



LATENT CLASS REGRESSION ANALYSIS
UNTUK DATA KATEGORIK DENGAN SATU KOVARIAT

SKRIPSI

Oleh

Haeruddin

NIM 081810101052

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2013



LATENT CLASS REGRESSION ANALYSIS
UNTUK DATA KATEGORIK DENGAN SATU KOVARIAT

SKRIPSI

diajukan guna melengkapi tugas akhir dan memenuhi salah satu syarat
untuk menyelesaikan Program Studi Matematika (S1)
dan mencapai gelar Sarjana Sains

Oleh

Haeruddin
NIM 081810101052

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
2013

PERSEMBAHAN

Skripsi ini saya persembahkan untuk:

1. Ibunda Tomyana, Ayahanda Samin dan nenekku Misnaya yang memberikan kasih sayang, doa dan restu dalam perjalanan hidupku;
2. kakak-kakakku Yusup Efendi dan Imam Hapid yang selalu memberikan motivasi tanpa henti;
3. guru-guruku sejak sekolah dasar sampai dengan perguruan tinggi yang telah mengajarku tentang pentingnya ilmu dalam kehidupan ini;
4. Almamater Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

MOTTO

Rasulullah SAW. Bersabda, “janganlah kalian menuntut ilmu untuk memanggakannya terhadap para ulama’ dan untuk diperdebatkan di kalangan orang-orang bodoh dan buruk perangainya. Jangan pula menuntut ilmu untuk penampilan dalam majelis (pertemuan atau rapat) dan untuk menarik perhatian orang-orang kepadamu. Barangsiapa seperti itu maka baginya neraka...neraka”.

(HR. Attirmidzi dan Ibnu Majah)¹⁾

Manusia, seribu orang di antara mereka seperti satu orang
Seseorang seperti seribu orang apabila memiliki keahlian.

(‘Aidh Al-Qarni)²⁾

¹⁾Gus AA. 2009. *Matematika AlQur’an (Mengungkap M’jizat dengan Bahasa Angka)*. Surakarta: Rahma Media Pustaka

²⁾Alqarni, ‘Aidh. 2010. *Detik-detik Usia yang Paling Mahal*. Yogyakarta: Bening

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

nama : Haeruddin

NIM : 081810101052

menyatakan dengan sesungguhnya bahwa karya ilmiah yang berjudul “Latent Class Regression Analysis Untuk Data Kategorik Dengan Satu Kovariat” adalah benar-benar hasil karya sendiri, kecuali kutipan yang sudah saya sebutkan sumbernya, belum pernah diajukan pada institusi manapun, dan bukan karya jiplakan. Saya bertanggung jawab atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmiah yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa ada tekanan dan paksaan dari pihak manapun serta bersedia mendapat sanksi akademik jika ternyata di kemudian hari pernyataan ini tidak benar.

Jember, 29 Januari 2013

Yang menyatakan,

Haeruddin

NIM. 081810101052

SKRIPSI

LATENT CLASS REGRESSION ANALYSIS
UNTUK DATA KATEGORIK DENGAN SATU KOVARIAT

Oleh
Haeruddin
NIM 081810101052

Pembimbing:

Dosen Pembimbing Utama : Prof. Drs. I Made Tirta, M.Sc, Ph.D.
Dosen Pembimbing Anggota : Yuliani Setia Dewi, S.Si, M.Si.

PENGESAHAN

Skripsi berjudul “Latent Class Regression Analysis Untuk Data Kategorik Dengan Satu Kovariat” telah diuji dan disahkan pada:

hari, tanggal :

tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas
Jember.

Tim Penguji:

Ketua,

Sekretaris

Prof. Drs. I Made Tirta, M.Sc., Ph.D.
NIP 195912201985031002

Yuliani Setia Dewi, S.Si, M.Si.
NIP 197407162000032001

Penguji I,

Penguji II,

Dr. Alfian Futuhul Hadi, S.Si, M.Si.
NIP 197407192000121001

Kiswara Agung Santoso, M.Kom.
NIP 197209071998031003

Mengesahkan

Dekan,

Prof. Drs. Kusno, DEA., Ph.D.
NIP 196101081986021001

RINGKASAN

Latent Class Regression Analysis Untuk Data Kategorik Dengan Satu Kovariat; Haeruddin, 081810101052; 2013; 42 halaman; Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Variabel laten merupakan variabel yang tidak dapat diukur secara langsung, melainkan harus melalui beberapa indikator. Salah satu metode statistik yang bisa digunakan dalam analisis variabel laten yaitu *latent class analysis*. Analisis ini merupakan metode untuk mengelompokkan objek penelitian ke dalam kelompok-kelompok laten, dimana variabel laten dan variabel indikator bersifat kategorik. Apabila pengelompokan objek tersebut dipengaruhi oleh variabel pengiring (kovariat), maka analisis yang digunakan yaitu *latent class regression analysis*. Metode yang digunakan untuk menduga parameter-parameter pada *latent class regression analysis*, yaitu Algoritma EM (ekspektasi-maksimisasi) kemudian dilanjutkan dengan metode Newton-Raphson. Pada algoritma EM, teorema Bayes diperlukan dalam menduga klasifikasi objek pada tahap ekspektasi, sedangkan metode maksimum likelihood digunakan pada tahap maksimisasi.

Pada penelitian ini, data kategorik yang digunakan yaitu data penentuan strategi pemasaran berdasarkan peluang (*opportunity*) dan ancaman (*treath*) pada PT. Panahmas Dwitama Distrindo Regional Jember. Tujuan penelitian ini adalah untuk mengetahui pengelompokan responden berdasarkan persepsinya terhadap *opportunity* dan *treath* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Kovariat yang digunakan yaitu lamanya berlangganan responden terhadap distributor produk tersebut. Analisis menggunakan *software* R dengan paket *poLCA*.

Pada *opportunity*, responden dikelompokkan menjadi tiga kelompok. kelompok pertama terdiri dari responden yang setuju bahwasannya minat masyarakat, modal, promosi dan produk baru merupakan *opportunity* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember,

dan tidak setuju apabila kualitas kerja menjadi *opportunity* distributor tersebut. Kelompok kedua terdiri dari responden yang setuju bahwasannya minat masyarakat, modal, promosi dan produk baru merupakan *opportunity* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember, dan menyatakan cukup apabila kualitas kerja menjadi *opportunity* distributor tersebut. Kelompok ketiga terdiri dari responden yang menyatakan setuju apabila minat masyarakat, kualitas kerja dan produk baru merupakan *opportunity* distributor tersebut, dan sangat setuju apabila modal dan promosi merupakan *opportunity* yang dimiliki oleh distributor tersebut. Berdasarkan lamanya berlangganan, jumlah responden di kelompok kedua lebih banyak dari pada kelompok lain untuk lamanya berlangganan antara 1 sampai 15 tahun, sedangkan kelompok pertama lebih banyak daripada kelompok lain untuk lamanya berlangganan antara 26 sampai 30 tahun.

Pada *treath*, responden dikelompokkan menjadi dua kelompok. Kelompok pertama terdiri dari responden yang menyatakan setuju apabila pesaing memberikan harga lebih murah, semakin besarnya biaya operasional perusahaan dan kondisi perekonomian yang tidak stabil karena kenaikan BBM merupakan *treath* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Sedangkan kelompok kedua menyatakan cukup apabila tiga variabel indikator ini merupakan *treath* yang dimiliki oleh distributor tersebut. Pada masing-masing lamanya berlangganan, responden di kelompok pertama lebih banyak dari pada di kelompok kedua. Kelompok yang memiliki responden paling banyak yaitu kelompok pertama dengan lama berlangganan antara 6 sampai 10 tahun, sebanyak 24 orang.

PRAKATA

Alhamdulillah, puji syukur ke hadirat Allah SWT, atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul “Latent Class Regression Analysis untuk Data Kategorik dengan Satu Kovariat”. Skripsi ini disusun untuk memenuhi salah satu syarat menyelesaikan pendidikan strata satu (S1) pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Prof. Drs. I Made Tirta, M.Sc., Ph.D., selaku Dosen Pembimbing Utama dan Yuliani Setia Dewi, S.Si, M.Si., selaku Dosen Pembimbing Anggota yang telah meluangkan waktu, pikiran, dan perhatian dalam penulisan skripsi ini;
2. Dr. Alfian Futuhul Hadi, S.Si, M.Si., dan Kiswara Agung Santoso, M.Kom., selaku dosen penguji atas saran-saran yang diberikan;
3. seluruh staf pengajar Jurusan Matematika Fakultas MIPA Universitas Jember yang telah memberikan ilmu serta bimbingannya sehingga penulis dapat menyelesaikan skripsi ini;
4. sahabat Imam Safii, Ika Agus Rini, Ahmad Turidi, Alfa Rijal Alfian, Zainal Abidin yang selalu memberikan semangat dalam penyelesaian tugas akhir ini;
5. Permata, Ana, Ika Nurul serta teman-teman angkatan 2008 Jurusan Matematika lainnya yang tidak bisa disebutkan satu persatu terima kasih atas motivasinya;
6. teman-teman semua angkatan di Jurusan Matematika dan semua pihak yang tidak dapat disebutkan satu per satu.

Penulis juga menerima segala kritik dan saran dari semua pihak demi kesempurnaan skripsi ini. Akhirnya penulis berharap, semoga skripsi ini dapat bermanfaat.

Jember, 29 Januari 2013

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PERSEMBAHAN	ii
HALAMAN MOTTO	iii
HALAMAN PERNYATAAN.....	iv
HALAMAN BIMBINGAN.....	v
HALAMAN PENGESAHAN.....	vi
RINGKASAN	vii
PRAKATA	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
BAB 1. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Permasalahan	3
1.3 Tujuan	3
1.4 Manfaat	4
BAB 2. TINJAUAN PUSTAKA.....	5
2.1 Analisis Multivariat	5
2.2 Metode Estimasi Parameter dengan Maksimum Likelihood	6
2.3 Teorema Bayes	7
2.4 Algoritma EM (Ekspektasi Maksimisasi)	8
2.5 Metode Newton-Raphson	9
2.6 <i>Latent Class Analysis</i>	10
2.6.1 Pengertian <i>Latent Class Model</i>.....	10

2.6.2	Estimasi Parameter <i>Latent Class Model</i>	11
2.7	<i>Latent Class Regression Analysis</i>	13
2.7.1	Pengertian <i>Latent Class Regression Model</i>	14
2.7.2	Estimasi Parameter <i>Latent Class Regression Model</i>	14
2.8	Uji Kecocokan Model	17
2.8.1	AIC (<i>Akaike's Information Criterion</i>)	17
2.8.2	BIC (<i>Bayesian Information Criterion</i>).....	18
BAB 3.	METODE PENELITIAN	19
3.1	Data	19
3.1.1	Sumber data	19
3.1.2	Identifikasi Variabel	19
3.2	Metode Analisis Data	22
3.2.1	Paket poLCA (<i>Polytomous Variable Latent Class Analysis</i>) .	22
3.2.2	Struktur Fungsi poLCA pada R.....	22
3.2.3	Penyajian grafik Keanggotaan Kelas Laten	23
3.2.4	Langkah-langkah Penelitian	23
BAB 4.	HASIL DAN PEMBAHASAN	25
4.1	Estimasi Parameter dalam <i>Latent Class Regression Analysis</i>	25
4.2	Uji Kecocokan Model	26
4.3	Estimasi Peluang Kelas Laten Berdasarkan Kompok dan Variabel Indikator	27
4.3.1	Peluang Kelas Laten Berdasarkan Kelompok dan Variabel Indikator untuk <i>Opportunity</i>	28
4.3.2	Peluang Kelas Laten Berdasarkan Kelompok dan Variabel Indikator untuk <i>Treath</i>	29
4.4	Model dalam <i>Latent Class Regression Analysis</i>	30
4.4.1	Regresi Kelas Laten untuk <i>Opportunity</i>	30
4.4.2	Regresi Kelas Laten untuk <i>Treath</i>	31
4.5	Peluang Kelas Laten Berdasarkan Kovariat	32

4.5.1 Peluang Kelas Laten Berdasarkan Kovariat untuk <i>Opportunity</i>	32
4.5.2 Peluang Kelas Laten Berdasarkan Kovariat untuk <i>Treath</i>	32
4.6 Plot Model Regresi Kelas Laten.....	33
4.6.1 Plot Model Regresi Kelas Laten untuk <i>Opportunity</i>	33
4.6.2 Plot Model Regresi Kelas Laten untuk <i>Treath</i>	34
4.7 Pembahasan	35
BAB 5. PENUTUP.....	41
5.1 Kesimpulan	41
5.2 Saran	41
DAFTAR PUSTAKA	43
LAMPIRAN	44

DAFTAR GAMBAR

	Halaman
2.1. Skema Algoritma EM dalam <i>Latent Class Regression Analysis</i>	16
3.1. Langkah-langkah Penelitian	23
4.1. Plot Grafik Regresi Kelas Laten pada Model Tiga Kelompok untuk <i>Opportunity</i>	34
4.2. Plot Grafik Regresi Kelas Laten pada Model Dua Kelompok untuk <i>Treath</i> (Persamaan 4.5)	34
4.3. Plot Grafik Regresi Kelas Laten pada Model Dua Kelompok untuk <i>Treath</i> (Persamaan 4.6)	35

DAFTAR TABEL

	Halaman
3.1. Definisi Operasional dari Peluang (<i>Opportunity</i>)	20
3.2. Definisi Operasional dari Ancaman (<i>Treath</i>)	21
4.1. Peluang Awal Kelas Laten untuk Masing-masing Model pada <i>Opportunity</i>	25
4.2. Peluang Awal Kelas Laten Untuk Masing-masing Model Pada <i>Treath</i>	26
4.3. Uji Kecocokan Model Pada <i>Opportunity</i>	26
4.4. Uji Kecocokan Model Pada <i>Treath</i>	27
4.5. Hasil Estimasi π_{jrk} untuk Model Tiga Kelompok pada <i>Opportunity</i>	28
4.6. Hasil Estimasi π_{jrk} untuk Model Dua Kelompok pada <i>Treath</i>	29
4.7. Regresi Kelas Laten pada Model Tiga Kelompok untuk <i>Opportunity</i>	30
4.8. Regresi Kelas Laten pada Model Dua Kelompok untuk <i>Treath</i>	31
4.9. Estimasi p_{ri} pada Model Tiga Kelompok untuk <i>Opportunity</i>	32
4.10. Estimasi p_{ri} pada Model Dua Kelompok untuk <i>Treath</i>	33
4.11. Pengelompokan Responden pada Masing-masing Kelompok untuk <i>Opportunity</i>	36
4.12. Pengelompokan Responden Berdasarkan Kovariat untuk <i>Opportunity</i>	37
4.13. Pengelompokan Responden pada Masing-masing Kelompok untuk <i>Treath</i>	38
4.14. Pengelompokan Responden Berdasarkan Kovariat untuk <i>Treath</i>	39
4.15. Pengelompokan Responden Berdasarkan Kovariat untuk <i>Treath</i> dan <i>Opportunity</i>	39

DAFTAR LAMPIRAN

	Halaman
A. Skrip poLCA dalam <i>software R</i>	45
B. Kuesioner untuk Eksternal Pelanggan (Outlet)	46
C. Skrip Program dan Output untuk <i>Opportunity</i> Menggunakan <i>Software R</i> ...	47
C. Skrip Program dan Output untuk <i>Treath</i> Menggunakan <i>Software R</i>	51

BAB 1. PENDAHULUAN

1.1 Latar Belakang

Dalam berbagai penelitian, seperti di bidang psikologi dan pendidikan, kadang nilai variabel yang diteliti tidak dapat diukur secara langsung, melainkan harus melalui beberapa indikator. Variabel yang tidak dapat diukur ini disebut variabel laten. Variabel laten dapat bersifat kontinu atau kategorik, dan dapat dibentuk oleh indikator yang bersifat kontinu atau kategorik pula (Novianti, 2008). Variabel indikator ini akan digunakan dalam metode statistik untuk menganalisis variabel laten tersebut. Salah satu metode statistik yang digunakan untuk menganalisis variabel laten yaitu analisis kelas laten (*latent class analysis*). Analisis ini digunakan untuk mengelompokkan objek-objek penelitian ke dalam kelompok-kelompok berdasarkan kesamaan karakteristik dari objek-objek tersebut.

