UNIVERSITAS INDONESIA

PENENTUAN BANYAK KELAS LATEN OPTIMAL PADA *LATENT PROFILE MODEL*

SKRIPSI

RIZQI MARLINDA
0606067780

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
DESEMBER 2010

UNIVERSITAS INDONESIA

PENENTUAN BANYAK KELAS LATEN OPTIMAL PADA *LATENT PROFILE MODEL*

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana sains

RIZQI MARLINDA
0606067780

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
DESEMBER 2010

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya sendiri,

dan semua sumber baik yang dikutip maupun dirujuk

telah saya nyatakan dengan benar.

Nama	:	Rizqi Marlinda
NPM	:	0606067780
Tanda Tangan	:	
Tanggal	:	22 Desember 2010

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :
Nama : Rizqi Marlinda
NPM : 0606067780
Program Studi : Sarjana Matematika
Judul Skripsi : Penentuan Banyak Kelas Laten Optimal pada *Latent Profile Model*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Sarjana Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia

DEWAN PENGUJI

Pembimbing	: Dra. Rianti Setiadi, M. Si.	( )
Pembimbing	: Fevi Novkaniza, S.Si, M. Si.	( )
Penguji	: Dra. Rianti Setiadi, M. Si.	( )
Penguji	: Dr. Sri Mardiyati, M.Kom	(Sri Mardiyati)
Penguji	: Dra. Siti Nurrohmah, M.Si	( )

Ditetapkan di : Depok
Tanggal : 22 Desember 2010

KATA PENGANTAR

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan skripsi ini. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Sience Jurusan Matematika pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

(1) Ibu Dra. Rianti Setiadi, M.Si dan Mba Fevi Novkaniza, M.Si., selaku dosen pembimbing yang telah menyediakan waktu, tenaga, dan pikiran untuk mengarahkan saya dalam penyusunan skripsi ini.

(2) Kedua orang tua tercinta yang telah mendidik penulis dengan penuh kasih sayang, selalu mendoakan dan memberikan motivasi kepada penulis.

(3) Adik tercinta, terima kasih atas pengertiannya karena tidak mengganggu penulis dalam menyelesaikan tugas akhir ini.

(4) Ibu Dr. Kiki Ariyanti Sugeng selaku pembimbing akademis penulis selama menjalani masa kuliah.

(5) Bapak dan Ibu dosen yang telah hadir dan memberikan saran-saran kepada penulis mulai dari sig 1 sampai kolokium, Ibu Dr. Dian Lestari, Ibu Dra. Rustina, Ibu Dra Saskya Mary, M. Si, Mba Mila Novita, M. Si, Ibu Dra. Titin Siswantining, DEA dan Ibu Ida Fitriani.

(6) Ibu Dra. Siti Nurrahmah, M.Si dan Ibu Dr. Sri Mardiyati, M. Kom selaku penguji sidang.

(7) Seluruh dosen beserta staf Departemen Matematika FMIPA UI atas kesabaran dan bimbingannya.

(8) Nadya dan Rahmanita yang telah berjuang bersama selama penyusunan tugas akhir ini. Terima kasih karena bersedia datang ke kampus untuk menemani penulis.

(9) Dhita, Alfa, Nisa, Dian, Farah, Mella, Tami, Widya, Milla, Nurgi, Ita, Puspa, Putri, Lena, Yunita, Rita P., dan lainnya. Terimakasih atas kebersamaannya selama ini.

(10) Teman-teman 2006 yang telah lulus, Ar, Oppie, Inne, Mei, Rahanti, Stefani, Anggha, Lee, Yuri, Rita Y., Syafirah, Tasya, Nobo, Tika, Lani, Poe, Bekti, Reza, Yuko, Tino, Teguh, Tisna, Oza, Bara, Doddy, Indra, Rafli, Rontu, Latief, Hot, Rifza, Rama, Rendy, Budi, Aliman, Michael, Stefano, Billy, Pangky, Dani, Ali.

(11) Kak Yanu, yang sudah sangat membantu penulis dalam menyelesaikan tugas akhir ini.

