warna biru. Responden ini merupakan responden laki-laki dengan presentase sebesar 51%. Sedangkan responden perempuan yang diwakili oleh warna jingga memiliki presentase sebesar 49%.



Gambar 3.2: Karakteristik Responden berdasarkan Jenis Kelamin

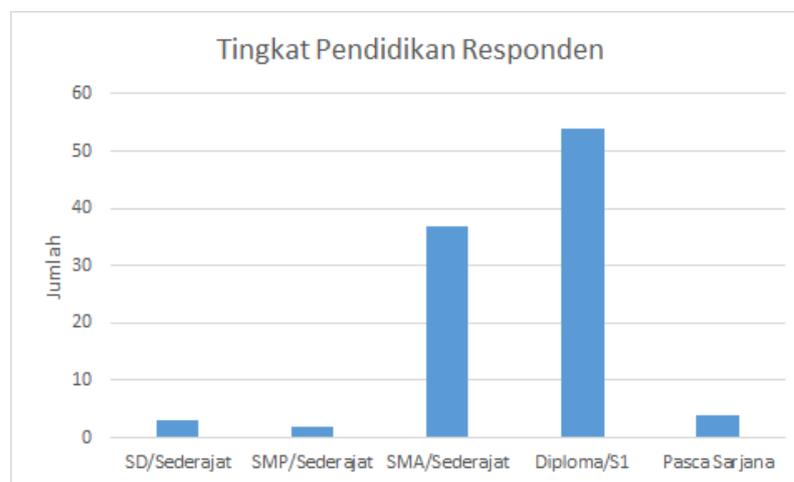
Karakteristik kedua yang dapat dijelaskan dalam penelitian ini merupakan usia yang dimiliki responden. Adapun sebaran usia responden tersebut terdapat dalam gambar 3.3.



Gambar 3.3: Karakteristik Responden berdasarkan Usia

Dalam diagram di atas, terlihat bahwa responden terbanyak memiliki usia kurang dari 24 tahun dengan jumlah responden sebanyak 64 orang. Responden dengan usia lebih dari 45 tahun menempati urutan kedua dengan jumlah responden sebanyak 24 tahun. Sedangkan untuk responden dengan usia antara 24-35 tahun menempati urutan ketiga dengan jumlah 11 responden, dan jumlah responden yang paling sedikit memiliki usia antara 36-45 tahun. Hal ini menunjukkan bahwa dalam penelitian ini, pengelompokan minat masyarakat terhadap asuransi akan dilakukan terhadap usia muda dan usia tua. Kemudian akan dilihat apakah terdapat pengaruh pengelompokan minat masyarakat usia muda dengan usia tua.

Karakteristik responden selanjutnya merupakan tingkat pendidikan responden. Karakteristik tersebut terdapat dalam gambar 3.4.

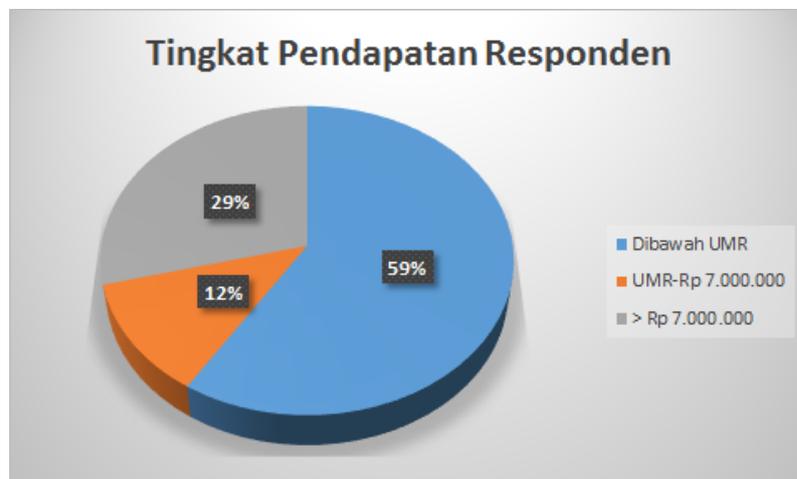


Gambar 3.4: Karakteristik Responden berdasarkan Tingkat Pendidikan

Dalam diagram di atas terlihat bahwa responden terbanyak memiliki tingkat pendidikan diploma atau S1 dengan jumlah responden sebanyak 54 orang. Selanjutnya, tingkat pendidikan dengan jumlah responden terbanyak dimiliki oleh SMA/Sederajat dengan jumlah 37 responden, diikuti oleh responden

dengan tingkat pendidikan pasca sarjana, SD/ sederajat, dan SMP/ sederajat. Hal ini menunjukkan bahwa dalam penelitian ini pengelompokan akan dilakukan untuk masyarakat dengan tingkat pendidikan yang cukup tinggi. Kemudian akan dilihat apakah terdapat pengaruh tingkat pendidikan terhadap pengelompokan minat masyarakat terhadap asuransi.

Karakteristik terakhir yang dapat dijelaskan adalah tingkat pendapatan. Hal tersebut dapat dilihat dalam gambar 3.5 berikut.



Gambar 3.5: Karakteristik Responden berdasarkan Tingkat Pendapatan

Dalam diagram di atas terlihat bahwa sebagian besar responden memiliki tingkat pendapatan dibawah UMR (Upah Minimum Regional) dengan presentase sebesar 59%. Responden yang memiliki pendapatan antara UMR-Rp 7.000.000 memiliki presentase 12%. Sedangkan responden dengan tingkat pendapatan lebih dari Rp 7.000.000 memiliki presentase sebesar 20%. Hal ini menunjukkan bahwa dalam penelitian ini akan dilakukan pengelompokan minat masyarakat terhadap asuransi untuk tingkat pendapatan rendah, sedang, dan tinggi. Setelah itu, akan dilihat apakah terdapat pengaruh tingkat pendidikan ini terhadap pengelompokan minat masyarakat terhadap asuransi.