Latent class analysis telah banyak digunakan dalam penelitian yang berhubungan dengan pengelompokan kelas laten. Pada tahun 2008, Novianti melakukan penelitian untuk mengelompokkan pasien demam berdarah RS. Dr. Oen 2 Solo berdasarkan gejalanya. Pengelompokan ini diukur menggunakan beberapa variabel indikator dengan respon biner. Dalam penelitian tersebut, Novianti (2008) menyimpulkan bahwasannya pasien dikelompokkan dalam tiga kelompok yaitu kelompok pertama untuk pasien dengan denyut nadi lemah, kelompok kedua untuk pasien yg membutuhkan transfusi, denyut nadi lemah, tekanan darah rendah dan trombositnya rendah, sedangkan kelompok ketiga untuk pasien dengan denyut nadi lemah dan trombositnya rendah.

Nainggolan (2009) meneliti tentang pengelompokan pasien Demam Dangué (DB) dan Demam Berdarah Dangué (DBD) berdasarkan kriteria WHO menggunakan *latent class analysis*. Jika dalam penelitian Novianti (2008) menggunakan variabel

indikator kategorik dengan respon biner, dalam penelitian yang dilakukan Nainggolan (2009) ini menggunakan variabel indikator kontinu dan variabel kategorik dengan respon biner. Variabel indikator yang digunakan yaitu hasil diagnosis dari pasien demam berdarah dangue yang berada di Rumah Sakit Cipto Mangunkusumo (RSCM) Jakarta.

Ada beberapa jenis *latent class analysis* antara lain *latent class cluster analysis* dan *latent class regression analysis*. *Latent class cluster analysis* merupakan analisis dengan pendekatan berbasis model yang memberikan kemungkinan untuk menguji kesesuaian antara model secara statistik. Tujuan dari analisis ini yaitu untuk mengidentifikasi jumlah kelas yang dibutuhkan dalam rangka menjelaskan hubungan variabel-variabel yang diamati (Vermut & Magidson, 2003). Salah satu penelitian yang menggunakan analisis ini dilakukan oleh Hanifah (2010) yaitu *latent class cluster analysis* untuk variabel indikator bertipe campuran dalam rangka pengelompokan desa. Dalam penelitian itu, faktor yang diduga menjadi penyebab kemajuan atau ketertinggalan suatu desa, yaitu faktor alam/lingkungan, faktor kelembagaan, faktor sarana/prasarana dan akses, serta faktor sosial ekonomi penduduk.

Vermunt (2004) menyatakan bahwa *latent class regression analysis* digunakan untuk mengetahui pengaruh variabel pengiring atau kovariat dalam pembentukan kelas-kelas laten. Berbeda dengan *latent class cluster analysis* yang menggunakan beberapa faktor, *latent class regression analysis* menggunakan kovariat yang berulang kali diamati. Adanya kovariat ini pula yang membedakan antara *latent class analysis* standar dengan *latent class regression analysis*.

Pada tahun 2008, Santoso meneliti tentang penentuan strategi pemasaran distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Penelitian tersebut dilakukan dengan menggunakan *EFAS (Eksternal Factor Analisis Summary)* dan Matrik *SWOT*. Data yang digunakan yaitu data kategorik tentang penilaian responden terhadap *opportunity* dan *treath* yang dimiliki oleh distributor tersebut. Analisis ini hanya untuk mengetahui strategi yang bisa digunakan oleh

distributor untuk perkembangan usaha. Akan tetapi, data ini tidak digunakan dalam rangka pengelompokan responden berdasarkan penilaian tersebut.

Berdasarkan uraian di atas, dapat diketahui bahwasannya *latent class regression analysis* merupakan analisis yang dibutuhkan untuk data yang bersifat kualitatif. Sehingga penelitian yang berkaitan dengan analisis data kualitatif ini dipandang perlu untuk dilakukan. Dalam penelitian ini, *latent class regression analysis* diterapkan untuk data kategorik dengan satu kovariat. Data kategorik yang digunakan dalam penelitian ini yaitu data penentuan strategi pemasaran berdasarkan peluang (*opportunity*) dan ancaman (*treath*) pada PT. Panahmas Dwitama Distrindo Regional Jember.

1.2 Permasalahan

Permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut:

1. bagaimana analisis pengelompokan responden berdasarkan persepsinya terhadap *opportunity* dan *treath* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember dengan menggunakan *latent class regression analysis*?
2. bagaimana analisis pengelompokan responden berdasarkan lamanya berlangganan terhadap distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember?

1.3 Tujuan

Adapun tujuan dari penelitian ini adalah untuk mengetahui pengelompokan responden berdasarkan persepsinya terhadap *opportunity* dan *treath* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember dengan menggunakan *latent class regression analysis*, serta untuk mengetahui pengelompokan responden berdasarkan lamanya berlangganan terhadap distributor tersebut.

1.4 Manfaat

Adapun manfaat yang dapat diambil dari penelitian ini adalah dapat menambah wawasan dan pengetahuan tentang analisis statistika khususnya *latent class regression analysis* dalam rangka klasifikasi persepsi pelanggan terhadap distributor suatu produk. Selain itu, hasil penelitian ini juga bisa dijadikan referensi sehingga mempermudah dalam menentukan strategi pemasaran distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember

BAB 2. TINJAUAN PUSTAKA

Latent class regression analysis merupakan bagian dari analisis multivariat untuk data bersifat kategorik dengan variabel yang tidak diketahui secara langsung. Metode yang digunakan untuk menduga parameter-parameter pada *latent class regression analysis* yaitu Algoritma EM (ekspektasi-maksimisasi) kemudian dilanjutkan dengan metode Newton-Raphson. Pada algoritma EM, teorema Bayes diperlukan dalam menduga klasifikasi objek pada tahap ekspektasi, sedangkan metode maksimum likelihood digunakan pada tahap maksimisasi untuk menentukan estimasi parameter yang paling baik.

2.1 Analisis Multivariat

Dalam metode statistik, salah satu teknik yang digunakan untuk menganalisis variabel adalah analisis multivariat. Analisis multivariat dibagi menjadi dua yaitu metode dependensi dan metode interdependensi. Metode dependensi adalah metode yang digunakan untuk mengelompokkan variabel menjadi variabel bebas yang mempengaruhi dan variabel tak bebas yang dipengaruhi. Sedangkan metode interdependensi adalah metode yang digunakan untuk mengelompokkan berdasarkan variabel dependennya saja. Metode interdependensi dilakukan untuk pengelompokkan atau mereduksi variabel yang banyak sekali menjadi variabel baru yang lebih sedikit, tetapi tidak mengurangi informasi yang terkandung dalam variabel asli. Analisis multivariat yang termasuk metode interdependensi antara lain analisis faktor, analisis multidimensi dan analisis kluster (Priatna, tanpa tahun).

Analisis kluster merupakan teknik analisis data yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan karakteristik yang dimilikinya,

sehingga objek yang terletak dalam satu kluster akan mempunyai sifat relatif homogen. Objek yang berada di dalam kluster harus memiliki kemiripan, sedangkan objek yang tidak berada dalam satu kluster tidak mempunyai kemiripan. Kluster-kluster yang terbentuk memiliki homogenitas internal yang tinggi dan heterogenitas eksternal yang tinggi.

Menurut Kaufman dan Rousseeuw (1990); Everitt (1993) (dalam Zhang, 2004), analisis kluster merupakan pembagian objek yang mempunyai sifat sama ke dalam kelompok-kelompok yang bermakna. Ketika objek yang diamati bersifat kontinu, analisis kluster disebut *latent profile analysis*. Ketika objek yang diamati bersifat kategorik, analisis kluster disebut *latent class analysis*. Selain itu, ada juga analisis kluster yang digunakan ketika objek yang diamati bersifat campuran (kontinu dan kategorik).

Apabila dalam *latent class analysis* tersebut dimasukkan suatu variabel pengiring (kovariat), maka analisis tersebut dinamakan *latent class regression analysis*. Dalam analisis ini, pengelompokan objek akan dilihat berdasarkan nilai kovariat tersebut.

2.2 Metode Estimasi Parameter dengan Maksimum Likelihood

Metode maksimum likelihood merupakan salah satu cara untuk mengestimasi parameter yang tidak diketahui. Prosedur estimasi ini menguji apakah estimasi maksimum yang tidak diketahui dari fungsi likelihood suatu sampel nilainya sudah memaksimumkan fungsi likelihood.

Misalkan x_1, x_2, \dots, x_n adalah variabel acak yang saling independen dan mempunyai fungsi kepadatan peluang $f(x, \theta)$ dimana θ merupakan parameter, maka fungsi likelihood dapat didefinisikan sebagai berikut:

$$L(x_1, x_2, \dots, x_n; \theta) = f(x_1, \theta)f(x_2, \theta) \dots f(x_n, \theta) \quad (2.1)$$

Bila fungsi likelihood terdifferensialkan pada θ , maka estimasi maksimum likelihood dapat diperoleh melalui persamaan berikut:

$$(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_n) \rightarrow \frac{\partial L(x_1, x_2, \dots, x_n; \theta)}{\partial \hat{\theta}_i} = 0$$

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

Dalam banyak kasus, penggunaan differensiasi akan lebih mudah digunakan pada logaritma natural dari $L(x_1, x_2, \dots, x_n; \theta)$, yaitu:

$$K(x_1, x_2, \dots, x_n; \theta) = \ln L(x_1, x_2, \dots, x_n; \theta) \quad (2.2)$$

Langkah-langkah untuk menentukan estimasi maksimum likelihood dari θ_i adalah:

1. Menentukan fungsi likelihood seperti pada persamaan (2.1)
2. Membentuk logaritma natural likelihood seperti pada persamaan (2.2)
3. Membentuk persamaan likelihood dan menyelesaikan:

$$\frac{\partial \ln L(x_1, x_2, \dots, x_n; \theta)}{\partial \hat{\theta}_i} = 0$$

4. Mendapatkan estimasi maksimum likelihood dari θ_i , yaitu $\hat{\theta}_i$

(Chandra, 2009)

2.3 Teorema Bayes

Misal A dan B adalah dua kejadian dalam ruang sampel S dan $P(B) \neq 0$, maka peluang terjadinya A apabila telah terjadi B dapat dirumuskan seperti pada persamaan (2.3).

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

Suatu himpunan B_i , dengan $i = 1, 2, \dots, m$, B_i dikatakan partisi dari ruang sampel S jika:

- a. $B_j \cap B_k = \emptyset$ untuk $j \neq k$ dimana $j = 1, 2, \dots, m$ dan $k = 1, 2, \dots, m$
- b. $\bigcup_{i=1}^m B_i = S$
- c. $P(B_i) > 0$ untuk $\forall i$

Bila A merupakan peristiwa bagian dari S sehingga $A = \bigcup_{i=1}^m (A \cap B_i)$ dimana masing-masing $(A \cap B_i)$ adalah saling lepas secara berpasangan, maka $P(A) =$

$P(\cup_{i=1}^m (A \cap B_i)) = \sum_{i=1}^m P(A \cap B_i)$ dan dengan menggunakan peluang bersyarat diperoleh:

$$P(A) = \sum_{i=1}^m P(A|B_i) P(B_i) \quad (2.4)$$

Secara umum, untuk semua i berlaku:

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

Pembagian dengan $P(A)$ menghasilkan:

$$\frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{P(A)} \quad (2.5)$$

Dengan mensubstitusikan persamaan (2.4) pada persamaan (2.5), maka didapatkan:

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

Persamaan (2.6) merupakan rumusan matematis dari teorema Bayes. Peristiwa-peristiwa B_i membentuk m hipotesis *prior* yang digunakan untuk mempertimbangkan peristiwa A . $P(B_i)$ disebut peluang *prior*. Sedangkan $P(B_i|A)$ disebut peluang *posterior* untuk hipotesis yang sama. Peluang *posterior* ini adalah peluang terjadinya peristiwa B_i , setelah atau ketika peristiwa A terjadi (Tirta, 2004).

2.4 Algoritma EM (Ekspektasi Maksimisasi)

Algoritma EM merupakan sebuah metode iteratif untuk estimasi Maksimum Likelihood (ML) yang berguna dalam permasalahan data yang tidak lengkap (*missing data*). Dalam setiap iterasi pada Algoritma EM ini terdapat 2 tahap, yaitu tahap Ekspektasi atau tahap E (*E step*) dan tahap Maksimisasi atau tahap M (*M step*).

E step bertujuan menemukan ekspektasi bersyarat dari *missing data* dengan syarat data yang diketahui nilainya (*observed*) dan penduga parameternya, kemudian mensubstitusikan nilai ekspektasi yang diperoleh terhadap *missing data*. Dalam hal ini *missing data* yang dimaksud bukanlah Y_{miss} tapi fungsi dari Y_{miss} yang muncul dalam *complete data* loglikelihood, yaitu $\ell(\theta | Y)$. Sedangkan M *step* bertujuan untuk

memaksimumkan fungsi loglikelihood dengan cara mencari turunan parsial dari fungsi *log-likelihood* tersebut. Secara ringkas, langkah-langkah dalam algoritma EM sebagai berikut,

1. E-step : estimasi statistik cukup (*sufficient statistic*) untuk data lengkap Y_t dengan cara menghitung nilai ekspektasinya.
2. M-step: Tentukan $\theta^{(t+1)}$ dengan metode MLE (*Maximum Likelihood Estimation*) dari Y_t
3. Iterasi sampai nilai $\theta^{(t)}$ konvergen, atau $\theta^{(t+1)} - \theta^{(t)}$ mendekati nol. Hasilnya adalah barisan dari nilai-nilai $\theta^{(0)} \geq \theta^{(1)} \geq \dots$ dimulai dari suatu nilai $\theta^{(0)}$ tertentu.