(12) Kak Murni dan Kak Handi, atas bantuannya sehingga terselesaikannya tugas akhir ini.

(13) Kak Bong dan Albert, atas bantuannya untuk mencarikan penulis jurnal-jurnal untuk mendukung tugas akhir ini.

Penulis juga mengucapkan terima kasih kepada pihak yang tidak dapat disebutkan satu per satu, yang telah membantu dalam penyusunan tugas akhir ini. Akhir kata, penulis mohon maaf apabila terdapat kesalahan atau kekurangan dalam tugas akhir ini. Semoga tugas akhir ini membawa manfaat bagi perkembangan ilmu.

Penulis

2010

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI
TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini:

Nama	:	Rizqi Marlinda
NPM	:	0606067780
Program Studi	:	Sarjana Matematika
Departemen	:	Matematika
Fakultas	:	Matematika dan Ilmu Pengetahuan Alam
Jenis karya	:	Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :
Penentuan Banyak Kelas Laten Optimal pada *Latent Profile Model*.

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok
Pada tanggal : 22 Desember 2010
Yang menyatakan

Marlinda

(Rizqi Marlinda)

ABSTRAK

Nama : Rizqi Marlinda
Program Studi : S1 Matematika
Judul : Penentuan Banyak Kelas Laten Optimal pada *Latent Profile Model*

Tugas akhir ini secara umum bertujuan untuk membahas *latent profile model* yaitu suatu model yang menghubungkan sejumlah variabel indikator yang bersifat kontinu dengan variabel laten kategorik yang dibentuknya. Kelas-kelas dari variabel laten pada *latent profile model* disebut kelas laten. Tiap kelas laten memiliki profil yang dapat diwakili oleh vektor mean dan vektor variansi dari variabel indikator pada tiap kelas. Dalam analisis laten profil, yang akan dilakukan adalah membentuk kelas dari variabel laten berdasarkan sejumlah variabel indikator kontinu sedemikian sehinga di dalam setiap kelas laten, variabel-variabel indikator akan saling bebas, kemudian menentukan mean dan variansi (profil) dari variabel-variabel indikator pada setiap kelas laten. Penaksiran parameter dalam *latent profile mode*l menggunakan taksiran maksimum likelihood, yang diselesaikan dengan algoritma EM (*Expectation-Maximization*). Kecocokan model dan banyaknya kelas laten optimal dalam *latent profile model* diuji dengan uji rasio likelihood. Metode tersebut akan diterapkan untuk membentuk kategori dari variabel laten “tingkat mengatur diri sendiri” berdasarkan variabel indikator “tingkat ketaktergantungan”, “skor tanggung jawab, dan “tingkat ketenangan” pada mahasiswa baru matematika FMIPA UI angkatan tahun 2010. Hasil analisis data menunjukkan bahwa tingkat mengatur diri sendiri pada mahasiswa baru matematika FMIPA UI angkatan 2010 dapat dikategorikan menjadi 2 kelas.

Kata kunci : *latent profile model*, variabel laten, algoritma EM, taksiran maksimum likelihood, uji rasio likelihood.
xii+51 halaman : 4 tabel
Daftar Pustaka : 10 (1976-2008)

ABSTRACT

Name : Rizqi Marlinda
Study Program : S1 Mathematics
Title : Determining the Number of Optimal Latent Classes in Latent Profile Model

This final project, in general aims to discuss the latent profile model. The latent profile model is a model that links a number of indicator variables that are continuous with the establishment of categorical latent variables. The classes of latent variables in latent profile model called latent class. Each latent class has a profile that can be represented by the vector mean and vector variance of indicator variables in each class. In the latent profile analysis to be done is to form a class of latent variable based on the number of variables indicators such continuous so that within each latent class, the indicator variables are independent, then determine the mean and variance (profile) of the indicator variable in each latent class. Parameters estimation in latent profile model using maximum likelihood estimation, which is solved by EM algorithm (Expectation-Maximization). The goodness of fit for model and the number of optimal latent classes in latent profile model was tested with likelihood ratio test. This method will be applied to form the category of latent variable "level of self management" based on the indicator variables "level of independency", "responsibility score, and " level of tranquility " in new students of Mathematic FMIPA UI 2010. The results of data analysis showed that levels of self management for new students of Mathematic FMIPA UI 2010 can be categorized into 2 classes.