3.6.2 Estimasi Peluang Awal Kelas Laten

Estimasi atau penaksiran peluang awal kelas laten diperlukan untuk mengklasifikasikan objek ke dalam kelas laten. Peluang awal (prior) ini merupakan nilai peluang banyaknya objek dalam suatu kelas terhadap keseluruhan kelas. Estimasi ini dilakukan dengan menggunakan metode maksimum *likelihood* yang dibantu dengan algoritma EM untuk memperoleh solusi optimal. Estimasi peluang awal kelas laten ini dilakukan untuk model dengan 2 dan 3 kelas dengan nilai *log-likelihood* maksimum. Hasil perhitungan nilai *log-likelihood* terbaik atau maksimal untuk model 2 kelas laten dan 3 kelas laten dengan bantuan *software* R terdapat dalam lampiran 4 dan 6. Adapun hasil estimasi yang diperoleh terdapat dalam Tabel 3.4 berikut :

Tabel 3.4: Peluang Awal Kelas Laten (π_k)

Model	π_1	π_2	π_3
2 Kelas	0.7476	0.2524	
3 Kelas	0.4305	0.2930	0.2766

Dari Tabel 3.3, terlihat bahwa untuk model dengan 2 kelas, estimasi peluang awal untuk kelas ke-1 (π_1) dan kelas ke-2 (π_2) masing-masing adalah 0.7476 dan 0.2524. Angka ini menyatakan bahwa sebanyak 74.76% dari responden yang dianalisis akan masuk ke kelas pertama dan sebanyak 25.24% responden akan masuk ke kelas kedua. Sedangkan untuk model dengan 3 kelas, estimasi peluang awal untuk kelas ke-1, kelas ke-2, dan kelas ke-3 berturut-turut adalah 0.4305, 0.2930, dan 0.2766. Interpretasi untuk ketiga nilai ini sama seperti model untuk 2 kelas. Estimasi model dengan 2 kelas dan 3 kelas selengkapnya terdapat dalam lampiran 5 dan 7.

3.6.3 Pemilihan Model Terbaik

Setelah dilakukan pengujian untuk model dengan 2 kelas laten dan 3 kelas laten, selanjutnya akan ditentukan model yang cocok untuk menggambarkan minat responden terhadap asuransi. Pemilihan ini didasarkan pada nilai AIC terkecil. Hasil perhitungan nilai AIC untuk masing-masing model terdapat dalam Tabel 3.5 berikut :

Tabel 3.5: Nilai AIC Untuk Setiap Model

Model	AIC	log-likelihood	Npar
2 Kelas	1849.265	-871.6325	53
3 Kelas	1785.630	-810.8150	82

Berdasarkan Tabel 3.4, diperoleh nilai AIC terkecil adalah 1785.630 yang dimiliki oleh model dengan 3 kelas laten. Model ini memiliki nilai log-likelihood sebesar -810.8150 dengan jumlah parameter yang diestimasi (Npar) sebanyak 82. Sedangkan model dengan 2 kelas laten memiliki nilai AIC yang lebih besar yaitu 1849.265 dengan nilai log-likelihood sebesar -871.6325 dan jumlah parameter yang diestimasi adalah 53. Pemilihan model terbaik didasarkan oleh nilai AIC terkecil, oleh karena itu model terbaik untuk mengelompokkan minat masyarakat terhadap asuransi adalah model dengan 3 kelas laten.

3.6.4 Estimasi Parameter Model Terpilih

Dari uraian sebelumnya, diketahui bahwa model terbaik yang dapat menggambarkan pengelompokan minat masyarakat terhadap asuransi merupakan model dengan 3 kelas laten. Masing-masing kelas memiliki informasi yang berbeda berkaitan dengan variabel indikator dan pilihan jawaban responden. Nilai estimasi parameter setiap variabel didasarkan pada tipe data yang digunakan. Dalam penelitian ini, seluruh variabel indikator yang digunakan

mempunyai tipe data yang sama yaitu ordinal dengan 3 kategori. Oleh karena itu, parameter yang diestimasi adalah $\hat{\alpha}_{jk(r)} = \hat{\gamma}_{jk(r)} - \hat{\gamma}_{jk(r-1)}$ yang merupakan nilai peluang ketika responden memberikan respon ke- r untuk variabel indikator ke- j pada kelas ke- k .

Penentuan karakteristik yang terdapat dalam setiap kelas dilakukan dengan melihat nilai peluang yang dimiliki oleh masing-masing kategori respon. Kategori respon untuk setiap variabel indikator dengan nilai peluang terbesar menjadi karakteristik bagi kelas tersebut. Hasil estimasi untuk model dengan 3 kelas laten dengan bantuan *software* R terdapat dalam Tabel 3.6 berikut.

Tabel 3.6: Estimasi Peluang Setiap Kategori pada Variabel Indikator

Variabel Indikator	Respon	Kelas ke-		
		1	2	3
Y ₁	Setuju	0.6212	0.85	0.3746
	Ragu-Ragu	0.3788	0	0.2781
	Tidak Setuju	0	0.15	0.3473
Y ₂	Setuju	0.8406	0.968	0.7397
	Ragu-Ragu	0.1594	0	0.1496
	Tidak Setuju	0	0.032	0.1107
Y ₃	Benar	0.7182	0.9085	0.3785
	Ragu-Ragu	0.2818	0.0526	0.265
	Salah	0	0.0389	0.3565
Y ₄	Benar	0.8532	0.8662	0.2132
	Ragu-Ragu	0.1286	0.1338	0.6343
	Salah	0.0182	0	0.1525
Y ₅	Benar	0.8894	0.8729	0.1859
	Ragu-Ragu	0.1106	0.1271	0.5972
	Salah	0	0	0.2169
Y ₆	Benar	0.9642	0.9319	0.5494
	Ragu-Ragu	0.0358	0.0681	0.4145
	Salah	0	0	0.0362
Y ₇	Benar	0.2167	0.704	0.0379
	Ragu-Ragu	0.7121	0.2149	0.4719
	Salah	0.0713	0.0811	0.4902
Y ₈	Benar	0.9528	1	0.7119
	Ragu-Ragu	0.0472	0	0.252
	Salah	0	0	0.0362
Y ₁₀	Benar	0.3369	0.6991	0.1451
	Ragu-Ragu	0.6159	0.2671	0.7832
	Salah	0.0471	0.0337	0.0717
Y ₁₁	Setuju	0.048	0.3396	0.0356
	Ragu-Ragu	0.7703	0	0.3558
	Tidak Setuju	0.1816	0.6604	0.6086
Y ₁₂	Setuju	0.0709	0.3396	0
	Ragu-Ragu	0.7818	0	0.4103
	Tidak Setuju	0.1473	0.6604	0.5897
Y ₁₃	Setuju	0.4291	0.7367	0.1065
	Ragu-Ragu	0.4722	0.0939	0.4311
	Tidak Setuju	0.0987	0.1694	0.4624

Dalam Tabel 3.5 terlihat bahwa pada kelas ketiga kelas yang terbentuk, responden meyakini bahwa peristiwa yang tidak pasti dalam hidup dapat menimbulkan kerugian. Hal ini dapat terlihat dari banyaknya jumlah responden yang menjawab pilihan setuju untuk variabel indikator tersebut pada masing-masing kelas, yaitu 62.12% pada kelas pertama, 85% pada kelas kedua, dan 37.46% pada kelas ketiga. Mayoritas responden pada kelas pertama, kedua, dan ketiga juga berkeinginan untuk memiliki dana guna mengantisipasi kerugian yang mungkin akan ditimbulkan tersebut, dengan presentase menjawab setuju untuk masing-masing kelas sebesar 84.06%, 96.8%, dan 73.97%. Dari sebaran jawaban pada kedua pertanyaan tersebut, secara umum, kesadaran responden akan tingkat kebutuhan akan perlindungan sudah baik.