(Budi, 2010).

2.5 Metode Newton-Raphson

Quinn (2001) menyatakan bahwa metode Newton-Raphson merupakan metode iteratif yang bisa digunakan untuk menghitung estimasi maksimum likelihood. Prodesur iteratif dari metode ini dilakukan sampai konvergen, sehingga didapatkan estimasi parameter yang stabil.

Menurut Tirta (2009), algoritma pokok dari metode Newton-Raphson ini dapat diuraikan sebagai berikut,

- a. Menentukan nilai awal, misalkan x_0
- b. melakukan iterasi dengan rumus,

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} \quad (2.7)$$

- c. melakukan iterasi sampai konvergen (sampai kriteria konvergensi terpenuhi).

Apabila peubah atau parameternya berdimensi tinggi, maka fungsi turunan pertamanya berupa vektor \mathbf{T} dan turunan keduanya berupa matrik yang disebut matrik Hessian \mathbf{H} . Bentuk multivariat dari metode Newton-Raphson yaitu sebagai berikut,

$$\mathbf{x}_1 = \mathbf{x}_0 - \mathbf{H}(\mathbf{x}_0)^{-1}\mathbf{T}(\mathbf{x}_0) \quad (2.8)$$

Misalkan $\mathbf{x} = (x_1, x_2, x_3, \dots, x_m)$, maka vektor \mathbf{T} dan matrik Hessian \mathbf{H} dapat ditunjukkan sebagai berikut:

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

2.6 Latent Class Analysis

Dalam kegiatan *latent class analysis* (analisis kelas laten), ada dua variabel yang digunakan dalam analisis tersebut yaitu variabel *manifest* (indikator) dan laten. Variabel *manifest* adalah variabel yang besaran kuantitatifnya dapat diketahui secara langsung, misalnya dari skor respon subjek terhadap instrumen pengukuran. Variabel laten adalah variabel yang nilai kuantitatifnya tidak dapat diketahui secara tampak (Widhiarso, 2011).

Vermunt dan Magidson (2003) menyatakan bahwa *latent class analysis* diperkenalkan Lazarfeld dan Henry (1968) sebagai suatu cara untuk merumuskan variabel laten dari survey dengan item dikotomous. Berbeda dengan analisis faktor yang menggunakan variabel laten bersifat kontinu, *latent class analysis* mengasumsikan variabel laten yang bersifat kategorik.

2.6.1 Pengertian Latent Class Model

Latent class model merupakan suatu model matematika yang menghubungkan probabilitas respon suatu individu untuk variabel indikator kategorik, dengan suatu variabel laten yang bersifat kategorik dalam beberapa kelompok (Novianti, 2008).

Misalkan J merupakan variabel manifest yang bersifat kategorik, tiap variabel manifest mempunyai K_j respon. Y_{ijk} merupakan hasil penelitian, dengan $j = 1, 2, \dots, J$, i dinotasikan sebagai individu ke- i ($i = 1, 2, \dots, N$) dan $k = 1, 2, \dots, K_j$. Dalam hal ini, $Y_{ijk} = 1$ jika individu ke- i memberikan respon ke- k pada variabel manifest ke- j , dan $Y_{ijk} = 0$ untuk yang lainnya.

Menurut Moutsaki dan Papageorgiu (2004), pada model kelas laten diasumsikan bahwa ruang vektor terdiri dari R kelompok. Untuk setiap kelompok r dihubungkan dengan p_r . Sebaran bersama dari peubah-peubah yang diamati adalah campuran terbatas (*finite mixture*) dari peluang,

$$P(Y_i|\pi, p) = \sum_{r=1}^R p_r f(Y_i; \pi_r) \quad (2.9)$$

dengan,

$f(Y_i|\pi_r)$ = Sebaran Y_i yang diberikan oleh parameter model π

p_r = peluang awal kelas laten atau kelompok R pada data y , dimana $p_r = \frac{n_r}{N}$

π_r =peluang suatu objek pada kelompok R

R = banyaknya kelompok ($r=1,2,\dots,R$), dimana $\sum_{r=1}^R \pi_r = 1$ dan $\sum_{r=1}^R p_r = 1$

Sebaran Y_i yang diberikan oleh parameter model π dapat dihasilkan (Linzer & Lewis, 2011) yaitu,

$$f(Y_i; \pi_r) = \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (2.10)$$

Sehingga fungsi kepadatan peluang dari semua kelas dapat dituliskan sebagai berikut,

$$P(Y_i|\pi, p) = \sum_{r=1}^R p_r \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (2.11)$$

dengan,

π_{jrk} = peluang suatu objek memberikan respon ke- k untuk variabel manifest ke- j dalam kelompok r

2.6.2 Estimasi Parameter *Latent Class Model*

Dua metode utama untuk menduga parameter-parameter pada model kelas laten adalah *maximum likelihood* (ML) dan metode Newton-Raphson. Fungsi log-likelihood yang disyaratkan pada pendekatan ML dapat diturunkan dari fungsi kepadatan peluang yang mendefinisikan model. Algoritma EM merupakan alat untuk

menduga ML dari parameter model kelas laten. Fungsi likelihood dari model kelas laten yaitu,

$$L = P(Y_1|\pi, p)P(Y_2|\pi, p) \dots P(Y_N|\pi, p)$$

$$= \prod_{i=1}^N \sum_{r=1}^R p_r \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (2.12)$$

Sehingga fungsi log-likelihood dari model kelas kelas tersebut dapat dinyatakan seperti di bawah ini,

$$\ln L = \ln \prod_{i=1}^N \sum_{r=1}^R p_r \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} = \sum_{i=1}^N \ln \sum_{r=1}^R p_r \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (2.13)$$

Menurut Nainggolan (2009), algoritma pertama mengikuti titik awal $\theta^{(0)} = (p^{(0)}, \pi^{(0)})$. Misalkan urutan pengulangan dinotasikan m . Proses pendugaan pada algoritma EM dimulai dengan iterasi dari EM. Setiap lingkaran proses pada algoritma EM terdiri dari dua langkah, yaitu pada langkah ekspektasi dan pemaksimalan dengan tahapan,

1. mendefinisikan nilai awal $\theta^{(0)} = (p^{(0)}, \pi^{(0)})$.
2. menghitung nilai dari:

$$P(Y_i|\pi, p)^{(m)} = \sum_{r=1}^R p_r^{(m)} \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}^{(m)}} \quad (2.14)$$

3. Tahapan E: menghitung $(\hat{P}(r_i|Y_i)^{(m)}, i = 1, \dots, n, r_i = 1, \dots, R)$ dimana $\hat{P}(r_i|Y_i)^{(m)}$ adalah peluang bersyarat yang menyatakan Y_i muncul dari R kelompok yang dirumuskan sebagai berikut,

$$\hat{P}(r_i|Y_i)^{(m)} = \frac{\hat{p}_r^{(m)} f(Y_i; \hat{\pi}_r)^{(m)}}{\sum_{q=1}^R \hat{p}_q^{(m)} f(Y_i; \hat{\pi}_q)^{(m)}} \quad (2.15)$$

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

$$\hat{p}_r^{(m+1)} = \frac{\sum_{i=1}^N \hat{P}(r_i|Y_i)^{(m)}}{N} \quad (2.16)$$

$$\hat{\pi}_r^{(m+1)} = \frac{\sum_{i=1}^N Y_{ij} \hat{P}(r_i|Y_i)^{(m)}}{\sum_{i=1}^N \hat{P}(r_i|Y_i)^{(m)}} \quad (2.17)$$

5. Ulangi tahap 2, 3 dan 4 sampai konvergen.

Ketika estimasi parameter dengan algoritma EM sudah konvergen, maka \hat{p}_r^{m+1} dan $\hat{\pi}_r^{m+1}$ yang baru dijadikan sebagai nilai dari \hat{p}_r dan $\hat{\pi}_r$.

2.7 Latent Class Regression Analysis

Latent class regression analysis merupakan pengembangan dari *latent class analysis* dengan cara melibatkan kovariat untuk memprediksi keanggotaan dalam kelompok. Analisis ini juga disebut teknik “satu langkah” untuk mengestimasi parameter akibat adanya kovariat, karena koefisien dari kovariat diestimasi secara simultan. Alternatif lain dari prosedur estimasi ini yaitu dengan menggunakan pendekatan “tiga langkah”. Prosedur pendekatan ini terdiri dari:

1. mengestimasi parameter model kelas laten;
2. menghitung prediksi peluang keanggotaan dalam kelompok;
3. menentukan model regresi kelas laten

(Linzer & Lewis, 2011).

Analisis ini mengasumsikan bahwa objek yang diamati berasal dari kelompok yang masih belum diketahui. Sehingga perlu dicari estimasi dari kelompok tersebut. Tiap kelompok mempunyai parameter regresi dan masing-masing objek yang diamati mempunyai probabilitas untuk menjadi anggota dari kelompok tersebut. Berbeda dengan *latent class analysis* yang mengasumsikan bahwa tiap objek mempunyai peluang awal kelas laten yang sama, *latent class regression analysis* mengasumsikan bahwa setiap objek mempunyai peluang kelas laten tergantung kovariat yang dimiliki oleh objek itu.

2.7.1 Pengertian *Latent Class Regression Model*

Misalkan X_i adalah pengamatan kovariat untuk responden i dan Y_i merupakan hasil pengamatan ke- i pada variabel manifest, β_r merupakan vektor dari koefisien yang menghubungkan kelompok ke- r , sedangkan p_{ri} adalah peluang suatu individu dalam keanggotaan kelas laten. Dalam hal ini, $\sum_r p_{ri} = 1$ untuk tiap individu yang diamati. Misal β_r dinotasikan sebagai vektor dari koefisien yang menghubungkan ke kelas laten ke- r . Dengan S kovariat, β_r mempunyai panjang $S + 1$, di sini satu koefisien dari setiap kovariat ditambah dengan sebuah konstanta. Karena kelas pertama digunakan sebagai acuan awal, maka β_1 didefinisikan bernilai nol. Sehingga,

$$\begin{aligned}\ln(p_{2i}/p_{1i}) &= X_i\beta_2 \\ \ln(p_{3i}/p_{1i}) &= X_i\beta_3 \\ &\vdots \\ \ln(p_{Ri}/p_{1i}) &= X_i\beta_R\end{aligned}$$

Sedangkan bentuk umumnya yaitu,

$$p_{ri} = p_r(X_i; \beta) = \frac{e^{X_i\beta_r}}{\sum_{q=1}^R e^{X_i\beta_q}} \quad (2.18)$$

$$p_r = \frac{1}{N} \sum_i \frac{e^{X_i\beta_r}}{\sum_{q=1}^R e^{X_i\beta_q}}$$

Estimasi parameter *latent class regression model* terdiri dari vektor dengan ukuran $R - 1$ yang berisi koefisien β_r , dan parameter seperti yang digunakan dalam *latent clas model* (p_r dan π_{jrk}). Rumus Bayes dari peluang suatu objek dalam *latent class regression model* yaitu,

$$\hat{P}(r_i|X_i; Y_i) = \frac{p_r(X_i; \hat{\beta})f(Y_i; \hat{\pi}_r)}{\sum_{q=1}^R p_q(X_i; \hat{\beta})f(Y_i; \hat{\pi}_q)} \quad (2.19)$$

2.7.2 Estimasi Parameter *Latent Class Regression Model*

Fungsi kepadatan peluang *latent class regression model* tidak jauh berbeda dengan persamaan (2.11), hanya saja p_r diganti dengan $p_r(X_i; \beta)$ seperti pada

persamaan (2.18). Fungsi kepadatan peluang dari *latent class regression model* dapat ditunjukkan seperti pada persamaan (2.20).

$$P(X_i; Y_i | \pi, p) = \sum_{r=1}^R p_r(X_i; \beta) \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \quad (2.20)$$

Fungsi likelihood dari persamaan (2.20) yaitu,

$$\begin{aligned} L &= P(X_1; Y_1 | \pi, p) P(X_2; Y_2 | \pi, p) \dots P(X_N; Y_N | \pi, p) \\ &= \prod_{i=1}^N \sum_{r=1}^R p_r(X_i; \beta) \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \end{aligned} \quad (2.21)$$

Sedangkan fungsi log-likelihood dari persamaan (2.20) yaitu,

$$\begin{aligned} \ln L &= \ln \prod_{i=1}^N \sum_{r=1}^R p_r(X_i; \beta) \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \\ &= \sum_{i=1}^N \ln \sum_{r=1}^R p_r(X_i; \beta) \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}} \end{aligned} \quad (2.22)$$

Proses estimasi parameter pada *latent class regression model* tidak jauh berbeda dengan *latent class model*, hanya saja ada penambahan kovariat dalam estimasinya. Proses estimasi parameter menggunakan algoritma EM dan metode Newton-Raphson dilakukan dengan langkah-langkah,

1. menendefinisikan nilai awal $\theta^{(0)} = (\pi_{jr}^{(0)}, \beta_r^0)$.
2. menghitung nilai dari,

$$p_{ri}^{(m)} = p_r(X_i; \beta)^{(m)} = \frac{e^{X_i \beta_r^{(m)}}}{\sum_{q=1}^R e^{X_i \beta_q^{(m)}} \quad (2.23)$$

$$P(Y_i | \pi, p)^{(m)} = \sum_{r=1}^R p_r(X_i; \beta)^{(m)} \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}^{(m)}} \quad (2.24)$$

3. Tahapan E: menghitung $(\hat{P}(r_i | X_i; Y_i)^{(m)}, i = 1, \dots, n, r_i = 1, \dots, R)$ dimana $\hat{P}(r_i | X_i; Y_i)^{(m)}$ adalah peluang bersyarat yang dirumuskan sebagai berikut,

$$\hat{P}(r_i|X_i; Y_i)^{(m)} = \frac{p_r(X_i; \hat{\beta})^{(m)} f(Y_i; \hat{\pi}_r)^{(m)}}{\sum_{q=1}^R p_q(X_i; \hat{\beta})^{(m)} f(Y_i; \hat{\pi}_q)^{(m)}} \quad (2.25)$$

4. Tahapan M : untuk mendapatkan nilai $\hat{\beta}_r$ dan $\hat{\pi}_{jr}$ dengan cara mencari maksimum likelihood dari persamaan (2.22), Algoritma dialihkan dengan menggunakan Algoritma Newton-Raphson, sehingga diperoleh penduga parameter yang baru,