Key words : latent profile model, latent variable, EM algorithm, maximum likelihood estimator, ratio likelihood test.
xii+51 pages : 4 tables
Bibliography : 10 (1976-2008)

DAFTAR ISI

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB 1
PENDAHULUAN

1.1 Latar Belakang

Dalam beberapa disiplin ilmu, seperti pendidikan, psikologi dan kesehatan, banyak penelitian dimana nilai dari variabel yang akan diamati atau diteliti tidak dapat diukur secara langsung, melainkan diukur melalui beberapa variabel lain. Variabel yang tidak dapat diamati secara langsung ini disebut sebagai variabel laten. Sedangkan variabel yang mengukur variabel laten dan dapat diukur secara langsung disebut sebagai variabel indikator.

Dalam bidang psikologi atau pendidikan terdapat beberapa variabel laten yang bersifat kategorik dan dibentuk oleh variabel indikator yang bersifat kontinu. Sebagai contoh pola asuh orang tua, yang mempunyai kategori otoriter, demokrasi, *permissive* dan *neglect*, diukur oleh variabel kontinu *demandingness* dan *responsiveness*. Terkadang terdapat variabel laten dimana kategori dari variabel tersebut lebih dipentingkan atau lebih memberikan arti dibandingkan nilai kontinunya jika nilai tersebut ada. Sebagai contoh adalah variabel laten tingkat mengatur diri sendiri yang dibentuk oleh variabel indikator tingkat ketaktergantungan, skor tanggung jawab dan tingkat ketenangan yang bersifat kontinu. Kategori dari variabel laten tingkat mengatur diri sendiri (rendah dan tinggi) lebih memberikan arti dibandingkan skor kontinu untuk variabel tersebut. Oleh karena itu, pembentukan kelas-kelas dari variabel laten perlu diperhatikan.

Model matematika yang bertujuan untuk menghubungkan sejumlah variabel indikator dengan suatu variabel laten yang dibentuknya disebut model variabel laten. Jika variabel indikator bersifat kontinu dan variabel laten yang dibentuknya bersifat kategorik maka model variabel latennya disebut *latent profile model*. Kelas-kelas dari variabel laten pada *latent profile model* disebut kelas laten.

Karena variabel laten merupakan variabel yang tidak dapat diukur secara langsung, tetapi diukur berdasarkan sejumlah variabel indikator, maka pembentukan kelas laten hanya dapat dilakukan berdasarkan nilai dari variabel indikatornya. Demikian juga ciri atau profil pada setiap kelas laten akan diwakili oleh ciri atau profil dari setiap variabel indikatornya (mean dan kovariansi dari variabel-variabel indikator).

Pada prinsipnya *latent profile model* akan membentuk kelas-kelas laten berdasarkan sejumlah variabel indikator kontinu sedemikian sehingga di dalam setiap kelas laten, variabel-variabel indikator akan saling bebas. Yang menjadi permasalahan dalam *latent profile model* adalah bagaimana menentukan banyak kelas laten optimal yang cukup untuk menjelaskan variasi dari variabel indikator dan bagaimana mengelompokkan individu ke dalam kelas laten yang terbentuk. Dalam tugas akhir ini pengelompokkan individu ke dalam kelas laten tidak dibahas.

1.2 Perumusan Masalah

Permasalahan dalam tugas akhir ini adalah bagaimana menentukan banyak kelas laten optimal yang cukup untuk menjelaskan variasi dari variabel indikator dalam *latent profile model*.

1.3 Tujuan Penulisan

Tujuan penulisan dari tugas akhir ini adalah menentukan banyak kelas laten optimal yang cukup untuk menjelaskan variasi dari variabel indikator dalam *latent profile model*.