Pada variabel indikator yang mengukur pengetahuan responden terhadap manfaat asuransi sebagai pengalih risiko. Ketiga kelas menunjukkan bahwa sebagian besar responden sudah mengetahui manfaat asuransi sebagai pengalih risiko, dengan presentase sebesar 71.82%, 90.9%, dan 37.85%. Responden pada kelas pertama dan kedua juga memiliki pengetahuan yang baik mengenai istilah penanggung dan tertanggung serta istilah polis dalam asuransi dengan presentase sebesar 85.32% dan 86.6% untuk istilah pihak tertanggung dan penanggung, dan 88.94% dan 87.3% untuk istilah polis. Sedangkan sebanyak 63.43% dan 59.72% responden pada kelompok ketiga masih ragu-ragu mengenai istilah pihak penanggung dan tertanggung dan istilah polis dalam asuransi. Hal yang berbeda ditunjukkan oleh respon terhadap pertanyaan mengenai istilah premi. Mayoritas responden pada ketiga kelas telah mengetahui istilah premi dalam asuransi dengan presentase masing-masing kelas terhadap jawaban setuju adalah 96.42%, 93.2%, dan 54.94%.

Sebanyak 70.4% responden pada kelas kedua sudah memahami prosedur dalam melakukan klaim, namun 71.21% responden pada kelas pertama dan

49.02% responden pada kelas ketiga masih ragu-ragu mengenai hal tersebut. Berbeda dengan pertanyaan yang menyatakan bahwa produk asuransi cukup beragam yang memiliki presentase 95.28% untuk kelas pertama, 100% untuk kelas kedua, dan 71.19% untuk kelas ketiga. Untuk pemahaman responden mengenai pemegang polis yang tidak selalu pihak tertanggung, sebanyak 61.59% responden pada kelas pertama dan 78.32% pada kelas ketiga masih ragu-ragu mengenai hal ini, sedangkan sebanyak 69.9% responden pada kelas kedua sudah memiliki pemahaman yang baik mengenai pertanyaan tersebut.

Kelas pertama berpendapat ragu-ragu bahwa asuransi memiliki lebih menguntungkan dibandingkan menabung/investasi dengan presentase sebesar 77.03%, sedangkan kelas kedua dan kelas ketiga menyatakan bahwa asuransi tidak lebih menguntungkan dibandingkan menabung dengan presentase masing-masing kelas yaitu 66% dan 60.86% . Kelas pertama juga menyatakan ragu-ragu bahwa asuransi memiliki lebih banyak manfaat dibandingkan menabung/investasi dengan presentase sebesar 78.18%, sedangkan kelas kedua dan kelas ketiga menyatakan bahwa asuransi tidak memiliki lebih banyak manfaat dibandingkan menabung/investasi dengan presentase masing-masing sebesar 66% dan 58.97%.

Untuk variabel indikator terakhir yang menyatakan kesediaan responden menyisihkan penghasilannya untuk membayar premi asuransi, sebanyak 47.22% responden pada kelas pertama menyatakan ragu-ragu, 73.7% responden pada kelas kedua menyatakan setuju, sedangkan 46.24% responden pada kelas kedua menyatakan tidak setuju.

3.6.5 Model Regresi Logistik Kelas Laten

Estimasi parameter model regresi logistik kelas laten untuk model dengan 3 kelas ditampilkan dalam nilai yang menjadi intersep dan koefisien X pada

tabel-tabel di bawah. Dalam model dengan 3 kelas laten, model regresi logistik yang terbentuk merupakan model regresi logistik antara kelas kedua dan kelas pertama dan model regresi logistik antara kelas ketiga dan kelas pertama. Adapun hasil untuk masing-masing model regresi logistik yang terbentuk terdapat dalam Tabel 3.7 dan 3.8 :

Tabel 3.7: Estimasi Parameter Model Regresi Logistik Kelas 2 dan 1

Keterangan	Koefisien	Std. error	Nilai t	Probabilitas
Intercep	0.33731	3.00152	0.112	0.912
Jenis Kelamin	-0.83116	0.84913	-0.979	0.341
Usia	0.0457	0.04071	1.122	0.276
Pendidikan	-0.69687	0.82459	-0.845	0.409
Pendapatan	1.01548	0.46159	2.2	0.041

Persamaan fungsi logit yang diperoleh dari tabel di atas dapat ditulis sebagai berikut :

$$\ln \left(\frac{\pi_{2i}}{\pi_{1i}} \right) = 0.33731 - 0.83116 \text{JenisKelamin}_i + 0.0457 \text{Usia}_i - 0.69687 \text{Pendidikan}_i + 1.01548 \text{Pendapatan}_i$$

Dalam tabel di atas, hanya nilai koefisien kovariat pendapatan yang memiliki nilai probabilitas lebih kecil dari $\alpha = 5\%$ yaitu 0.041. Hal ini menunjukkan bahwa pendapatan memiliki pengaruh yang signifikan terhadap pengelompokan minat responden terhadap asuransi di kelas kedua dan pertama. Sehingga fungsi logit yang dihasilkan adalah :

$$\ln \left(\frac{\pi_{2i}}{\pi_{1i}} \right) = 1.01548 \text{Pendapatan}_i \quad (3.7)$$

Sedangkan persamaan fungsi logit yang diperoleh dari Tabel 3.7 adalah sebagai berikut :

$$\ln \left(\frac{\pi_{3i}}{\pi_{1i}} \right) = -1.22907 - 1.1985 \text{Jenis Kelamin}_i + 0.09188 \text{Usia}_i + 0.02693 \text{Pendidikan}_i - 0.02284 \text{Pendapatan}_i$$