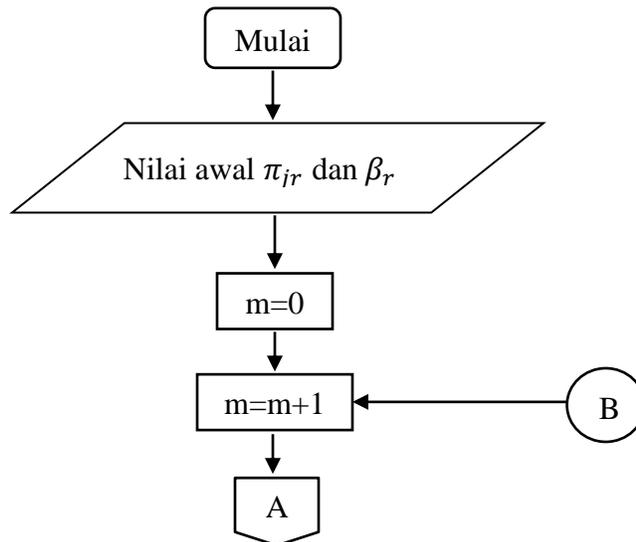
$$\hat{\beta}_r^{(m+1)} = \hat{\beta}_r^{(m)} - \mathbf{H}^{-1} \mathbf{T}_\beta \quad (2.26)$$

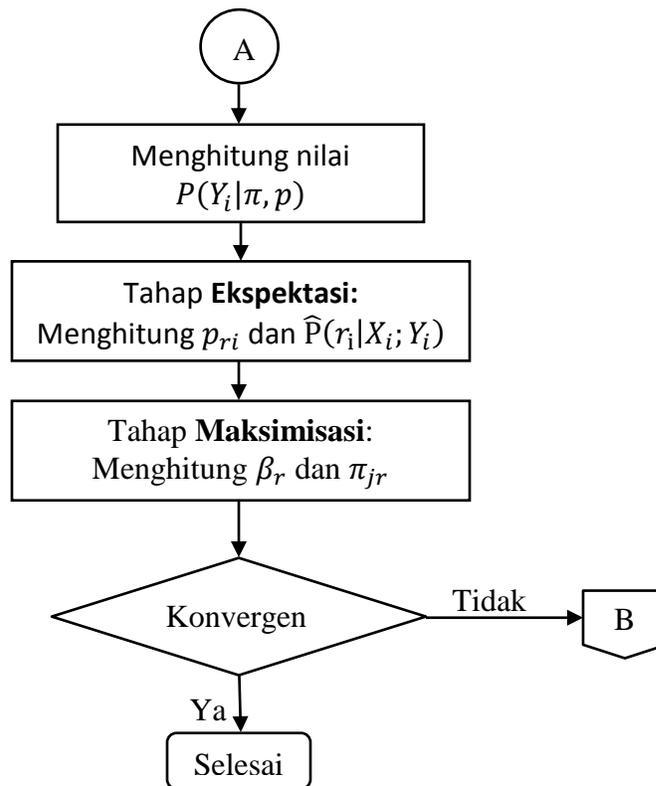
$$\hat{\pi}_{jr}^{(m+1)} = \frac{\sum_{i=1}^N Y_{ij} \hat{P}(r_i|X_i; Y_i)^{(m)}}{\sum_{i=1}^N \hat{P}(r_i|X_i; Y_i)^{(m)}} \quad (2.27)$$

Pada setiap iterasi, \mathbf{T} dinotasikan sebagai vektor gradien dari fungsi log-likelihood ke semua parameter yang dievaluasi pada $\hat{\theta}^m$. \mathbf{H} adalah matrik *Hessian* yang berisi turunan kedua dari seluruh parameter. Ketika suatu kriteria Newton-Raphson disesuaikan, $-\mathbf{H}^{-1}$ menghasilkan suatu penurunan dari log likelihood, tahapan ukuran dikurangi hingga tidak panjang. Matrik $-\mathbf{H}^{-1}$ dievaluasi sampai menghasilkan $\hat{\theta}$ akhir.

5. Ulangi tahap 2, 3, dan 4 sampai konvergen

Diagram dari algoritma EM untuk *latent class regression analysis* dapat ditunjukkan seperti pada Gambar 2.1.





Gambar 2.1. Skema Algoritma EM dalam *Latent Class Regression Analysis*

2.8 Uji Kecocokan Model

Asumsi kebebasan lokal merupakan asumsi yang mendasar dalam *latent class analysis*. Ketidaccocokan model dalam *latent class analysis* disebabkan oleh pelanggaran pada asumsi ini (Hanifah, 2010). Apabila ada lokal dependensi, maka hal ini dapat diatasi dengan menambah jumlah kelompok atau mengurangi variabel indikator sehingga didapatkan model yang cocok. Uji kecocokan model *latent regression class analysis* dapat dilakukan dengan beberapa uji statistik, diantaranya yaitu AIC (*Akaike's Information Criterion*) dan BIC (*Bayesian Information Criterion*).

2.8.1 AIC (*Akaike's Information Criterion*)

Menurut Fraley dan Raftery (1998) (dalam Nainggolan, 2008), Pendekatan paling luas dalam memilih kecocokan model pada *latent class analysis* yaitu dengan

menggunakan kriteria informasi Akaike (AIC, *Akaike's Information Criterion*) dengan rumus:

$$AIC = -2l(\hat{\theta}) + 2q \quad (2.28)$$

Dengan $l(\hat{\theta})$ adalah nilai maksimum *log-likelihood* dari model dan q adalah banyaknya parameter dalam model.

Secara umum, semakin kecil nilai AIC maka model yang dipakai semakin cocok. Model yang dianggap terbaik adalah model dengan nilai AIC minimum. Namun demikian, dengan pertimbangan aspek lain, perbedaan AIC yang tidak terlalu besar mungkin dapat diabaikan (Tirta, 2009).

2.8.2 BIC (*Bayesian Information Criterion*)

Selain menggunakan kriteria statistik AIC (*Akaike's Information Criterion*), untuk memilih model yang terbaik bisa digunakan juga BIC (*Bayesian Information Criterion*). Menurut Hanifah (2010), *Bayesian Information Criterion* (BIC) adalah sebuah kriteria statistik untuk memilih model. Nilai BIC mencerminkan peningkatan nilai jumlah kuadrat residual dan jumlah parameter dari model yang digunakan. Variasi yang tidak bisa dijelaskan oleh model mengakibatkan peningkatan BIC. Model dengan nilai BIC yang lebih kecil dipilih sebagai model yang terbaik, karena nilai BIC yang lebih kecil menunjukkan bahwa model yang dihasilkan lebih bisa menjelaskan variasi dari data. BIC yang dihitung berdasarkan *log likelihood* dirumuskan sebagai berikut:

$$BIC = -2l(\hat{\theta}) + \ln(N)q \quad (2.29)$$

Dengan $l(\hat{\theta})$ adalah nilai maksimum fungsi *log likelihood* dari suatu model yang diestimasi, N adalah banyaknya observasi dan q adalah banyaknya parameter.

BAB 3. METODE PENELITIAN

Pada bab ini, akan dijelaskan mengenai deskripsi serta analisis yang digunakan dalam penelitian ini. Metode penelitian ini meliputi data dan metode analisis data menggunakan *latent class regression analysis*.

3.1 Data

3.1.1 Sumber Data

Sumber data yang digunakan dalam penelitian ini yaitu data sekunder dari Santoso (2008) tentang penentuan strategi pemasaran pada distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Dalam penelitian tersebut, data dianalisis dengan menggunakan *EFAS (Eksternal Factor Analisis Summary)* dan Matrik *SWOT (Strength, Weakness, Opportunity, Threat)*. Mekanisme yang dilakukan dalam pembuatan matrik *SWOT* dimulai dari identifikasi peluang dan ancaman kemudian dirangkum dalam satu tabel yaitu *EFAS*. Dari analisis tersebut, kemudian ditentukan strategi pemasaran yang tepat bagi distributor tersebut.

Pada penelitian ini, data yang dianalisis hanya faktor eksternal yang mempengaruhi penentuan strategi pemasaran pada distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Adapun faktor eksternal tersebut yaitu *opportunity* dan *treath*.

3.1.2 Identifikasi variabel

Santoso (2008) menggunakan sampel berjumlah 107 outlet, masing-masing outlet diwakili oleh satu orang sebagai responden. Karena ada satu responden mengandung informasi yang kurang lengkap (satu responden tidak mengandung kovariat), maka data yang digunakan dalam penelitian ini sebanyak 106 responden.

Tingkat pengukuran dalam penelitian ini menggunakan skala *likert* dengan rentang skala yang digunakan untuk setiap indikator variabel adalah 1 (satu) sampai 5 (lima) dengan tingkatan sebagai berikut:

- a. Sangat setuju bernilai 5
- b. Setuju bernilai 4
- c. Cukup bernilai 3
- d. Tidak setuju bernilai 2
- e. Sangat tidak setuju bernilai 1

Skala *likert* ini digunakan sebagai penilaian responden terhadap variabel indikator dalam mengukur variabel laten. Sehingga, dari penilaian tersebut diketahui informasi mengenai pengelompokan objek berdasarkan karakteristik yang sama.

Dalam penelitian ini, variabel laten yang digunakan yaitu Peluang (*opportunity*) dan Ancaman (*treath*). Peluang (*opportunity*) merupakan faktor eksternal perusahaan yang tercipta dari kelemahan *competitor* dan merupakan keadaan konsumen yang dapat dimanfaatkan secara maksimal oleh perusahaan. Variabel indikator dari peluang ini dapat ditunjukkan seperti pada Tabel 3.1

Tabel 3.1 Definisi Operasional dari Peluang (*opportunity*)

No.	Variabel	Indikator pengukuran	Skala	Simbol
1.	Minat Masyarakat	Semakin diminatinya produk Unilever di masyarakat	Ordinal	Y11
2.	Modal	Kemudahan untuk mendapat modal untuk perluasan usaha	Ordinal	Y12
3.	Promosi	Promosi yang baik dari perusahaan pada setiap media informasi	Ordinal	Y13
4.	Kualitas Kerja	Pengembangan dan peningkatan kualitas kerja	Ordinal	Y14
5.	Produk Baru	Dapat menambah produk baru yang didistribusikan	Ordinal	Y15

Sedangkan ancaman (*treath*) merupakan faktor eksternal yang dapat merugikan perusahaan itu sendiri. Variabel indikator dari ancaman ini dapat dilihat pada Tabel 3.2.

Tabel 3.2 Definisi Operasional dari Ancaman (*treath*)

No.	Variabel	Indikator pengukuran	Skala	Simbol
1.	Pesaing	Munculnya pesaing dengan produk sejenis	Ordinal	Y21
2.	Harga	Pesaing memberikan harga lebih rendah atau murah	Ordinal	Y22
3.	Biaya	Semakin besarnya biaya operasional perusahaan	Ordinal	Y23
4.	Kondisi Perekonomian	Kondisi perekonomian tidak stabil karena kenaikan BBM	Ordinal	Y24
5.	Distributor Baru	Munculnya banyak distributor baru dengan produk baru dengan pelayanan yang variatif	Ordinal	Y25

Munculnya pesaing dengan produk sejenis dalam berbagai merk dapat mengganggu pendistribusian produk yang dilakukan perusahaan. Sedangkan adanya distributor baru merupakan ancaman yang muncul sehingga membuat para pelanggan berfikir dalam melakukan transaksi dengan perusahaan tersebut (Santoso, 2008).

Berdasarkan identitas responden, variabel yang menjadi pengamatan untuk kovariat yaitu lamanya responden berlangganan terhadap produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Dalam penelitian ini, kovariat tersebut diberi simbol X dengan ketentuan sebagai berikut:

- a. Lama berlangganan antara 1 sampai 5 tahun, X bernilai 1
- b. Lama berlangganan antara 6 sampai 10 tahun, X bernilai 2
- c. Lama berlangganan antara 11 sampai 15 tahun, X bernilai 3

- d. Lama berlangganan antara 16 sampai 20 tahun, X bernilai 4
- e. Lama berlangganan antara 21 sampai 25 tahun, X bernilai 5
- f. Lama berlangganan antara 26 sampai 30 tahun, X bernilai 6

3.2 Metode Analisis Data

Metode analisis yang digunakan dalam penelitian ini adalah *latent class regression analysis*. Adapun *software* yang digunakan yaitu *software open source* (paket R) versi 2.14.1. Tujuan dari penggunaan aplikasi ini yaitu untuk mengelompokkan responden pada kelas-kelas laten berdasarkan indikator-indikator. Nama paket R yang digunakan dalam penelitian ini adalah poLCA (*Polytomous Variable Latent Class Analysis*).

3.2.1 Paket poLCA (*Polytomous Variable Latent Class Analysis*)

Paket poLCA merupakan paket yang digunakan untuk mengestimasi model kelas laten dan model regresi kelas laten di program R. Paket poLCA bisa didapatkan dengan cara *men-download* di <http://CRAN.R-project.org/> atau di website <http://userwww.sservice.emory.edu/~dlinzer/poLCA>. Pada paket ini, *latent class regression analysis* berguna untuk menentukan keanggotaan kelas laten dan model regresi pada masing-masing kelompok.

3.2.2 Struktur Fungsi poLCA pada R

Model kelas laten mempunyai lebih dari satu variabel indikator, sehingga penulisan skrip mengenai variabel ini dalam *software* R misalnya `cbind(Y11, Y12, Y13, Y14, Y15)`. Pada *latent class regression model*, contoh fungsi yang digunakan yaitu:

```
formula <- cbind(Y11, Y12, Y13, Y14, Y15) ~ X
```

Dalam hal ini, (Y11, Y12, Y13, Y14, Y15) merupakan variabel indikator sedangkan X adalah kovariat untuk pengamatan. Struktur fungsi poLCA yang digunakan adalah sebagai berikut,

```
lc.peluang <- poLCA(formula, data, nclass = 2, maxiter = 1000,
  graphs = FALSE, tol = 1e-10, na.rm = TRUE,
  probs.start = NULL, nrep = 1, verbose = TRUE, calc.se
  = TRUE)
```

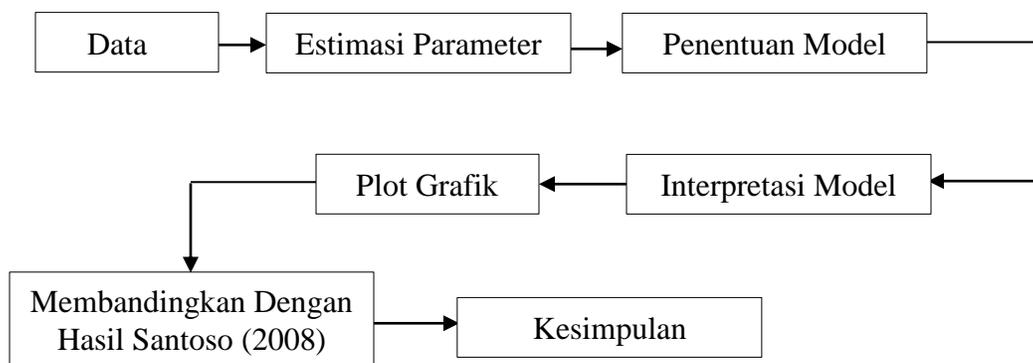
uraian lengkap mengenai perintah-perintah fungsi ini bisa dilihat pada lampiran A.