1.4 Pembatasan Masalah

Pembatasan masalah dalam tulisan ini adalah:

1. Kelas laten yang terbentuk bersifat kategorik nominal (keordinalan kelas laten tidak diperhitungkan).
2. Pengelompokkan individu ke dalam kelas laten yang terbentuk tidak dibahas dalam tugas akhir ini.

1.5 Sistematika Penulisan

Sistematika penulisan pada tugas akhir ini adalah sebagai berikut:

Bab 1 Pendahuluan

Bab ini berisi latar belakang masalah, permasalahan, pembatasan masalah, tujuan penulisan, dan sistematika penulisan.

Bab 2 Landasan Teori

Bab ini berisi pembahasan mengenai konsep dasar yang akan digunakan dalam pembentukkan *latent profile model*, meliputi *mixture* distribusi, taksiran *joint* maksimum likelihood, algoritma EM (*Expectation-Maximization*), dan uji rasio likelihood.

Bab 3 *Latent Profile Model*

Bab ini berisi pembahasan mengenai *latent profile model*, penaksiran parameter, algoritma EM (*Expectation-Maximization*), uji kecocokan model dan menentukan banyak kelas optimal dari variabel laten yang dibentuk.

Bab 4 Contoh Aplikasi Pembentukan Kelas Laten pada Variabel "Tingkat Mengatur Diri Sendiri"

Bab ini berisi penerapkan *latent profile model* pada kasus "Membentuk kategori dari variabel laten "tingkat mengatur diri sendiri" berdasarkan variabel indikator “ tingkat ketaktergantungan”, “ skor tanggung jawab”, dan “tingkat ketenangan” pada mahasiswa baru Departemen Matematika FMIPA UI angkatan tahun 2010”.

Bab 5 Penutup

Bab ini berisi kesimpulan dan saran.

BAB 2
LANDASAN TEORI

Bab ini membahas beberapa pengertian dasar yang diperlukan pada pembahasan bab-bab berikutnya, yaitu mengenai *mixture* distribusi, taksiran *joint* maksimum likelihood, algoritma EM (*Expectation-Maximization*) dan uji rasio likelihood.

2.1 *Mixture* Distribusi

Misalkan terdapat variabel random $X_1, X_2, \ldots, X_C$ dengan pdf $f_{x1}, f_{x2}, \ldots, f_{xC}$. Sebuah variabel random Y disebut *mixture* dari variabel random $X_1, X_2, \ldots, X_C$ jika mempunyai pdf sebagai berikut

$$f_Y(y) = a_1 f_{X1}(y) + a_2 f_{X2}(y) + \ldots + a_C f_{XC}(y) \qquad (2.1)$$

Dimana $a_j > 0$ dan $\Sigma\, a_j = 1$; a_j disebut bobot dari "pencampuran" atau *mixing* probabilitas (Stuart A. Klugman).

2.2 Taksiran *Joint* Maksimum Likelihood

Misalkan $Y_1, Y_2, \ldots, Y_n$ adalah suatu sampel random berukuran n dari suatu distribusi dengan pdf $f(y;\boldsymbol{\theta})$ yang bergantung pada $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m) \in \boldsymbol{\Omega}$, $\boldsymbol{\Omega}$ adalah ruang parameter. Karena $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ merupakan suatu vektor maka salah satu metode untuk mencari taksiran untuk $\boldsymbol{\theta}$ adalah taksiran *joint* maksimum likelihood. Dalam melakukan penaksiran *joint* maksimum likelihood ada beberapa tahapan yang harus dilakukan.