Tabel 3.8: Estimasi Parameter Model Regresi Logistik Kelas 3 dan 1

Keterangan	Koefisien	Std. error	Nilai t	Probabilitas
Intercep	-1.22907	3.72126	-0.33	0.745
Jenis Kelamin	-1.1985	1.41054	-0.85	0.407
Usia	0.09188	0.03862	2.379	0.029
Pendidikan	0.02693	0.71294	0.038	0.97
Pendapatan	-0.02284	0.79677	-0.029	0.977

Persamaan fungsi logit antara kelas ketiga dan kelas pertama di atas memiliki nilai koefisien usia yang signifikan dengan nilai probabilitas kurang dari $\alpha = 5\%$ yaitu sebesar 0.029. Sedangkan nilai untuk intercep, koefisien kovariat jenis kelamin, pendidikan, dan pendapatan memiliki nilai probabilitas yang lebih dari α , sehingga tidak signifikan. Hal ini menunjukkan bahwa hanya kovariat usia yang memiliki pengaruh terhadap pengelompokan minat responden terhadap asuransi pada kelas ketiga dan kelas pertama. Sehingga persamaan yang terbentuk menjadi :

$$\ln \left(\frac{\pi_{3i}}{\pi_{1i}} \right) = 0.09188 \text{Usia}_i \quad (3.8)$$

3.6.6 Interpretasi Model

Dari sub bahasan sebelumnya, diketahui bahwa model fungsi logit yang dihasilkan untuk kelas kedua dan kelas pertama adalah :

$$\ln \left(\frac{\pi_{2i}}{\pi_{1i}} \right) = 1.01548 \text{Pendapatan}_i$$

Model tersebut dapat dituliskan menjadi bentuk berikut :

$$\left(\frac{\pi_{2i}}{\pi_{1i}} \right) = e^{1.01548 \text{pendapatan}_i} \quad (3.9)$$

dengan pendapatan_i berarti bahwa nilai pendapatan yang dimiliki oleh responden ke- i . Dalam penelitian ini pendapatan responden dikelompokkan menjadi 3 kategori, kategori 1 untuk pendapatan di bawah UMR, kategori 2 untuk

pendapatan antara UMR sampai dengan Rp 7.000.000 dan kategori 3 untuk pendapatan lebih dari Rp 7.000.000. Misal responden ke-1 memiliki tingkat pendapatan dalam kategori pertama, maka model dalam Persamaan (3.9) dapat diinterpretasikan bahwa kecendrungan seseorang untuk masuk ke kelas kedua dibanding kelas pertama adalah 2.761 lebih besar untuk pendapatan di bawah UMR. Angka ini didapat dengan mensubstitusikan nilai $pendapatan_i$ dengan angka 1 ke dalam Persamaan (3.9).

Sedangkan untuk model fungsi logit untuk kelas ketiga dan kelas pertama dengan model :

$$\ln \left(\frac{\pi_{3i}}{\pi_{1i}} \right) = 0.09188 \text{Usia}_i \quad (3.10)$$

Interpretasi terhadap model tersebut dilakukan dengan mengasumsikan bahwa model tersebut linier. Sehingga penambahan 1 nilai usia, akan menambah nilai $\ln \left(\frac{\pi_{3i}}{\pi_{1i}} \right)$ sebesar 0.09188.

BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan pembahasan pada BAB sebelumnya, dapat ditarik kesimpulan sebagai berikut :

1. Analisis regresi logistik kelas laten dapat digunakan untuk mengelompokkan objek ke dalam kelas-kelas laten yang terbentuk dan secara bersamaan dapat mengetahui latar belakang dari masing-masing kelompok tersebut.
2. Minat masyarakat terhadap Asuransi di wilayah Jakarta Timur dapat dikelompokkan menjadi 3 kelas laten.
3. Secara umum responden pada kelas pertama telah menyadari kebutuhannya dalam hal perlindungan terhadap risiko, hal ini dapat diketahui dari nilai peluang menjawab setuju untuk pertanyaan bahwa peristiwa dalam hidup dapat menimbulkan kerugian dan keinginan untuk memiliki dana untuk mengantisipasi kerugian tersebut. Pengetahuan mereka akan asuransi juga sudah baik, hal ini tercermin dari nilai peluang menjawab setuju untuk pertanyaan mengenai manfaat asuransi, istilah pihak tertanggung dan penanggung, istilah polis, dan premi dalam asuransi. Selain itu pemahaman mereka mengenai asuransi dan ketertarikannya terhadap asuransi juga cukup baik. Hal ini terlihat dari nilai peluang terhadap jawaban atas pertanyaan mengenai pemahaman prosedur me-

lakukan klaim, pemahaman bahwa produk asuransi beragam, pemahaman bahwa pemegang polis tidak selalu pihak tertanggung, persepsi bahwa asuransi lebih menguntungkan dibanding menabung, asuransi memiliki lebih banyak manfaat dibanding asuransi, dan kesediaan menyisihkan sebagian pendapatan untuk membayar premi.

4. Responden pada kelas kedua telah menyadari kebutuhannya dalam hal perlindungan terhadap risiko. Pengetahuan dan pemahaman akan asuransi yang dimiliki pada responden di kelas ini juga sudah baik. Namun, ketertarikan mereka terhadap asuransi kurang baik.
5. Responden pada kelas ketiga masih belum terlalu menyadari kebutuhannya dalam hal perlindungan terhadap risiko. Pengetahuan dan pemahaman akan asuransi yang dimilikinya pun kurang baik. Selain itu, ketertarikan yang dimiliki akan asuransi juga belum baik.
6. Hanya variabel usia sajalah yang mempengaruhi pengelompokan responden pada kelas ketiga dan kelas pertama. Sedangkan variabel pendapatan menjadi variabel yang mempengaruhi pengelompokan responden pada kelas kedua dan kelas pertama.

4.2 Saran

Penelitian ini menggunakan analisis regresi logistik kelas laten dengan 4 variabel kovariat dengan 100 responden, pada penelitian selanjutnya diharapkan dapat digunakan variabel yang lebih banyak dengan jumlah responden yang lebih banyak agar diperoleh hasil yang lebih akurat dalam mengelompokkan minat masyarakat terhadap asuransi.