3.2.3 Penyajian Grafik *Latent Class Regression Model*

Pada model regresi kelas laten, anggota dari kelompok yang terbentuk merupakan objek yang dipengaruhi oleh kovariat. Probabilitas banyaknya objek dalam kovariat ke- i pada masing-masing kelompok dapat digambarkan dengan menggunakan fungsi `matplot()`.

3.2.4 Langkah-langkah Penelitian

Langkah-langkah yang akan dilakukan dalam penelitian ini dapat dilihat pada Gambar 3.1



Gambar 3.1 Langkah-langkah Penelitian

Langkah penelitian pada Gambar 3.1 dapat dijelaskan sebagai berikut:

1. Data

Data yang digunakan dalam penelitian ini seperti yang dijelaskan pada bagian 3.1.2. kemudian data tersebut dimasukkan dalam *software* program R.2.14.1

2. Estimasi parameter

Estimasi parameter ini menggunakan metode maksimum likelihood melalui algoritma EM (Ekspektasi Maksimisasi), kemudian dilanjutkan dengan metode Newton-Raphson. Parameter yang diestimasi yaitu p_r , π_{jrk} , p_{ri} , dan β_r dengan kelompok-kelompok yang berbeda. Parameter ini digunakan untuk mengetahui model dari klasifikasi responden berdasarkan *treath* dan *opportunity*.

3. Menentukan model yang cocok

Menentukan model yang cocok bisa dilakukan dengan membandingkan nilai AIC dan BIC pada model kelompok yang satu dengan kelompok yang lain. Semakin kecil nilai AIC atau BIC, maka semakin cocok model yang didapatkan.

4. Interpretasi model yang cocok

Dari model yang didapatkan pada langkah 2, kemudian menginterpretasikan masing-masing nilai parameter dari model tersebut.

5. Plot grafik keanggotaan kelas laten

Grafik ini digunakan untuk mengetahui pola hubungan peluang keanggotaan kelas laten berdasarkan kovariat (p_{ri}) pada masing-masing kelompok.

6. Membandingkan dengan hasil Santoso (2008)

Setelah analisis matematis tentang pengelompokan telah dilakukan, maka selanjutnya membandingkan hasil penelitian ini dengan hasil Santoso 2008. Dalam kasus ini, yang dibandingkan hanya terbatas pada *opportuntiy* dan *treath* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember.

7. Menarik kesimpulan

Kesimpulan bisa didapatkan berdasarkan hasil interpretasi pada langkah 4.

BAB 4. HASIL DAN PEMBAHASAN

Pada bab ini, akan dibahas mengenai hasil dari penelitian tentang *latent class regression analysis* untuk pengelompokan objek berdasarkan karakteristik yang sama. Selain variabel indikator, penggunaan kovariat dalam analisis ini merupakan hal penting yang akan mempengaruhi banyaknya objek dalam kelompok tersebut.

4.1 Estimasi Parameter dalam *Latent Class Regression Analysis*

Estimasi parameter merupakan sesuatu yang sangat diperlukan untuk mengetahui pengelompokan kelas laten dari masing-masing model yang diujikan. Model untuk *opportunity* dan *treath* yang diujikan masing-masing terdiri dari dua kelompok, tiga kelompok dan empat kelompok. Dengan menggunakan *software R*, parameter yang diestimasi antara lain p_r , π_{jrk} , β_r , dan p_{ri} .

Dalam hal ini, $\sum_{r=1}^R p_r = 1$ dengan p_r merupakan peluang awal kelas laten untuk kelompok ke- r dan R merupakan banyaknya kelompok pada masing-masing model. Adapun peluang kelas laten untuk penilaian responden terhadap *opportunity* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember dapat dilihat pada Tabel 4.1.

Tabel 4.1 Peluang Awal Kelas Laten Untuk Masing-masing Model Pada *Opportunity*

R	p_1	p_2	p_3	p_4
2	0.9253	0.0747		
3	0,1869	0,4646	0,3485	
4	0,1773	0,2147	0,2109	0,3972

Sedangkan peluang awal kelas laten untuk penilaian responden terhadap *treath* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember dapat dilihat pada Tabel 4.2.

Tabel 4.2 Peluang Awal Kelas Laten untuk Masing-masing Model pada *Treath*

R	p_1	p_2	p_3	p_4
2	0,7377	0,2623		
3	0,665	0,1966	0,1384	
4	0,1702	0,1722	0,0727	0,5849

4.2 Uji Kecocokan Model

Setelah model yang didapatkan, kemudian menentukan model yang cocok berdasarkan nilai AIC dan BIC masing-masing model. Semakin kecil nilai tersebut, maka semakin baik model yang didapatkan. Nilai AIC dan BIC untuk *opportunity* dan *treath* masing-masing dapat dilihat pada Tabel 4.3 dan 4.4.

Tabel 4.3 Uji Kecocokan Model pada *Opportunity*

R	AIC	BIC
2	1259,058	1370,922
3	1185,911	1356,371
4	1189,734	1418,79

Pada Tabel 4.3, dapat diketahui bahwa model empat kelompok memiliki nilai BIC paling tinggi yaitu sebesar 1418,79, sedangkan nilai BIC paling kecil dimiliki oleh model tiga kelompok yaitu sebesar 1356,371. Selain itu, model dua kelompok memiliki nilai AIC paling tinggi sebesar 1259,058, sedangkan model yang memiliki AIC paling kecil yaitu model tiga kelompok sebesar 1185,911. Sehingga model yang cocok dalam rangka pengelompokan responden berdasarkan penilaian terhadap *opportunity* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember adalah model 3 kelompok.

Tabel 4.4 Uji Kecocokan Model Pada *Treath*

<i>R</i>	AIC	BIC
2	1071,605	1183,469
3	1066,678	1237,138
4	1081,826	1310,882

Berdasarkan informasi pada Tabel 4.4, dapat diketahui bahwa model 4 kelompok memiliki nilai BIC paling tinggi yaitu sebesar 1310,882, sedangkan nilai BIC paling kecil dimiliki oleh model dua kelompok yaitu sebesar 1183,469. sehingga model yang cocok bernilai BIC yaitu model dua kelompok. Selain itu, nilai AIC paling tinggi dimiliki oleh model empat kelompok sebesar 1081,826, sedangkan nilai AIC paling kecil dimiliki oleh model tiga kelompok sebesar 1066,678. Banyaknya kelompok yang direkomendasikan oleh BIC berbeda dengan model pada AIC terkecil. Dalam kasus ini, model yang cocok dalam rangka pengelompokan responden berdasarkan penilaian terhadap *treath* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember adalah model 2 kelompok.

4.3 Estimasi Peluang Kelas Laten Berdasarkan Kelompok Dan Variabel Indikator

Pada Subbab 4.1, telah ditunjukkan peluang awal kelas laten masing-masing kelompok. Setiap kelompok mengandung informasi yang berkaitan dengan variabel indikator dan pilihan jawaban responden. Peluang kelas laten (π_{jrk}) dihitung berdasarkan variabel indikator dan pilihan jawaban responden. Dalam hal ini, $\sum_k^{K_j} \pi_{jrk} = 1$ dengan π_{jrk} merupakan peluang kelas laten ketika responden menjawab pilihan ke- k untuk variabel indikator ke- j pada masing-masing kelas dan K_j merupakan banyak pilihan jawaban untuk variabel indikator ke- j .

4.3.1. Peluang Kelas Laten Berdasarkan Kelompok dan Variabel Indikator untuk *Opportunity*

Pada Tabel 4.1, diketahui peluang awal kelas laten untuk *opportunity* pada masing-masing kelompok. Setelah ditentukan model yang cocok dari nilai AIC dan BIC, hasil estimasi π_{jrk} untuk model tiga kelompok untuk *opportunity* dapat ditunjukkan pada Tabel 4.5.

Tabel 4.5 Hasil Estimasi π_{jrk} untuk Model Tiga Kelompok pada *Opportunity*

r	j	π_{jrk}				
		π_{jr1}	π_{jr2}	π_{jr3}	π_{jr4}	π_{jr5}
1	1	0	0,0000	0,000	0,5035	0,4965
	2	0	0,0000	0,0000	1,0000	0,0000
	3	0	0,1952	0,0000	0,8048	0,0000
	4	0,0000	0,4968	0,5032	0,0000	0,0000
	5	0	0,2019	0,2225	0,4828	0,0928
2	1	0	0,2843	0,264	0,4518	0,0000
	2	0	0,0406	0,3452	0,6142	0,0000
	3	0	0,0000	0,2031	0,7969	0,0000
	4	0,0406	0,0000	0,4552	0,3939	0,1103
	5	0	0,2234	0,3550	0,4216	0,0000
3	1	0	0,0000	0,000	0,6707	0,3293
	2	0	0,0000	0,0000	0,3503	0,6497
	3	0	0,0577	0,0541	0,2655	0,6227
	4	0,0000	0,0584	0,1249	0,6389	0,1778
	5	0	0,0000	0,1112	0,5054	0,3834

Dari hasil estimasi pada Tabel 4.5 dapat diketahui bahwa banyaknya responden pada kelompok pertama ketika menjawab variabel indikator pertama dengan pilihan jawaban keempat yaitu sebanyak 50,35%, sedangkan yang menjawab variabel indikator pertama dengan pilihan jawaban kelima sebanyak 49,65%.

Pada tabel 4.5 dapat dilihat bahwasannya pada kelompok pertama dan kelompok ketiga tidak ada responden yang menggunakan pilihan jawaban pertama untuk menjawab masing-masing variabel indikator. Sedangkan responden yang menggunakan pilihan pertama hanya terdapat pada kelompok kedua ketika menjawab variabel indikator yaitu sebesar 4,06%.

4.3.2. Peluang Kelas Laten Berdasarkan Kelompok dan Variabel Indikator untuk *Treath*

Hasil estimasi π_{jrk} untuk model dua kelompok pada *treath* didapatkan seperti pada Tabel 4.6.

Tabel 4.6 Hasil Estimasi π_{jrk} untuk Model Dua Kelompok pada *Treath*

r	j	π_{jrk}				
		π_{jr1}	π_{jr2}	π_{jr3}	π_{jr4}	π_{jr5}
1	1	0	0,0000	0,1968	0,6753	0,1279
	2	0	0,0256	0,1964	0,7013	0,0767
	3	0	0,0000	0,2625	0,6608	0,0767
	4	0,0000	0,0000	0,1815	0,7929	0,0256
	5	0,0000	0	0,0778	0,5298	0,3924
2	1	0	0,0719	0,1299	0,7263	0,0719
	2	0	0,1438	0,7782	0,0779	0,0000
	3	0	0,4315	0,4847	0,0838	0,0000
	4	0,0719	0,2877	0,6404	0,0000	0,0000
	5	0,0719	0	0,3566	0,4882	0,0833

Dari Tabel 4.6 dapat diketahui bahwa pada kelompok pertama sebagian besar responden menggunakan pilihan keempat ketika menjawab masing-masing variabel indikator. Hal ini dapat dilihat dari banyaknya responden yang menggunakan pilihan keempat untuk variabel indikator, masing-masing sebesar 67,53%, 70,13%, 66,08%,

79,29% dan 52,98%. Hal yang membedakan antara kelompok pertama dan kelompok kedua yaitu sebagian besar responden menggunakan pilihan keempat untuk menjawab variabel indikator kedua, ketiga dan keempat, sedang-kan pada kelompok dua sebagian besar responden menggunakan pilihan ketiga.

4.4 Model dalam *Latent Class Regression Analysis*

Parameter lain yang didapatkan ketika melakukan proses estimasi dalam analisis regresi kelas laten yaitu $\hat{\beta}_r$. Nilai $\hat{\beta}_r$ ini yang yang menjadi intersep dan koefisien X dalam persamaan regresi kelas laten. Adapun hasil dari estimasi $\hat{\beta}_r$ untuk *opportunitiy* dan *treath* ditunjukkan pada Subsubbab 4.4.1 dan 4.4.2.

4.4.1 Regresi Kelas Laten untuk *Opportunity*.

Estimasi $\hat{\beta}_r$ dari regresi kelas laten pada model tiga kelompok untuk *opportunity* ditunjukkan padat tabel 4.7.

Tabel 4.7 Regresi Kelas Laten pada Model Tiga Kelompok untuk *Opportunity*

		Estimasi	Std. error	Nilai t	p-value
$\left(\frac{p_{2i}}{p_{1i}}\right)$	Intersep	3,78331	1,24893	3,029	0,004
	X	-0,72031	0,26127	-2,757	0,009
$\left(\frac{p_{3i}}{p_{1i}}\right)$	Intersep	3,72920	1,28980	2,891	0,006
	X	-0,79664	0,28517	-2,794	0,008

Persamaan regresi kelas laten dari Tabel 4.7 dapat ditulis sebagai berikut:

$$\ln\left(\frac{p_{2i}}{p_{1i}}\right) = 3,78331 - 0,72031X_i \quad (4.1a)$$

$$\ln\left(\frac{p_{3i}}{p_{1i}}\right) = 3,72920 - 0,79664X_i \quad (4.1b)$$

Persamaan regresi antara kelompok dua dan kelompok pertama pada model tiga kelompok untuk *opportunity* mempunyai intersep dan koefisien dari X yang sangat signifikan. Hal ini disebabkan nilai *p-value* masing-masing lebih besar dari

5%, yaitu sebesar 0,4% dan 0,9%. Persamaan regresi kelas laten antara kelompok ketiga dan kelompok pertama juga mempunyai intersep dan koefisien dari X yang sangat signifikan, dengan *p-value* masing-masing yaitu 0,6% dan 0,8%.

4.4.2. Model Regresi Kelas Laten untuk *Treath*

Selain dari hasil estimasi $\hat{\beta}_r$ pada model tiga kelompok untuk *opportunity*, didapatkan juga hasil estimasi $\hat{\beta}_r$ pada model dua kelompok untuk *treath* seperti ditunjukkan pada Tabel 4.8.