Pertama, cari pdf bersama dari $Y_1, Y_2, \ldots, Y_n$ yaitu $f(y_1, y_2, \ldots, y_n; \boldsymbol{\theta})$. Karena $Y_1, Y_2, \ldots, Y_n$ merupakan sampel random maka pdf bersama dari $Y_1, Y_2, \ldots, Y_n$ dapat dinyatakan sebagai:

$$f(y_1, y_2, \ldots, y_n; \boldsymbol{\theta}) = f(y_1; \boldsymbol{\theta}).f(y_2; \boldsymbol{\theta}). \cdots .f(y_n; \boldsymbol{\theta}) \tag{2.2}$$

Kedua, cari fungsi likelihood yang merupakan pdf bersama dari $Y_1, Y_2, \ldots, Y_n$ yang dianggap sebagai fungsi dari $\boldsymbol{\theta}$, sehingga persamaan (2.2) dapat ditulis sebagai berikut, sebut $L(\boldsymbol{\theta})$.

$$\begin{aligned} L(\boldsymbol{\theta}) &= f(y_1, y_2, \ldots, y_n; \boldsymbol{\theta}) \\ &= f(y_1; \boldsymbol{\theta}).f(y_2; \boldsymbol{\theta}). \cdots .f(y_n; \boldsymbol{\theta}) \\ &= \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta}) \end{aligned} \tag{2.3}$$

Ketiga, cari taksiran dari $\boldsymbol{\theta}$. Dalam metode penaksiran *joint* maksimum likelihood, taksiran dari $\boldsymbol{\theta}$ diperoleh dengan menemukan nilai $\boldsymbol{\theta}$, sebut $\widehat{\boldsymbol{\theta}}$, yang memaksimumkan fungsi likelihood. Maka $\widehat{\boldsymbol{\theta}}$ disebut taksiran *joint* maksimum likelihood dari $\boldsymbol{\theta}$.

Karena nilai $\boldsymbol{\theta}$ yang memaksimumkan $L(\boldsymbol{\theta})$ sama dengan nilai $\boldsymbol{\theta}$ yang memaksimumkan $\ln L(\boldsymbol{\theta})$ sebut $l(\boldsymbol{\theta})$ (bukti diberikan di lampiran 1) maka seringkali taksiran *joint* maksimum likelihood dari $\boldsymbol{\theta}$ dicari dengan memaksimumkan $l(\boldsymbol{\theta})$, sehingga persamaan (2.3) menjadi:

$$\begin{aligned} l(\boldsymbol{\theta}) = \ln L(\boldsymbol{\theta}) &= ln\left(\prod_{i=1}^{n} f(y_i; \boldsymbol{\theta})\right) \\ &= \sum_{i=1}^{n} ln\ f(y_i; \boldsymbol{\theta}) \end{aligned} \tag{2.4}$$

Nilai $\boldsymbol{\theta}$ yang memaksimumkan $l(\boldsymbol{\theta})$ dapat diperoleh dengan mencari solusi simultan dari persamaan

$$\begin{aligned} S_j(\boldsymbol{\theta}) &= \frac{\partial}{\partial \theta_j} l(\boldsymbol{\theta}), \qquad untuk\ j = 1, \ldots, m \\ &= \frac{1}{L(\boldsymbol{\theta})} . \frac{\partial}{\partial \theta_j} L(\boldsymbol{\theta}) = 0 \end{aligned} \tag{2.5}$$

Adakalanya sistem persamaan (2.5) dapat diselesaikan secara analitik. Jika tidak, suatu prosedur numerik dapat digunakan.

2.3 Algoritma EM (*Expectation-Maximization*)

Algoritma EM merupakan suatu algoritma yang bersifat iteratif yang biasanya digunakan untuk mencari MLE dari parameter dalam model variabel laten. Misalkan X adalah suatu variabel laten dengan C kategori dan $Y_1, Y_2, \ldots, Y_n$ adalah variabel indikator yang mempunyai joint pdf $f(\boldsymbol{y};\theta)$. Misalkan $l(\boldsymbol{y}, \theta)$ adalah fungsi log likelihood dari $\boldsymbol{Y}$.