DAFTAR PUSTAKA

- Agresti A. 2007. *Categorical Data Analysis*. USA: John Wiley & Sons, Inc.
- Chung H, Flaherty B. P, dan Schafer J.L. "Latent Class Logistic Regression: Application to marijuana use and attitudes among high school seniors", *Journal of The Royal Statistical Society: Series A(Statistic in Society)*. 2006.
- Collins, Linda M. dan Stephanie T. Lanza. 2010. *Latent Class and Latent Transition Analysis*. USA: John Wiley & Sons, Inc.
- Ersyad, G.A. 2016. Penggunaan Analisis Regresi Kelas Laten untuk Menentukan Tingkat Kepuasan Pelanggan E-commerce. *Skripsi*. Universitas Negeri Jakarta Fakultas Matematika dan Ilmu Pengetahuan Alam. Jakarta.
- Haeruddin. 2013. Latent Class Regression Analysis Untuk Data Kategorik Dengan Satu Kovariat. *Skripsi*. Universitas Jember Fakultas Matematika dan Ilmu Pengetahuan Alam. Jember.
- Hosmer, David W. dan Lemeshow Stanley. 2000. *Applied Logistic Regression*. Edisi ke-2. USA: John Wiley & Sons, Inc.
- Kurniawan, Henry, "Model Regresi Logistik Kelas Laten Pada Performa Studi Penerima Beasiswa", *Kumpulan Makalah Seminar Semirata 2013: 265-273*. 2013.
- Linzer, D.A. dan Lewis, J.B. "poLCA : An R package for Polytomous Variable Latent Class Analysis", *Journal of Statistical Software Vol.42 No.10: 1-29*. 2011.
- Moustaki, I. dan Papageorgiou, I. "Latent Class Models for Mixed Variables with Applications in Archaeometry", *Computational Statistics & Data Analysis : 1-17*. 2004.

- Bonifasius, Nainggolan MH. Perbandingan Analisis Laten Kelas Dengan Kriteria WHO untuk Penggerombolan Pasien Demam Dengue (DD) dan Demam Berdarah Dengue (DBD). *Tesis*. Institut Pertanian Bogor. Depok.
- Nasution, As'at, "Analisis Minat Masyarakat Kabupaten Mandailing Natal Terhadap Jasa Dan Fasilitas Lembaga Asuransi (Studi Kasus Pegawai Negeri Sipil)", *Jurnal Ekonomi dan Keuangan Vol.3 No.2: 125-140*. 2015.
- Novianti. 2008. Latent Class Model. *Skripsi*. Universitas Indonesia Fakultas Matematika dan Ilmu Pengetahuan Alam. Depok.
- Nurnaningsih, Ratih. 2012. Pengaruh Promosi, Harga, dan Kualitas Layanan Terhadap Keputusan Pembelian Jasa Asuransi Jiwa. *Skripsi*. Universitas Diponegoro Fakultas Ekonomika dan Bisnis. Semarang.
- Riswan, "Pengelompokkan Prestasi Matematika Siswa Indonesia Berdasarkan Hasil Survey TIMSS Menggunakan Analisis Logistik Kelas Laten", *Dinamika Ilmu Vol.13 No.1: 67-87*. 2013.
- Tricahyo, Gustus. 2012. Keefektifan Penggunaan Model Pembelajaran Kooperatif Tipe *Team Assisted Individualization* (TAI) dalam Meningkatkan Minat Belajar Siswa dalam Pembelajaran PKM Kelas XI Mesin di SMK PIRI Sleman. *Skripsi*. Universitas Negeri Yogyakarta.
- Vermunt, J.K. dan Magidson J. 2006. *Technical Guide for Latent GOLD 4.0: Basic and Advanced*. Belmont Massachusetts: Statistical Innovations Inc.
- Vermunt, J. K. dan Magidson J, "Latent Class Cluster Analysis", *Applied Latent Class Analysis*. New York: Cambridge University Press.
- Yamane, Taro. 1967. *Elementary Sampling Theory*. USA: Prentice-Hall, Inc.

LAMPIRAN-LAMPIRAN

Lampiran 1. Kuisioner Penelitian

Karakteristik Responden Tuliskan dan berikan tanda silang (X) pada jawaban yang sesuai dengan data Anda :

1. Jenis Kelamin :

- (a) Laki-Laki
- (b) Perempuan

2. Usia :

3. Pendidikan :

- (a) SD/Sederajat
- (b) SMP/Sederajat
- (c) SMA/Sederajat
- (d) Diploma/S1
- (e) Pasca Sarjana

4. Pendapatan :

- (a) Dibawah UMR
- (b) UMR- Rp 7.000.000
- (c) > Rp 7.000.000

Minat Terhadap Asuransi Berilah tanda centang (V) pada kolom jawaban yang sesuai dengan pendapat Anda:

1. Saya yakin bahwa peristiwa yang tidak pasti dalam hidup dapat menimbulkan kerugian
 - (a) Setuju
 - (b) Ragu-Ragu
 - (c) Tidak Setuju

2. Saya ingin memiliki dana untuk mengantisipasi kerugian tersebut
 - (a) Setuju
 - (b) Ragu-Ragu
 - (c) Tidak Setuju

3. Asuransi merupakan salah satu bentuk jasa dalam pengalihan risiko
 - (a) Benar
 - (b) Ragu-Ragu
 - (c) Salah

4. Pemegang polis merupakan pihak bertanggung sedangkan perusahaan asuransi merupakan pihak penanggung
 - (a) Benar
 - (b) Ragu-Ragu
 - (c) Salah