Tabel 4.8 Regresi Kelas Laten pada Model Dua Kelompok untuk *Treath*

		Estimasi	Std. error	Nilai t	<i>p-value</i>
$\left(\frac{p_{2i}}{p_{1i}}\right)$	Intersep	-1,70139	0,73266	-2,322	0,023
	X	0,19027	0,18473	1,030	0,307

Persamaan regresi kelas laten pada model dua kelompok untuk *treath* dapat dinyatakan seperti di bawah ini:

$$\ln\left(\frac{p_{2i}}{p_{1i}}\right) = -1,70139 + 0,19027X_i \quad (4.2)$$

Persamaan regresi antara kelompok kedua dan kelompok pertama pada model dua kelompok untuk *treath* mempunyai intersep yang signifikan dan koefisien X yang tidak signifikan, dengan nilai *p-value* masing-masing sebesar 2,3% dan 30,7%. Sehingga persamaan regresi kelas laten pada persamaan (4.3) dapat dinyatakan seperti di bawah ini:

$$\ln\left(\frac{p_{2i}}{p_{1i}}\right) = -1,70139 \quad (4.3)$$

4.5 Peluang Kelas Laten Berdasarkan Kovariat

Setelah hasil estimasi $\hat{\beta}_r$ dari masing-masing model untuk *opportunity* dan *treath* didapatkan, maka bisa ditentukan pula nilai dari peluang kelas laten berdasarkan kovariat yaitu p_{ri} (untuk kelompok ke-*r* dan kovariat ke-*i*). Hubungan

antara p_{ri} dan $\hat{\beta}_r$ ditunjukkan pada persamaan (2.18). Dalam hal ini, untuk kovariat ke- i berlaku $\sum_{r=1}^R p_{ri} = 1$.

4.5.1. Peluang Kelas Laten Berdasarkan Kovariat untuk *Opportunity*

Besarnya nilai $\hat{\beta}_r$ pada persamaan (4.1) menentukan juga besarnya p_{ri} . Adapun hasil estimasi p_{ri} pada model dua kelompok untuk *opportunity* ditunjukkan pada Tabel 4.9.

Tabel 4.9 Estimasi p_{ri} pada Model Tiga Kelompok untuk *Opportunity*

i	p_{1i}	p_{2i}	p_{3i}
1	0,02429119	0,5196242	0,4560846
2	0,05031704	0,5237541	0,4259289
3	0,10120026	0,5125848	0,3862149
4	0,19285563	0,4753223	0,3318220
5	0,33613358	0,4031249	0,2607416
6	0,51724991	0,3018561	0,1808940

Berdasarkan tabel 4.9 didapatkan bahwa untuk level kovariat satu sampai lima sebagian besar responden berada pada kelompok kedua. Tetapi untuk kovariat keenam, sebagian besar responden berada pada kelompok pertama.

4.5.2. Peluang Kelas Laten Berdasarkan Kovariat untuk *Treath*

Hasil estimasi p_{ri} pada model dua kelompok untuk *treath* didapatkan dengan menggunakan persamaan (4.3) dan (4.4), sehingga dihasilkan seperti pada Tabel 4.10.

Tabel 4.10 Estimasi p_{ri} pada Model Dua Kelompok untuk *Treath*

i	(Persamaan 4.2)		(Persamaan 4.3)	
	p_{1i}	p_{2i}	p_{1i}	p_{2i}
1	0,8192265	0,1807735	0,8457162	0,1542838
2	0,7893217	0,2106783	0,8457162	0,1542838

3	0,7559436	0,2440564	0,8457162	0,1542838
4	0,7191589	0,2808411	0,8457162	0,1542838
5	0,6791829	0,3208171	0,8457162	0,1542838
6	0,6363935	0,3636065	0,8457162	0,1542838

Estimasi p_{ri} yang dihitung menggunakan persamaan (4.2) menunjukkan bahwa sebagian besar responden berada di kelompok pertama. Lebih dari 60% responden berada di kelompok 1 untuk masing-masing kovariat.

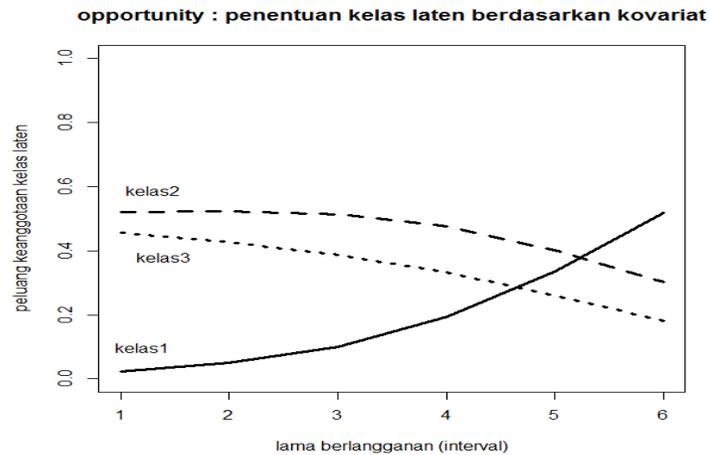
Hasil estimasi p_{ri} yang dihitung menggunakan persamaan (4.3) menunjukkan bahwa sebagian besar responden berada di kelompok pertama juga. Akan tetapi pada hasil ini besarnya kovariat tidak mempengaruhi persentase responden dalam kovariat untuk masing-masing kelompok. Hal ini dikarenakan nilai koefisien dari X pada persamaan (4.3) bernilai 0.

4.6 Plot Model Regresi Kelas Laten

Plot model ini merupakan hasil dari peluang kelas laten berdasarkan kovariat sebagaimana yang dicantumkan pada subbab 4.5. Dari hasil plot ini, akan diketahui trend dari masing masing kelompok.

4.6.1. Plot Model Regresi Kelas Laten untuk *Opportunity*

Tabel 4.9 memberikan informasi p_{ri} untuk *opportunity* model tiga kelompok. Gambar 4.1 berikut menunjukkan plot dari p_{ri} untuk model tersebut.

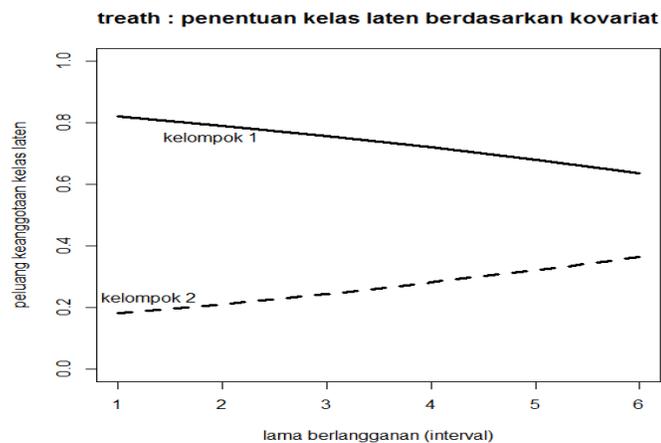


Gambar 4.1 Plot Regresi Kelas Laten Pada Model Tiga Kelompok Untuk *Opportunity*

Dari Gambar 4.1, diketahui bahwa grafik dari kelompok pertama dan kelompok kedua semakin turun seiring semakin besarnya kovariat, turunnya trend ini semakin kelihatan pada kovariat empat sampai enam. Namun kelompok ketiga menunjukkan semakin naik seiring bertambahnya kovariat.

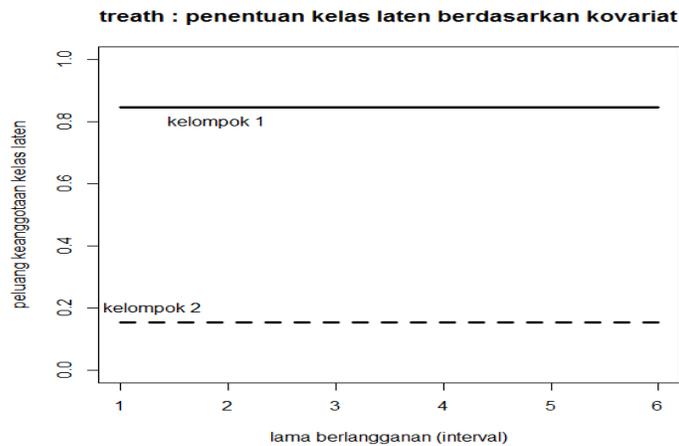
4.6.2. Plot Regresi Kelas Laten untuk *Treath*

Plot pada Subsubbab 4.6.2 didapatkan dari peluang kelas laten berdasarkan kovariat pada Subsubbab 4.5.2. Plot ini ditunjukkan di Gambar 4.3 dan 4.4.



Gambar 4.2 Plot Regresi Kelas Laten pada Model Dua Kelompok untuk *Treath* (persamaan 4.5)

Kelompok pertama mengalami penurunan persentase seiring bertambah besarnya kovariat, akan tetapi tidak signifikan. Kelompok kedua mengalami kenaikan seiring bertambahnya kovariat.



Gambar 4.3 Plot Regresi Kelas Laten pada Model Dua Kelompok untuk *Treath* (persamaan 4.6)

Pada plot yang ditunjukkan pada Gambar 4.3, diketahui bahwa grafik dari masing-masing kelompok adalah konstan. Hal ini menunjukkan bahwa besarnya kovariat tidak berpengaruh terhadap persentase responden pada masing-masing kelompok.

4.7 Pembahasan

Santoso (2008) menyatakan bahwa responden setuju apabila semakin diminatinya produk Unilever di masyarakat, kemudahan mendapatkan modal dalam perluasan usaha, promosi yang baik dari perusahaan pada setiap media informasi, pengembangan dan peningkatan kualitas kerja, dan dapat menambah produk baru yang didistribusikan merupakan *opportunity* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Responden juga menyatakan setuju apabila munculnya pesaing dengan produk yang sejenis, pesaing memberikan harga yang lebih rendah, semakin besarnya biaya operasional

perusahaan, kondisi perekonomian yang tidak stabil karena kenaikan BBM, dan munculnya banyak distributor baru dengan produk baru dan pelayanan variatif merupakan *treath* dari distributor tersebut.

Informasi mengenai *opportunity* dan *treath* ini akan dianalisis lebih lanjut berdasarkan variabel indikator dan kovariat dengan menggunakan *latent class regression analysis*.

Pada *opportunity*, responden dikelompokkan menjadi tiga kelompok. Kelompok pertama sebanyak 18,69%, kelompok kedua sebanyak 46,46% dan kelompok ketiga sebanyak 34,85%. Masing-masing kelompok mempunyai persentase responden berdasarkan variabel indikator seperti yang ditunjukkan pada Tabel 4.11.

Tabel 4.11 Pengelompokan Responden pada Masing-masing Kelompok untuk *Opportunity*

Variabel Indikator	Kelompok 1	Kelompok 2	Kelompok 3
Minat masyarakat	50,35% (setuju)	45,18% (setuju)	67,07% (setuju)
Modal	100% (setuju)	61,42% (setuju)	64,97% (sangat setuju)
Promosi	80,48 (setuju)	79,69% (setuju)	62,27% (sangat setuju)
Kualitas kerja	49,68% (tidak setuju)	45,52% (cukup)	63,89% (setuju)
Produk baru	48,28% (setuju)	42,16% (setuju)	50,54% (setuju)

Berdasarkan Tabel 4.11, kelompok pertama terdiri dari responden yang setuju bahwasannya minat masyarakat, modal, promosi dan produk baru merupakan *opportunity* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember, dan tidak setuju apabila kualitas kerja menjadi *opportunity* distributor tersebut. Kelompok kedua terdiri dari responden yang setuju bahwasannya minat masyarakat, modal, promosi dan produk baru merupakan *opportunity* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember, dan menyatakan cukup apabila kualitas kerja menjadi

opportunity distributor tersebut. Kelompok ketiga terdiri dari responden yang menyatakan setuju apabila minat masyarakat, kualitas kerja dan produk baru merupakan *opportunity* distributor tersebut, dan sangat setuju apabila modal dan promosi merupakan *opportunity* yang dimiliki oleh distributor tersebut.

Informasi pada Tabel 4.11 dapat memperjelas bahwasannya distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember harus benar-benar mempertimbangkan minat masyarakat, modal, promosi, kualitas kerja dan produk baru dalam rangka pengembangan strategi perusahaan untuk pemasaran produk Unilever.

Berdasarkan lamanya berlangganan, banyaknya responden pada masing-masing kelompok ditunjukkan seperti pada Tabel 4.12.

Tabel 4.12 Pengelompokan Responden Berdasarkan Kovariat Untuk *Opportunity*

Lama Berlangganan (tahun)	Kelompok 1	Kelompok 2	Kelompok 3
1 – 5	0	6	3
6 – 10	1	16	13
11 – 15	2	13	8
16 – 20	3	6	7
20 – 25	4	5	1
26 – 30	10	5	3

Secara umum, responden di kelompok kedua lebih banyak dari pada kelompok yang lain. Jumlah responden di kelompok pertama, kedua dan ketiga masing-masing sebanyak 20, 51 dan 35. Jumlah responden di kelompok kedua lebih banyak dari pada kelompok lain untuk lamanya berlangganan antara 1 sampai 15 tahun, sedangkan kelompok pertama lebih banyak daripada kelompok lain untuk lamanya berlangganan antara 26 sampai 30 tahun. Kelompok kedua dengan lamanya berlangganan antara 6 sampai 10 tahun merupakan kelompok yang mempunyai responden paling banyak daripada kelompok dengan lama berlangganan yang lain,

yaitu sebanyak 16 orang. Selain itu, diketahui pula bahwa tidak ada responden yang berada di kelompok pertama dengan lama berlangganan antara 1 sampai 5 tahun.

Pada *treath*, responden dikelompokkan menjadi dua kelompok. Kelompok pertama sebanyak 73,77% dan kelompok kedua sebanyak 26,23%. Masing-masing kelompok mempunyai persentase responden berdasarkan variabel indikator seperti yang dinyatakan pada Tabel 4.13.

Tabel 4.13 Pengelompokan Responden pada Masing-masing Kelompok untuk *Treath*

Variabel Indikator	Kelompok 1	Kelompok 2
Pesaing	67,53% (setuju)	72,63% (setuju)
Harga	70,13% (setuju)	77,82% (cukup)
Biaya	66,08% (setuju)	48,47% (cukup)
Kondisi Perekonomian	79,29% (setuju)	64,04% (cukup)
Distributor Baru	52,98% (setuju)	48,82% (setuju)

Berdasarkan Tabel 4.13, kelompok pertama terdiri dari responden yang menyatakan setuju apabila pesaing memberikan harga lebih murah, semakin besarnya biaya operasional perusahaan dan kondisi perekonomian yang tidak stabil karena kenaikan BBM merupakan *treath* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Sedangkan kelompok kedua menyatakan cukup apabila tiga variabel indikator ini merupakan *treath* yang dimiliki oleh distributor tersebut.

Informasi pada Tabel 4.13 memperjelas bahwa adanya pesaing, pesaing memberikan harga lebih murah, semakin besarnya biaya operasional perusahaan, kondisi perekonomian yang tidak stabil karena kenaikan BBM dan munculnya banyak distributor baru merupakan *treath* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember.

Pengelompokan responden berdasarkan lamanya berlangganan pada masing-masing kelompok dapat ditunjukkan pada Tabel 4.14.