$$l(\boldsymbol{y}, \theta) = \log [f(\boldsymbol{y};\theta)] \tag{2.6}$$

Misalkan $f(\boldsymbol{y}, x, \theta)$ adalah pdf bersama dari $\boldsymbol{Y}$ dan X, dengan θ adalah parameter dalam model. Karena, seperti yang telah dinyatakan pada pemisalan awal, X adalah variabel laten, maka salah satu cara untuk mencari taksiran θ yang memaksimumkan fungsi likelihood dari $\boldsymbol{Y}$ adalah dengan menggunakan algoritma EM. Prinsip dari algoritma EM dapat dijelaskan menjadi 2 bagian sebagai berikut:

1. E-*Step*

   E-*step* dilakukan untuk mencari $E\left[log[f(x, y, \boldsymbol{\theta}_t)] \middle| y, \widehat{\boldsymbol{\theta}}_{t-1}\right]$ dimana:

   $\widehat{\boldsymbol{\theta}}_{t-1}$ adalah taksiran $\boldsymbol{\theta}$ pada iterasi ke-(t-1), t =1, 2,…

   $\boldsymbol{\theta}_t$ adalah nilai $\boldsymbol{\theta}$ pada iterasi ke-(t).

   $\boldsymbol{\theta}_0$ adalah suatu nilai taksiran awal yang diberikan.

2. M-*Step*

   Pada M-*step*, maksimumkan $E\left[log[f(x, y, \boldsymbol{\theta}_t)] \middle| y, \widehat{\boldsymbol{\theta}}_{t-1}\right]$ dengan cara menurunkannya terhadap θ_t untuk mendapatkan taksiran θ_t pada iterasi ke-(t), sebut $\hat{\theta}_t$.

Proses E-*step* dan M-*step* ini akan dilakukan terus secara iteratif sampai didapatkan suatu estimasi untuk θ yang konvergen atau $\left|\hat{\theta}_t - \hat{\theta}_{t-1}\right|$ cukup kecil. Iterasi algoritma EM seperti yang dijelaskan melalui E-*step* dan M-*step* diatas akan meningkatkan nilai $l(\boldsymbol{y}, \theta)$, pada setiap iterasinya (Welling).

2.4 Uji Rasio Likelihood

Uji rasio likelihood adalah suatu metode yang digunkan untuk menguji hipotesis H_0: $\boldsymbol{\theta} \epsilon \boldsymbol{\omega}$ terhadap hipotesis alternatif H_1: $\boldsymbol{\theta} \epsilon \boldsymbol{\Omega}$, dengan $\boldsymbol{\Omega}$ adalah ruang parameter keseluruhan dan $\boldsymbol{\omega}$ adalah ruang parameter dalam H_0, $\boldsymbol{\omega} \subset \boldsymbol{\Omega}$.

Misalkan terdapat variabel random $Y_1, Y_2, \ldots, Y_n$ yang memiliki pdf $f(y_i; \theta_1, \theta_2, \cdots, \theta_m)$ untuk i = 1, 2, …, n. Parameter-parameter dari populasi $\theta_1, \theta_2, \cdots, \theta_m$ dimisalkan berada pada ruang parameter $\boldsymbol{\Omega}$. Misalkan ω merupakan subset dari $\boldsymbol{\Omega}$ dan akan diuji hipotesis H_0: $(\theta_1, \theta_2, \cdots, \theta_m) \epsilon \boldsymbol{\omega}$ terhadap H_1: $(\theta_1, \theta_2, \cdots, \theta_m) \epsilon \boldsymbol{\Omega}$. Definisikan fungsi likelihood sebagai berikut:
Sebut pdf bersama dari $Y_1, Y_2, \ldots, Y_n$ pada saat $(\theta_1, \theta_2, \cdots, \theta_m) \epsilon \boldsymbol{\omega}$:

$$L(\boldsymbol{\omega}) = \prod_{i=1}^{n} f_i(y_i; \theta_1, \theta_2, \cdots, \theta_m)$$

dan sebut pdf bersama dari $Y_1, Y_2, \ldots, Y_n$ pada saat $(\theta_1, \theta_2, \cdots, \theta_m) \epsilon \boldsymbol{\Omega}$:

$$L(\boldsymbol{\Omega}) = \prod_{i=1}^{n} f_i(y_i; \theta_1, \theta_2, \cdots, \theta_m)$$

Pandang rasio dari kedua fungsi likelihood di atas adalah sebagai berikut:

$$\lambda^* = \frac{L(\boldsymbol{\omega})}{L(\boldsymbol{\Omega})}$$

Nilai λ^* tidak dapat digunakan sebagai statistik uji untuk menguji H_0 terhadap H_1 karena nilai $L(\boldsymbol{\omega})$ dan $L(\boldsymbol{\Omega})$ biasanya tidak diketahui.