5. Polis merupakan perjanjian yang mengatur segala hak dan kewajiban pihak bertanggung dan penanggung

(a) Benar

(b) Ragu-Ragu

(c) Salah

6. Premi merupakan iuran yang dibayarkan pihak tertanggung kepada penanggung

(a) Benar

(b) Ragu-Ragu

(c) Salah

7. Prosedur dalam melakukan klaim sederhana

(a) Benar

(b) Ragu-Ragu

(c) Salah

8. Produk asuransi cukup beragam

(a) Benar

(b) Ragu-Ragu

(c) Salah

9. Produk asuransi dapat dibeli di Bank

(a) Benar

(b) Ragu-Ragu

(c) Salah

10. Pemegang polis dalam asuransi tidak selalu pihak tertanggung

- (a) Benar
- (b) Ragu-Ragu
- (c) Salah

11. Menurut saya, asuransi lebih menguntungkan dibandingkan menabung/investasi

- (a) Setuju
- (b) Ragu-Ragu
- (c) Tidak Setuju

12. Menurut saya, asuransi memiliki lebih banyak manfaat dibandingkan menabung/investasi

- (a) Setuju
- (b) Ragu-Ragu
- (c) Tidak Setuju

13. Saya bersedia menyisihkan sebagian penghasilan untuk membayar premi asuransi

- (a) Setuju
- (b) Ragu-Ragu
- (c) Tidak Setuju

X_1	X_2	X_3	X_4	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}
2	18	1	1	1	2	1	2	2	2	2	2	2	2	2	2	2
2	22	3	1	3	3	2	1	3	3	2	3	3	3	1	1	2
2	23	4	2	3	3	3	3	3	3	2	3	3	2	2	1	1
2	22	4	1	3	3	3	2	2	3	2	2	1	1	1	1	2
2	50	2	1	3	3	3	3	3	3	3	3	3	2	1	1	3
1	22	4	2	2	3	2	2	3	2	1	3	2	2	2	2	2
2	21	4	1	3	3	3	2	3	3	3	3	2	3	1	1	2
1	58	3	1	3	3	3	3	3	3	2	3	3	2	2	2	3
1	23	4	1	3	3	1	2	2	2	2	2	2	2	1	1	1
2	28	4	3	3	3	2	3	3	3	2	3	3	2	2	2	3
1	22	3	3	3	3	3	3	3	3	1	3	3	3	1	1	3
1	21	3	3	3	3	3	3	3	3	3	3	1	3	1	1	3
1	21	3	3	3	3	3	3	3	3	2	3	3	3	2	2	3
2	21	3	1	2	3	3	3	2	3	2	3	2	2	2	2	2
2	22	4	1	3	3	3	3	2	3	2	3	2	3	2	2	2
1	23	4	1	3	3	1	1	1	1	1	1	2	1	1	1	1
2	22	3	1	3	3	3	3	3	3	3	3	2	1	2	2	3
2	26	4	1	3	3	3	3	3	3	3	3	3	3	1	1	2
2	21	3	1	3	3	3	3	3	3	3	3	2	3	2	2	3
2	25	4	3	3	3	3	3	3	3	3	3	3	3	1	1	3
2	21	3	1	3	3	3	3	3	3	3	3	3	2	2	2	2
2	43	4	2	3	3	3	3	3	3	3	3	3	3	3	3	3
1	22	3	3	1	3	3	2	2	3	3	3	3	2	3	3	3
1	46	4	2	3	3	3	3	3	3	2	3	1	3	3	3	3
2	22	3	1	1	3	3	3	3	3	2	3	3	2	1	1	3
1	26	4	1	2	2	2	2	2	3	1	3	3	3	3	1	2
2	49	4	2	3	3	3	2	3	2	3	3	2	3	3	3	3
2	59	4	3	3	3	3	2	2	3	2	3	2	2	2	2	2
2	22	4	3	3	3	3	2	2	3	1	3	2	2	1	1	2
2	23	4	1	3	2	2	2	3	3	2	3	3	3	2	2	2
1	21	3	1	1	3	3	3	3	2	3	3	3	3	1	1	1
2	45	4	1	1	1	1	2	2	2	2	3	2	2	1	2	1
1	27	4	3	3	3	3	3	3	3	3	3	3	3	1	1	3
2	22	5	1	2	2	3	3	3	3	2	3	2	3	2	2	2
1	23	4	1	2	3	2	2	2	3	2	3	3	3	3	3	3
2	22	4	1	2	3	2	2	3	2	3	3	2	3	1	2	3
2	20	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
1	21	3	1	1	3	1	3	3	3	2	3	2	2	1	1	1
1	48	4	3	2	3	3	3	3	3	2	3	2	2	2	2	1
1	23	4	1	3	3	1	2	2	2	2	2	2	2	1	1	1
1	21	4	3	3	3	3	3	3	3	2	3	1	3	1	1	1
2	24	4	3	2	2	2	3	3	3	2	3	2	2	1	1	1
1	22	4	1	1	3	1	1	1	3	1	3	3	2	1	1	3
2	21	4	1	2	3	3	3	3	3	2	3	2	2	1	2	2
1	25	3	1	3	3	3	3	3	3	3	3	3	3	1	1	3
1	49	2	2	2	3	2	3	2	2	1	2	2	2	2	1	1
1	22	3	1	3	3	3	3	3	3	3	3	3	3	3	3	3
1	32	4	3	3	3	2	2	2	2	1	3	3	2	2	2	2
1	27	5	3	2	3	3	2	2	2	1	3	3	2	2	2	2
1	56	4	3	2	3	3	2	1	2	2	3	3	2	2	2	2
1	49	4	1	3	3	3	2	2	3	1	3	3	2	1	2	2
1	21	3	3	3	3	2	3	3	3	3	3	3	1	1	1	3

Lampiran 3. Hasil Uji Validitas & Reliabilitas

Correlations															
		Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8	Y9	Y10	Y11	Y12	Y13	Minat
Y1	Pearson Correlation	1	.266	.412	.176	.283	.133	-.216	-.056	-.092	.395	-.254	.214	.254	.529
	Sig. (2-tailed)		.155	.024	.351	.129	.483	.251	.767	.630	.031	.176	.257	.176	.003
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y2	Pearson Correlation	.266	1	.390	.175	.102	.099	.235	-.112	.068	.028	.067	.071	.353	.389
	Sig. (2-tailed)	.155		.033	.354	.590	.602	.212	.555	.720	.883	.724	.710	.056	.034
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y3	Pearson Correlation	.412	.390	1	.539	.551	.384	.486	.372	.038	.372	.278	.293	.390	.756
	Sig. (2-tailed)	.024	.033		.002	.002	.036	.006	.043	.843	.043	.136	.116	.033	.000
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y4	Pearson Correlation	.176	.175	.539	1	.789	.277	.420	.032	-.132	.231	-.043	-.045	-.100	.420
	Sig. (2-tailed)	.351	.354	.002		.000	.138	.021	.867	.485	.219	.822	.812	.598	.021
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y5	Pearson Correlation	.283	.102	.551	.789	1	.207	.645	.310	-.122	.473	-.015	.139	.088	.576
	Sig. (2-tailed)	.129	.590	.002	.000		.273	.000	.096	.520	.008	.939	.465	.644	.001
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y6	Pearson Correlation	.133	.099	.384	.277	.207	1	.070	.484	.132	.116	.151	.060	.322	.412
	Sig. (2-tailed)	.483	.602	.036	.138	.273		.714	.007	.486	.542	.424	.754	.083	.024
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y7	Pearson Correlation	.216	.235	.486	.420	.645	.070	1	.274	.224	.492	.134	.377	.425	.701
	Sig. (2-tailed)	.251	.212	.006	.021	.000	.714		.143	.233	.006	.479	.040	.019	.000
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y8	Pearson Correlation	-.056	-.112	.372	.032	.310	.484	.274	1	.188	.286	.096	.270	.385	.414
	Sig. (2-tailed)	.767	.555	.043	.867	.096	.007	.143		.319	.126	.613	.149	.036	.023
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y9	Pearson Correlation	-.092	.068	.038	-.132	-.122	.132	.224	.188	1	.116	.169	.041	.416	.287
	Sig. (2-tailed)	.630	.720	.843	.485	.520	.486	.233	.319		.542	.372	.829	.022	.124
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y10	Pearson Correlation	.395	.028	.372	.231	.473	.116	.492	.286	.116	1	.337	.270	.465	.640
	Sig. (2-tailed)	.031	.883	.043	.219	.008	.542	.006	.126	.542		.069	.149	.010	.000
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y11	Pearson Correlation	.254	.067	.278	-.043	-.015	.151	.134	.096	.169	.337	1	.799	.394	.584
	Sig. (2-tailed)	.176	.724	.136	.822	.939	.424	.479	.613	.372	.069		.000	.031	.001
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y12	Pearson Correlation	.214	.071	.293	-.045	.139	.060	.377	.270	.041	.270	.799	1	.465	.624
	Sig. (2-tailed)	.257	.710	.116	.812	.465	.754	.040	.149	.829	.149	.000		.010	.000
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Y13	Pearson Correlation	.254	.353	.390	-.100	.088	.322	.425	.385	.416	.465	.394	.465	1	.704
	Sig. (2-tailed)	.176	.056	.033	.598	.644	.083	.019	.036	.022	.010	.031	.010		.000
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30
Minat	Pearson Correlation	.529	.389	.756	.420	.576	.412	.701	.414	.287	.640	.584	.624	.704	1
	Sig. (2-tailed)	.003	.034	.000	.021	.001	.024	.000	.023	.124	.000	.001	.000	.000	.000
	N	30	30	30	30	30	30	30	30	30	30	30	30	30	30