Tabel 4.14 Pengelompokan Responden Berdasarkan Kovariat Untuk *Treath*

Lama Berlangganan (tahun)	Kelompok 1	Kelompok 2
1 – 5	9	0
6 – 10	24	6
11 – 15	15	8
16 – 20	10	6
20 – 25	7	3
26 – 30	13	5

Pada masing-masing lamanya berlangganan, responden di kelompok pertama lebih banyak daripada di kelompok kedua. Kelompok yang memiliki responden paling banyak yaitu kelompok pertama dengan lama berlangganan antara 6 sampai 10 tahun, sebanyak 24 orang. Di kelompok kedua, tidak ada responden yang lama berlangganan antara 1 sampai 5 tahun.

Apabila responden dikelompokkan berdasarkan persepsinya terhadap *opportunity* dan *treath*, maka didapatkan hasil seperti pada Tabel 4.16.

Tabel 4.15 Pengelompokan Responden Berdasarkan Kovariat Untuk *Treath* dan *Opportunity*

Lama Berlangganan (tahun)	Kelompok Berdasarkan <i>Treath</i>	Kelompok Berdasarkan <i>Opportunity</i>		
		1	2	3
1 – 5	1	0	6	3
	2	0	0	0
6 – 10	1	0	13	11
	2	1	3	2
11 – 15	1	2	6	7
	2	0	7	1
16 – 20	1	3	3	4
	2	0	3	3

20 – 25	1	3	3	1
	2	1	2	0
26 – 30	1	7	3	3
	2	3	2	0

Berdasarkan Tabel 4.15, jumlah responden paling banyak berada di kelompok kedua untuk *opportunity* dengan kelompok pertama untuk *treath* pada lama berlangganan antara 6 sampai 10 tahun, sebanyak 13 orang. Pada kelompok ketiga untuk *opportunity*, tidak ada responden pada kelompok kedua untuk *treath* yang berlangganan antara 1 sampai 5 tahun, antara 20 sampai 25 tahun dan 26 sampai 30 tahun.

BAB 5. PENUTUP

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan pada Bab 4, dapat ditarik kesimpulan sebagai berikut:

1. Persepsi responden terhadap *opportunity* distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember dapat dikelompokkan menjadi tiga kelompok, sedangkan *treath* menjadi dua kelompok.
2. Pada *opportunity*, jumlah responden di kelompok kedua lebih banyak dari pada kelompok lain untuk lamanya berlangganan antara 1 sampai 15 tahun, sedangkan kelompok satu lebih banyak daripada kelompok lain untuk lamanya berlangganan antara 26 sampai 30 tahun. Selain itu, kelompok satu mengalami kenaikan jumlah responden seiring semakin tingginya level lamanya berlangganan. Pada *treath*, jumlah responden di kelompok satu lebih banyak daripada di kelompok dua untuk masing-masing lamanya berlangganan.
3. Minat masyarakat, modal, promosi, kualitas kerja dan produk baru merupakan *opportunity* yang dimiliki oleh distributor produk Unilever di PT. Panahmas Dwitama Distrindo Regional Jember. Sedangkan munculnya pesaing dengan produk sejenis, pesaing memberikan harga lebih murah, semakin besarnya biaya operasional perusahaan, kondisi perekonomian yang tidak stabil dan munculnya distributor baru dengan produk baru merupakan *treath* yang dimiliki oleh distributor tersebut.

5.2 Saran

Pada skripsi ini, analisis yang digunakan adalah *latent class regression analysis* untuk data kategorik dengan satu kovariat. Ada kesempatan bagi para peneliti untuk

mengembangkan analisis ini, misalnya dengan menggunakan data kategorik dengan dua kovariat atau lebih.

DAFTAR PUSTAKA

- Budi. 2010. *Analisis Missing Data Menggunakan Algoritma EM*.
<http://statistikakomputasi.wordpress.com/2010/04/08/analisis-missing-data-menggunakan-algoritma-em-2/> [3 April 2012]
- Candra, Y. 2009. Pembentukan Model Probit Bivariat. Skripsi. Semarang: FMIPA Universitas Diponegoro
- Hanifah, E. 2010. Metode Latent Class Cluster Untuk Variabel Indikator Bertipe Campuran Dalam Rangka Pengelompokan Desa. Tesis. Bandung: Universitas Padjadjaran
- Linzer, D. A, & Lewis, J. B. 2011. poLCA: An R Package for Polytomous Variable Latent Class Analysis. *Journal of Statistical Software*: **42**: 10: 3- 12
- Magidson, J. & Vermunt, J. K. 2003. *Latent Class Models*. Artikel. Statistical Innovations Inc
- Moustaki, I. & Papageorgiou .2004. *Latent Class Models for Mixed Variables with Applications in Archaeometry*. Yunani: Athens University
- Nainggolan, B. M. H. 2009. Perbandingan Analisis Laten Kelas dengan Kriteria WHO Untuk Penggerombolan Pasien Demam Dangué (DD) dan Demam Berdarah Dangué (DBD). Tesis. Bogor: Institut Pertanian Bogor
- Novianti. 2008. Latent Class Model. Skripsi. Depok: Universitas Indonesia

- Priatna, B. A. Tanpa tahun. *Teknik-Teknik Analisis Multivariat Terkini Yang Sering Digunakan Dalam Penelitian*. Yogyakarta: Universitas Pendidikan Indonesia
- Quinn, K. 2001. *The Newton Raphson Algorithm for Function Optimization*. USA: University of Washington
- Santoso, P. A. B. 2008. Penentuan Strategi Pemasaran Pada Distributor Produk Unilever Di PT. Panahmas Dwitama Distrindo Regional Jember. Skripsi. Jember: Universitas Jember
- Tirta, I. M. 2004. *Pengantar Statistika Matematika*. Jember: FMIPA Universitas Jember
- Tirta, I. M. 2009. *Analisis Regresi dengan R*. Jember: Jember University Press
- Vermunt, J. K. 2004. *Latent Class Regression Analysis*. Jena University
- Widhiarso, W. 2011. Berkenalan Dengan Variabel Laten. Tidak Dipublikasikan. Yogyakarta: Universitas Gajah Mada
- Zhang, N. L. Hierarchical Latent Class Model for Cluster Analisis. *Journal Of Machine Learning Research* 5: 4: 02: 697

LAMPIRAN A. SKRIP *poLCA* DALAM *SOFTWARE R*

```
poLCA(formula, data, nclass = 2, maxiter = 1000, graphs =  
FALSE, tol = 1e-10, na.rm = TRUE, probs.start = NULL,  
nrep = 1, verbose = TRUE, calc.se = TRUE)
```

`nclass` : banyaknya kelas-kelas laten yang dibentuk dalam model

`maxiter` : iterasi (pengulangan) maksimum ketika menjalankan algoritma, seperti algoritma EM.

`graphs` : memberikan tampilan (grafik) dari hasil estimasi model. Jika `TRUE`, maka grafik akan ditampilkan. Jika `FALSE`, maka tidak ditampilkan.

`tol` : nilai toleransi kekonvergenan dalam algoritma

`na.rm` : mengatasi kasus untuk variabel manifest yang responnya ada yang tidak ada. Jika `TRUE`, maka variabel manifest dalam kasus ini akan dihapus sebelum estimasi model dijalankan.

`prob.start` : nilai awal dalam rangka menjalankan proses estimasi menggunakan algoritma EM

`nreps` : banyaknya model yang ingin didapatkan, sehingga bisa dipilih model yang cocok berdasarkan besarnya maksimum log-likelihood.

`verbose` : mengeluarkan hasil model dalam layar. Jika `FALSE`, tidak ada keluaran yang dihasilkan.

`calc.se` : menghitung standard error dari model. Jika `FALSE`, maka mengestimasi model tanpa adanya variabel cocok yang dikhususkan dalam formula.

`condition` : menggambarkan respon dari variabel manifest untuk mencari distribusi frekuensi dari suatu variabel manifest yang lain.

LAMPIRAN B. KUESIONER UNTUK EKSTERNAL PELANGGAN (OUTLET)

Responden

Usia :.....Tahun

Lama Berlangganan :

Jenis kelamin : Laki-laki/Perempuan

Alamat Distrik :

EFAS (<i>Eksternal Strategic Factors Analysis Summary</i>)	Bobot				
	Sangat Tidak Setuju (1)	Tidak Setuju (2)	Cukup (3)	Setuju (4)	Sangat Setuju (5)
<u>Peluang</u> 1. Semakin diminatinya produk Unilever di masyarakat 2. Kemudahan mendapatkan modal untuk perluasan usaha 3. Promosi yang baik dari perusahaan pada setiap media informasi 4. Pengembangan dan peningkatan kualitas kerja 5. Dapat menambah produk baru yang didistribusikan <u>Ancaman</u> 1. Munculnya pesaing dengan produk yang sejenis 2. Pesaing memberikan harga yang lebih rendah atau murah 3. Semakin besarnya biaya operasional perusahaan 4. Kondisi perekonomian tidak stabil karena kenaikan BBM 5. Munculnya banyak distributor baru dengan produk baru dengan pelayanan yang variatif					

Sumber: Santoso (2008)

LAMPIRAN C. SKRIP PROGRAM DAN OUTPUT UNTUK *OPPORTUNITY*
MENGUNAKAN *SOFTWARE R*

a. Skrip program untuk model dua kelompok

```
> fpeluang<-cbind(Y11,Y12,Y13,Y14,Y15)~X
> lclpeluang<-poLCA(fpeluang,PASAR,nclass=2)
Conditional item response (column) probabilities,
  by outcome variable, for each class (row)
$Y11
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:    0 0.0819 0.1127 0.5811 0.2243
class 2:    0 0.7538 0.2462 0.0000 0.0000

$Y12
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:    0 0.0204 0.1531 0.5818 0.2447
class 2:    0 0.0000 0.2501 0.7499 0.0000

$Y13
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:    0 0.0612 0.0416 0.6627 0.2345
class 2:    0 0.0000 1.0000 0.0000 0.0000

$Y14
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0000 0.1223 0.3570 0.3983 0.1223
class 2: 0.2526 0.0000 0.2501 0.4973 0.0000

$Y15
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:    0 0.1529 0.1844 0.4996 0.1631
class 2:    0 0.0000 1.0000 0.0000 0.0000

Estimated class population shares
0.9253 0.0747

Predicted class memberships (by modal posterior prob.)
0.9245 0.0755

=====
Fit for 2 latent classes:
=====
2 / 1
      Coefficient Std. error t value Pr(>|t|)
(Intercept)  -2.85489    1.12693   -2.533    0.014
X              0.09659    0.29414    0.328    0.744
=====

number of observations: 106
number of estimated parameters: 42
residual degrees of freedom: 64
maximum log-likelihood: -587.5289

AIC(2): 1259.058
BIC(2): 1370.922
X^2(2): 1376.35 (Chi-square goodness of fit)
```

b. Skrip program untuk model tiga kelompok

```

> fpeluang<-cbind(Y11,Y12,Y13,Y14,Y15)~X
> lclpeluang<-poLCA(fpeluang,PASAR,nclass=3)
Conditional item response (column) probabilities,
  by outcome variable, for each class (row)
$Y11
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.0000 0.000 0.5035 0.4965
class 2:   0 0.2843 0.264 0.4518 0.0000
class 3:   0 0.0000 0.000 0.6707 0.3293

$Y12
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.0000 0.0000 1.0000 0.0000
class 2:   0 0.0406 0.3452 0.6142 0.0000
class 3:   0 0.0000 0.0000 0.3503 0.6497

$Y13
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.1952 0.0000 0.8048 0.0000
class 2:   0 0.0000 0.2031 0.7969 0.0000
class 3:   0 0.0577 0.0541 0.2655 0.6227

$Y14
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0000 0.4968 0.5032 0.0000 0.0000
class 2: 0.0406 0.0000 0.4552 0.3939 0.1103
class 3: 0.0000 0.0584 0.1249 0.6389 0.1778

$Y15
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.2019 0.2225 0.4828 0.0928
class 2:   0 0.2234 0.3550 0.4216 0.0000
class 3:   0 0.0000 0.1112 0.5054 0.3834

Estimated class population shares
0.1869 0.4646 0.3485

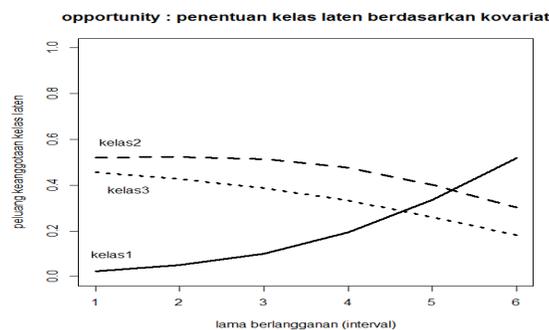
Predicted class memberships (by modal posterior prob.)
0.1887 0.4811 0.3302

=====
Fit for 3 latent classes:
=====
2 / 1
      Coefficient Std. error t value Pr(>|t|)
(Intercept)   3.78331   1.24893   3.029   0.004
X              -0.72031   0.26127  -2.757   0.009
=====
3 / 1
      Coefficient Std. error t value Pr(>|t|)
(Intercept)   3.72920   1.28980   2.891   0.006
X              -0.79664   0.28517  -2.794   0.008
=====
number of observations: 106
number of estimated parameters: 64

```

```
residual degrees of freedom: 42
maximum log-likelihood: -528.9556
```

```
AIC(3): 1185.911
BIC(3): 1356.371
X^2(3): 600.7943 (Chi-square goodness of fit)
> matplot(c(1:6), (cbind(1,exb)/(1+rowSums(exb))),
+ main="opportunity : penentuan kelas laten berdasarkan kovariat",
+ xlab="lama berlangganan (interval)",ylab="peluang keanggotaan kelas laten",
+ ylim=c(0,1),type="l",lwd=3,col=1)
> text(1.2,0.1,"kelas1")
> text(1.3,0.59,"kelas2")
> text(1.4,0.38,"kelas3")
```



```
> lclpeluang$predclass
[1] 3 1 1 2 3 2 2 1 3 2 2 2 2 2 2 2 3 2 3 2 1 1 2 1 3 3 3 3 1 1 1 2 1
[38] 3 3 2 2 2 3 3 2 2 3 2 2 2 2 1 2 3 2 2 2 1 3 2 3 2 2 2 2 2 2 3 2 3 2 1 1 2
[75] 1 3 3 3 3 3 3 1 1 1 2 1 3 3 2 2 2 3 3 2 2 3 1 2 2 3 2 2 2 3 2 2
> cbind(1,exb)/(1+rowSums(exb))
      [,1]      [,2]      [,3]
[1,] 0.02429119 0.5196242 0.4560846
[2,] 0.05031704 0.5237541 0.4259289
[3,] 0.10120026 0.5125848 0.3862149
[4,] 0.19285563 0.4753223 0.3318220
[5,] 0.33613358 0.4031249 0.2607416
[6,] 0.51724991 0.3018561 0.1808940
```

c. Skrip program untuk model empat kelompok

```
> fpeluang<-cbind(Y11,Y12,Y13,Y14,Y15)~X
> lclpeluang<-poLCA(fpeluang,PASAR,nclass=4)
Conditional item response (column) probabilities,
by outcome variable, for each class (row)
$Y11
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.0000 0.0000 1.0000 0.0000
class 2:   0 0.0000 0.0000 0.4728 0.5272
class 3:   0 0.0000 0.0000 0.5526 0.4474
class 4:   0 0.3325 0.3088 0.3587 0.0000

$Y12
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.1064 0.0980 0.7295 0.066
class 2:   0 0.0000 0.0000 0.0000 1.000
class 3:   0 0.0000 0.0000 1.0000 0.000
```

class 4: 0 0.0000 0.3601 0.6399 0.000

\$Y13

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0	0.1064	0.0000	0.5147	0.3788
class 2:	0	0.0000	0.0879	0.2143	0.6978
class 3:	0	0.1790	0.0000	0.8210	0.0000
class 4:	0	0.0000	0.2375	0.7625	0.0000

\$Y14

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0.0000	0.0000	0.0000	0.9442	0.0558
class 2:	0.0000	0.0879	0.1757	0.4740	0.2624
class 3:	0.0000	0.4474	0.5526	0.0000	0.0000
class 4:	0.0475	0.0000	0.4905	0.3437	0.1183

\$Y15

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0	0.0000	0.1080	0.7681	0.1239
class 2:	0	0.0000	0.0866	0.4005	0.5128
class 3:	0	0.1790	0.2421	0.4895	0.0895
class 4:	0	0.2613	0.3940	0.3447	0.0000

Estimated class population shares

0.1773 0.2147 0.2109 0.3972

Predicted class memberships (by modal posterior prob.)