Definisi 2.1 (Hog & Craig, 1995)

Misalkan $L(\hat{\boldsymbol{\omega}})$ adalah nilai dari fungsi likelihood dengan $\hat{\boldsymbol{\omega}}$ yang memaksimumkan $L(\boldsymbol{\omega})$ dan $L(\widehat{\boldsymbol{\Omega}})$ merupakan nilai dari fungsi likelihood dengan $\widehat{\boldsymbol{\Omega}}$ yang memaksimumkan $L(\boldsymbol{\Omega})$. Rasio dari $L(\hat{\boldsymbol{\omega}})$ terhadap $L(\widehat{\boldsymbol{\Omega}})$ disebut rasio likelihood dan dinotasikan sebagai berikut

$$\lambda = \frac{L(\hat{\boldsymbol{\omega}})}{L(\widehat{\boldsymbol{\Omega}})} \tag{2.7}$$

Nilai λ dapat dicari dan dapat digunakan sebagai statistik uji untuk menguji H_0 terhadap H_1. $L(\hat{\boldsymbol{\omega}})$ dan $L(\hat{\boldsymbol{\Omega}})$ bernilai positif sehingga λ juga akan bernilai positif, atau $\lambda \geq 0$. Kemudian karena $\hat{\boldsymbol{\omega}}$ adalah subset dari $\hat{\boldsymbol{\Omega}}$, maka $L(\hat{\boldsymbol{\omega}}) \leq L(\hat{\boldsymbol{\Omega}})$. Karena $L(\hat{\boldsymbol{\omega}}) \leq L(\hat{\boldsymbol{\Omega}})$, maka nilai $\lambda \leq 1$. Jadi, diperoleh $0 \leq \lambda = \frac{L(\hat{\boldsymbol{\omega}})}{L(\hat{\boldsymbol{\Omega}})} \leq 1$. Apabila $\lambda = 0$ maka $L(\hat{\boldsymbol{\omega}}) = 0$, H_0 ditolak. Jadi, H_0 akan ditolak apabila λ bernilai kecil (mendekati 0). Misalkan λ_0 adalah suatu bilangan positif sedemikian sehingga H_0: $(\theta_1, \theta_2, \cdots, \theta_m) \in \boldsymbol{\omega}$ akan ditolak apabila $\lambda(x_1, x_2, \ldots, x_n) = \lambda \leq \lambda_0$.

Misalkan α adalah tingkat signifikansi yang dipakai dalam pengujian.

$$\alpha = Pr[\lambda(x_1, x_2, \cdots, x_n) \leq \lambda_0; H_0].$$

Jika pdf dari statistik $\lambda = \lambda(X_1, X_2, \ldots, X_n)$ dapat diketahui untuk H_0 benar, λ_0 dapat ditentukan, sedemikian sehingga $\alpha = Pr[\lambda(x_1, x_2, \cdots, x_n) \leq \lambda_0; H_0]$. Namun, seringkali, di bawah H_0 benar, sulit untuk menentukan distribusi dari $\lambda = \lambda(X_1, X_2, \ldots, X_n)$. Oleh karena itu tidak mungkin untuk menemukan λ_0 sedemikian sehingga $\alpha = Pr[\lambda(x_1, x_2, \cdots, x_n) \leq \lambda_0; H_0]$.