*. Correlation is significant at the 0.05 level (2-tailed).
 **. Correlation is significant at the 0.01 level (2-tailed).

Reliability Statistics

Cronbach's Alpha	N of Items
.810	12

Lampiran 4. Nilai Log-*likelihood* Terbaik untuk Model 2 Kelas Laten

Model 1: llik = -871.6325 ... best llik = -871.6325

Model 2: llik = -886.8787 ... best llik = -871.6325

Model 3: llik = -881.1626 ... best llik = -871.6325

Model 4: llik = -881.1626 ... best llik = -871.6325

Model 5: llik = -881.1626 ... best llik = -871.6325

Model 6: llik = -886.8787 ... best llik = -871.6325

Model 7: llik = -873.2557 ... best llik = -871.6325

Model 8: llik = -873.2557 ... best llik = -871.6325

Model 9: llik = -885.679 ... best llik = -871.6325

Model 10: llik = -873.2557 ... best llik = -871.6325

Model 11: llik = -873.2557 ... best llik = -871.6325

Model 12: llik = -873.2557 ... best llik = -871.6325

Model 13: llik = -873.2557 ... best llik = -871.6325

Model 14: llik = -871.6325 ... best llik = -871.6325

Model 15: llik = -881.1626 ... best llik = -871.6325

Lampiran 5. Hasil untuk Model 2 Kelas Laten dengan *software* R

Conditional item response (column) probabilities,
by outcome variable, for each class (row)

\$Y1

Pr(1) Pr(2) Pr(3)

class 1: 0.0740 0.2347 0.6912

class 2: 0.3354 0.2556 0.4090

\$Y2

Pr(1) Pr(2) Pr(3)

class 1: 0.0264 0.0930 0.8806

class 2: 0.0804 0.1603 0.7593

\$Y3

Pr(1) Pr(2) Pr(3)

class 1: 0.0341 0.1958 0.7701

class 2: 0.3350 0.2520 0.4130

\$Y4

Pr(1) Pr(2) Pr(3)

class 1: 0.0128 0.1188 0.8684

class 2: 0.1603 0.7180 0.1217

\$Y5

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.1264 0.8736

class 2: 0.2377 0.6163 0.1460

\$Y6

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.0523 0.9477

class 2: 0.0396 0.4395 0.5209

\$Y7

Pr(1) Pr(2) Pr(3)

class 1: 0.0682 0.5172 0.4146

class 2: 0.5509 0.4491 0.0000

\$Y8

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.0266 0.9734

class 2: 0.0396 0.2779 0.6825

\$Y10

Pr(1) Pr(2) Pr(3)

class 1: 0.0403 0.4928 0.4669

class 2: 0.0786 0.7592 0.1622

\$Y11

Pr(1) Pr(2) Pr(3)

class 1: 0.3841 0.4551 0.1609

class 2: 0.6057 0.3556 0.0386

\$Y12

Pr(1) Pr(2) Pr(3)

class 1: 0.3585 0.4676 0.1739

class 2: 0.6022 0.3978 0.0000

\$Y13

Pr(1) Pr(2) Pr(3)

class 1: 0.1513 0.3091 0.5395

class 2: 0.4235 0.4710 0.1055

Estimated class population shares

0.7476 0.2524

Predicted class memberships (by modal posterior prob.)

0.75 0.25

=====
Fit for 2 latent classes:
=====

2 / 1

Coefficient	Std. error	t value	Pr(> t)
(Intercept)	-3.78006	3.43118	-1.102 0.276
X1	-1.16790	1.13647	-1.028 0.309
X2	0.07036	0.03776	1.864 0.069
X3	0.84172	0.64864	1.298 0.201
X4	-0.47148	0.69355	-0.680 0.500

=====

number of observations: 100

number of estimated parameters: 53

residual degrees of freedom: 47

maximum log-likelihood: -871.6325

AIC(2): 1849.265

BIC(2): 1987.339

$X^2(2)$: 1154722 (Chi-square goodness of fit)

Lampiran 6. Nilai Log-likelihood Terbaik untuk Model 3 Kelas Laten

Model 1: llik = -810.815 ... best llik = -810.815
Model 2: llik = -823.2568 ... best llik = -810.815
Model 3: llik = -823.2568 ... best llik = -810.815
Model 4: llik = -814.1534 ... best llik = -810.815
Model 5: llik = -819.7193 ... best llik = -810.815
Model 6: llik = -819.9584 ... best llik = -810.815
Model 7: llik = -810.9465 ... best llik = -810.815
Model 8: llik = -815.5929 ... best llik = -810.815
Model 9: llik = -812.0557 ... best llik = -810.815
Model 10: llik = -836.7446 ... best llik = -810.815
Model 11: llik = -855.3251 ... best llik = -810.815
Model 12: llik = -813.2113 ... best llik = -810.815
Model 13: llik = -847.0638 ... best llik = -810.815
Model 14: llik = -825.8822 ... best llik = -810.815
Model 15: llik = -825.5693 ... best llik = -810.815