0.1792 0.2264 0.2264 0.3679

=====
Fit for 4 latent classes:
=====

2 / 1

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	1.03351	1.00190	1.032	0.315
X	-0.27612	0.31934	-0.865	0.397

3 / 1

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	-2.04769	1.23232	-1.662	0.112
X	0.55385	0.26924	2.057	0.053

4 / 1

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	1.18448	0.74814	1.583	0.129
X	-0.11757	0.20484	-0.574	0.572

=====
number of observations: 106
number of estimated parameters: 86
residual degrees of freedom: 20
maximum log-likelihood: -508.8669

AIC(4): 1189.734

BIC(4): 1418.79

X^2(4): 423.8201 (Chi-square goodness of fit)

LAMPIRAN D. SKRIP PROGRAM DAN OUTPUT UNTUK *TREATH* MENGGU-
NAKAN *SOFTWARE R*

a. Skrip program untuk model dua kelompok

```
> fancam<-cbind(Y21,Y22,Y23,Y24,Y25)~X
> lclancam<-poLCA(fancam,PASAR,nclass=2)
Conditional item response (column) probabilities,
by outcome variable, for each class (row)

$Y21
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.0000 0.1968 0.6753 0.1279
class 2:   0 0.0719 0.1299 0.7263 0.0719

$Y22
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.0256 0.1964 0.7013 0.0767
class 2:   0 0.1438 0.7782 0.0779 0.0000

$Y23
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.0000 0.2625 0.6608 0.0767
class 2:   0 0.4315 0.4847 0.0838 0.0000

$Y24
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0000 0.0000 0.1815 0.7929 0.0256
class 2: 0.0719 0.2877 0.6404 0.0000 0.0000

$Y25
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.0000   0 0.0778 0.5298 0.3924
class 2: 0.0719   0 0.3566 0.4882 0.0833

Estimated class population shares
0.7377 0.2623

Predicted class memberships (by modal posterior prob.)
0.7358 0.2642

=====
Fit for 2 latent classes:
=====
2 / 1
      Coefficient Std. error t value Pr(>|t|)
(Intercept)  -1.70139   0.73266  -2.322   0.023
X              0.19027   0.18473   1.030   0.307
=====
number of observations: 106
number of estimated parameters: 42
residual degrees of freedom: 64
maximum log-likelihood: -493.8025

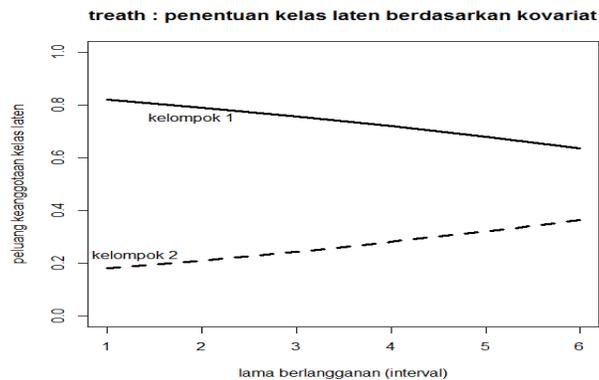
AIC(2): 1071.605
BIC(2): 1183.469
```

```
X^2(2): 1222.058 (Chi-square goodness of fit)
```

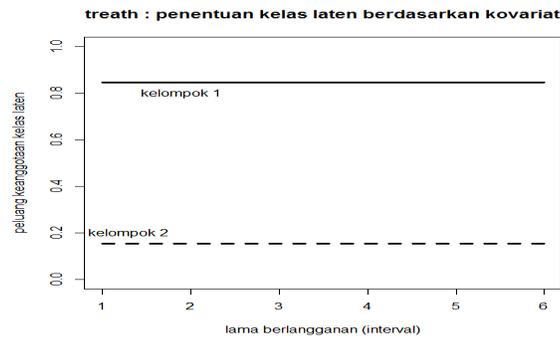
```
> ancmat <- cbind(1,c(1:6))
> exb <- exp(ancmat %*% lclancam$coeff)
> cbind(1,exb)/(1+rowSums(exb))
      [,1]      [,2]
[1,] 0.8192265 0.1807735
[2,] 0.7893217 0.2106783
[3,] 0.7559436 0.2440564
[4,] 0.7191589 0.2808411
[5,] 0.6791829 0.3208171
[6,] 0.6363935 0.3636065

> lclancam$predclass
[1] 2 1 1 1 1 1 1 2 1 1 1 1 2 1 1 1 1 2 1 1 2 2 1 1 1 1 1 1 1 1 1 2 1 1 2 1
[38] 1 2 1 2 2 1 1 2 2 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 2 1 1 2 2 1 1
[75] 1 1 1 1 1 1 1 2 1 1 2 1 1 2 1 2 2 1 1 2 2 1 2 1 1 2 1 1 1 1 1 1 1 1
```

```
> matplot(c(1:6), (cbind(1,exb)/(1+rowSums(exb))),
+         main="treath : penentuan kelas laten berdasarkan kovariat",
+         xlab="lama berlangganan (interval)",
+         ylab="peluang keanggotaan kelas laten",
+         ylim=c(0,1),type="l",lwd=3,col=1)
text(1.9,0.75,"kelompok 1")
text(1.3,0.23,"kelompok 2")
```



```
exb <- exp(ancmat %*% matrix(c(-1.70139,0),2,1))
matplot(c(1:6), (cbind(1,exb)/(1+rowSums(exb))),
+         main="treath : penentuan kelas laten berdasarkan kovariat",
+         xlab="lama berlangganan (interval)",
+         ylab="peluang keanggotaan kelas laten",
+         ylim=c(0,1),type="l",lwd=3,col=1)
text(1.9,0.80,"kelompok 1")
text(1.3,0.20,"kelompok 2")
```



```
> cbind(1,exb)/(1+rowSums(exb))
      [,1]      [,2]
[1,] 0.8457162 0.1542838
[2,] 0.8457162 0.1542838
[3,] 0.8457162 0.1542838
[4,] 0.8457162 0.1542838
[5,] 0.8457162 0.1542838
[6,] 0.8457162 0.1542838
```

b. Skrip program untuk model tiga kelompok

```
> fancam<-cbind(Y21,Y22,Y23,Y24,Y25)~X
> lclancam<-poLCA(fancam,PASAR,nclass=3)
Conditional item response (column) probabilities,
by outcome variable, for each class (row)
```

```
$Y21
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.096 0.1780 0.7260 0.0000
class 2:   0 0.000 0.0000 0.7274 0.2726
class 3:   0 0.000 0.2169 0.6696 0.1135
```

```
$Y22
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.1919 0.7040 0.1041 0.0000
class 2:   0 0.0000 0.5911 0.0000 0.4089
class 3:   0 0.0284 0.1937 0.7779 0.0000
```

```
$Y23
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:   0 0.2454 0.6505 0.1041 0.0000
class 2:   0 0.4692 0.0000 0.5308 0.0000
class 3:   0 0.0000 0.2900 0.6249 0.0851
```

```
$Y24
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.096 0.3838 0.5202 0.0000 0.0000
class 2: 0.000 0.0000 0.8637 0.0000 0.1363
class 3: 0.000 0.0000 0.1204 0.8796 0.0000
```

```
$Y25
      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.096   0 0.2406 0.6634 0.0000
```

```
class 2: 0.000      0 0.3329 0.0000 0.6671
class 3: 0.000      0 0.0866 0.5841 0.3293
```

```
Estimated class population shares
0.1966 0.1384 0.665
```

```
Predicted class memberships (by modal posterior prob.)
0.1887 0.1509 0.6604
```

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

```
2 / 1
```

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	0.32493	1.29249	0.251	0.803
X	-0.18864	0.29229	-0.645	0.522

```
3 / 1
```

	Coefficient	Std. error	t value	Pr(> t)
(Intercept)	1.97910	1.00829	1.963	0.056
X	-0.21425	0.22701	-0.944	0.351

```
=====  
number of observations: 106  
number of estimated parameters: 64  
residual degrees of freedom: 42  
maximum log-likelihood: -469.3391
```

```
AIC(3): 1066.678  
BIC(3): 1237.138  
X^2(3): 332.9499 (Chi-square goodness of fit)
```

```
> probs<-poLCA.reorder(lclancam$probs.start, order(lclancam$P,  
decreasing = TRUE))
```

```
> lc2ancam<-poLCA(fancam,PASAR,nclass=3,probs.start=probs)
```

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

```
$Y21
```

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0	0.000	0.2169	0.6696	0.1135
class 2:	0	0.096	0.1780	0.7260	0.0000
class 3:	0	0.000	0.0000	0.7274	0.2726

```
$Y22
```

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0	0.0284	0.1937	0.7779	0.0000
class 2:	0	0.1919	0.7040	0.1041	0.0000
class 3:	0	0.0000	0.5911	0.0000	0.4089

```
$Y23
```

	Pr(1)	Pr(2)	Pr(3)	Pr(4)	Pr(5)
class 1:	0	0.0000	0.2900	0.6249	0.0851
class 2:	0	0.2454	0.6505	0.1041	0.0000
class 3:	0	0.4692	0.0000	0.5308	0.0000

```
$Y24
```

```

          Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.000 0.0000 0.1204 0.8796 0.0000
class 2: 0.096 0.3838 0.5202 0.0000 0.0000
class 3: 0.000 0.0000 0.8637 0.0000 0.1363

$Y25
          Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.000      0 0.0866 0.5841 0.3293
class 2: 0.096      0 0.2406 0.6634 0.0000
class 3: 0.000      0 0.3329 0.0000 0.6671

Estimated class population shares
0.665 0.1966 0.1384

Predicted class memberships (by modal posterior prob.)
0.6604 0.1887 0.1509

=====
Fit for 3 latent classes:
=====
2 / 1
          Coefficient Std. error t value Pr(>|t|)
(Intercept)   -1.97910    1.00829   -1.963    0.056
X              0.21425    0.22701    0.944    0.351
=====
3 / 1
          Coefficient Std. error t value Pr(>|t|)
(Intercept)   -1.65417    0.75153   -2.201    0.033
X              0.02561    0.21510    0.119    0.906
=====
number of observations: 106
number of estimated parameters: 64
residual degrees of freedom: 42
maximum log-likelihood: -469.3391

AIC(3): 1066.678
BIC(3): 1237.138
X^2(3): 332.9499 (Chi-square goodness of fit)

```

c. Skrip program untuk model empat kelompok

```

> fancam<-cbind(Y21,Y22,Y23,Y24,Y25)~X
> lclancam<-poLCA(fancam,PASAR,nclass=4)
Conditional item response (column) probabilities,
by outcome variable, for each class (row)

$Y21
          Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:      0 0.1109 0.2218 0.6674 0.0000
class 2:      0 0.0000 0.0000 0.8904 0.1096
class 3:      0 0.0000 0.0000 0.7405 0.2595
class 4:      0 0.0000 0.2419 0.6290 0.1290

$Y22
          Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:      0 0.2218 0.7782 0.0000 0.0000

```

```

class 2:      0 0.0000 0.1235 0.5478 0.3287
class 3:      0 0.0000 1.0000 0.0000 0.0000
class 4:      0 0.0323 0.2097 0.7581 0.0000

```

\$Y23

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1:    0 0.1163 0.7728 0.1109 0.0000
class 2:    0 0.1202 0.1129 0.7670 0.0000
class 3:    0 1.0000 0.0000 0.0000 0.0000
class 4:    0 0.0000 0.2903 0.6129 0.0968

```

\$Y24

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.1109 0.4435 0.4456    0 0.0000
class 2: 0.0000 0.0000 0.8904    0 0.1096
class 3: 0.0000 0.0000 1.0000    0 0.0000
class 4: 0.0000 0.0000 0.0000    1 0.0000

```

\$Y25

```

      Pr(1) Pr(2) Pr(3) Pr(4) Pr(5)
class 1: 0.1109    0 0.2272 0.6619 0.0000
class 2: 0.0000    0 0.0000 0.3320 0.6680
class 3: 0.0000    0 0.7656 0.0000 0.2344
class 4: 0.0000    0 0.0968 0.5968 0.3065

```

Estimated class population shares

```
0.1702 0.1722 0.0727 0.5849
```

Predicted class memberships (by modal posterior prob.)

```
0.1698 0.1698 0.0755 0.5849
```

```
=====
Fit for 4 latent classes:
=====
```

2 / 1

```

      Coefficient Std. error t value Pr(>|t|)
(Intercept)  -0.06321    1.01207   -0.062    0.951
X              0.02065    0.26391    0.078    0.938
=====

```

3 / 1

```

      Coefficient Std. error t value Pr(>|t|)
(Intercept)  -0.87702    1.43500   -0.611    0.548
X              0.00740    0.37798    0.020    0.985
=====

```

4 / 1

```

      Coefficient Std. error t value Pr(>|t|)
(Intercept)   1.75843    0.87366    2.013    0.058
X             -0.15335    0.23383   -0.656    0.519
=====

```

number of observations: 106

number of estimated parameters: 86

residual degrees of freedom: 20

maximum log-likelihood: -454.913

AIC(4): 1081.826

BIC(4): 1310.882

X²(4): 314.7547 (Chi-square goodness of fit)