Lampiran 7. Hasil untuk Model 3 Kelas Laten dengan *software* R

Conditional item response (column) probabilities,
by outcome variable, for each class (row)

\$Y1

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.3788 0.6212

class 2: 0.1500 0.0000 0.8500

class 3: 0.3473 0.2781 0.3746

\$Y2

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.1594 0.8406

class 2: 0.0320 0.0000 0.9680

class 3: 0.1107 0.1496 0.7397

\$Y3

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.2818 0.7182

class 2: 0.0389 0.0526 0.9085

class 3: 0.3565 0.2650 0.3785

\$Y4

Pr(1) Pr(2) Pr(3)

class 1: 0.0182 0.1286 0.8532

class 2: 0.0000 0.1338 0.8662

class 3: 0.1525 0.6343 0.2132

\$Y5

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.1106 0.8894

class 2: 0.0000 0.1271 0.8729

class 3: 0.2169 0.5972 0.1859

\$Y6

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.0358 0.9642

class 2: 0.0000 0.0681 0.9319

class 3: 0.0362 0.4145 0.5494

\$Y7

Pr(1) Pr(2) Pr(3)

class 1: 0.0713 0.7121 0.2167

class 2: 0.0811 0.2149 0.7040

class 3: 0.4902 0.4719 0.0379

\$Y8

Pr(1) Pr(2) Pr(3)

class 1: 0.0000 0.0472 0.9528

class 2: 0.0000 0.0000 1.0000

class 3: 0.0362 0.2520 0.7119

\$Y10

Pr(1) Pr(2) Pr(3)

class 1: 0.0471 0.6159 0.3369

class 2: 0.0337 0.2671 0.6991

class 3: 0.0717 0.7832 0.1451

\$Y11

Pr(1) Pr(2) Pr(3)

class 1: 0.1816 0.7703 0.0480

class 2: 0.6604 0.0000 0.3396

class 3: 0.6086 0.3558 0.0356

\$Y12

```

Pr(1) Pr(2) Pr(3)
class 1: 0.1473 0.7818 0.0709
class 2: 0.6604 0.0000 0.3396
class 3: 0.5897 0.4103 0.0000

```

\$Y13

```

Pr(1) Pr(2) Pr(3)
class 1: 0.0987 0.4722 0.4291
class 2: 0.1694 0.0939 0.7367
class 3: 0.4624 0.4311 0.1065

```

Estimated class population shares

```
0.4305 0.293 0.2766
```

Predicted class memberships (by modal posterior prob.)

```
0.42 0.3 0.28
```

```
=====
Fit for 3 latent classes:
=====
```

2 / 1

Coefficient	Std. error	t value	Pr(> t)	
(Intercept)	0.33731	3.00152	0.112	0.912
X1	-0.83116	0.84913	-0.979	0.341
X2	0.04570	0.04071	1.122	0.276
X3	-0.69687	0.82459	-0.845	0.409
X4	1.01548	0.46159	2.200	0.041

```
=====
```

3 / 1

Coefficient	Std. error	t value	Pr(> t)	
(Intercept)	-1.22907	3.72126	-0.330	0.745
X1	-1.19850	1.41054	-0.850	0.407
X2	0.09188	0.03862	2.379	0.029
X3	0.02693	0.71294	0.038	0.970
X4	-0.02284	0.79677	-0.029	0.977

=====

number of observations: 100

number of estimated parameters: 82

residual degrees of freedom: 18

maximum log-likelihood: -810.815

AIC(3): 1785.63

BIC(3): 1999.254

 $X^2(3)$: 1727273 (Chi-square goodness of fit)

SURAT PERNYATAAN KEASLIAN SKRIPSI

Dengan ini saya yang bertanda tangan di bawah ini, mahasiswa Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta:

Nama : Irena Purwa Ningrida
No. Registrasi : 3125136343
Program Studi : Matematika

Menyatakan bahwa skripsi ini yang saya buat dengan judul "**Analisis Regresi Logistik Kelas Laten Minat Masyarakat terhadap Asuransi**" adalah :

1. Dibuat dan diselesaikan oleh saya sendiri.
2. Bukan merupakan duplikat skripsi yang pernah dibuat oleh orang lain atau jiplakan karya tulis orang lain.

Pernyataan ini dibuat dengan sesungguhnya dan saya bersedia menanggung segala akibat yang timbul jika pernyataan saya tidak benar.

Jakarta, Agustus 2017

Yang membuat pernyataan



Irena Purwa Ningrida

DAFTAR RIWAYAT HIDUP



IRENA PURWA NINGRIDA. Lahir di Tangerang, 24 Mei 1995. Anak pertama dari pasangan Bapak Her-yansyah dan Ibu Lilis Yuningsih. Saat ini bertempat tinggal di Jalan Dahlia IX Blok F27 No.11, Pondok Indah, Kutabumi, Tangerang 15561.

No. Ponsel : 089658769210

Email : irenapurwaningrida@gmail.com

Riwayat Pendidikan : Penulis mengawali pendidikan di TK Islam Al-Kamal selama 2 tahun, dan kemudian melanjutkan pendidikan di SD Negeri Kutabumi 3 pada tahun 2001 - 2007. Setelah itu, penulis melanjutkan ke SMP Negeri 2 Tangerang hingga tahun 2010. Kemudian kembali melanjutkan ke SMA Negeri 1 Tangerang dan lulus tahun 2013. Di Tahun yang sama penulis melanjutkan ke Universitas Negeri Jakarta (UNJ), program studi Matematika, melalui jalur UMB. Di pertengahan tahun 2017 penulis telah memperoleh gelar Sarjana Sains untuk Program Studi Matematika, FMIPA, UNJ.

Riwayat Organisasi : Selama di bangku perkuliahan, penulis aktif di berbagai organisasi kemahasiswaan. Dalam tahun pertama, penulis mendapat kepercayaan sebagai staff Departemen Informasi dan Komunikasi BEMJ Matematika, khususnya dalam pengelolaan buku tahunan dan sosial media. Pada tahun selanjutnya, penulis mendapat kepercayaan sebagai staff ahli Departemen Kaderisasi BEMJ Matematika.

Riwayat Pekerjaan : Penulis mulai menjadi pengajar matematika sejak tahun 2017.