Karena distribusi dari statistik $\lambda = \lambda(X_1, X_2, \ldots, X_n)$ sulit ditentukan, maka dicari statistik uji lain yaitu $L^2 = -2 \ln \lambda$ (log rasio likelihood). George G. Roussas dalam bukunya yang berjudul *A Course in Mathematical Statistics* menunjukkan bahwa $-2 \ln \lambda$ akan berdistribusi χ_r^2, dimana r = dimensi $\boldsymbol{\Omega}$ – dimensi $\boldsymbol{\omega}$.

BAB 3
PENENTUAN BANYAK KELAS LATEN OPTIMAL PADA *LATENT PROFILE MODEL*

Pada bab ini akan dibahas mengenai pembentukan kelas pada *latent profil model*, penaksiran parameter pada model, uji kecocokan model dan menentukan banyak kelas laten optimal.

3.1 Model

Latent profile model merupakan suatu model yang menghubungkan sejumlah variabel indikator yang bersifat kontinu dengan variabel laten kategorik yang dibentuknya. Kelas-kelas dari variabel laten pada *latent profile model* disebut kelas laten. Misalkan $Y_1, Y_2, \dots, Y_n$ adalah variabel indikator yang bersifat kontinu dan membentuk suatu variabel laten X yang bersifat kategorik dengan C kategori. Asumsikan bahwa setiap variabel indikator berdistribusi normal pada setiap kelas. Suatu individu akan memberikan jawaban (respon) untuk setiap variabel indikator. Misalkan y_{ik} adalah respon individu ke-k terhadap variabel indikator Y_i, $k = 1, 2, \dots, K$; $i = 1, 2, \dots, n$. Sebut $\boldsymbol{y_k} = (y_{1k}, y_{2k}, \dots, y_{nk})$ adalah respon individu ke-k untuk n variabel indikator $Y_1, Y_2, \dots, Y_n$. Sebut:

$P(x)$ = probabilitas suatu individu berada pada kelas laten ke-x, dimana $x = 1,2,\dots, C$

Misalkan $f(\boldsymbol{y})$ adalah pdf bersama dari n variabel indikator. Sebut $f(\boldsymbol{y_k})$ adalah nilai $f(\boldsymbol{y})$ untuk individu ke-k. Misalkan $f(\boldsymbol{y}|X = x)$ adalah pdf bersama dari n variabel indikator jika diketahui $X = x$ dan $f(\boldsymbol{y_k}|X = x)$ adalah nilai dari $f(\boldsymbol{y}|X = x)$ untuk individu ke-k. Tiap kelas pada variabel laten memiliki vektor mean dan vektor variansi yaitu

$\boldsymbol{\mu}_x$ = vektor mean dari variabel indikator pada kelas ke-*x*, contohnya:

$$\boldsymbol{\mu}_1 = \begin{bmatrix} \mu_{11} \\ \vdots \\ \mu_{n1} \end{bmatrix}, \qquad \boldsymbol{\mu}_2 = \begin{bmatrix} \mu_{12} \\ \vdots \\ \mu_{n2} \end{bmatrix}, \ \ldots, \qquad \boldsymbol{\mu}_C = \begin{bmatrix} \mu_{1C} \\ \vdots \\ \mu_{nC} \end{bmatrix}$$

μ_{ix} = mean dari variabel indikator ke-*i* pada kelas ke-*x*, $i = 1,\ldots,n$; $x = 1,\ldots,C$

dan

$\boldsymbol{\sigma}_x{}^2$ = vektor variansi dari variabel indikator pada kelas ke-*x*, contohnya:

$$\boldsymbol{\sigma}_1{}^2 = \begin{bmatrix} \sigma_{11}{}^2 \\ \vdots \\ \sigma_{n1}{}^2 \end{bmatrix}, \qquad \boldsymbol{\sigma}_2{}^2 = \begin{bmatrix} \sigma_{12}{}^2 \\ \vdots \\ \sigma_{n2}{}^2 \end{bmatrix}, \ \ldots, \qquad \boldsymbol{\sigma}_C{}^2 = \begin{bmatrix} \sigma_{1C}{}^2 \\ \vdots \\ \sigma_{nC}{}^2 \end{bmatrix}$$

$\sigma_{ix}{}^2$ = variansi dari variabel indikator ke-*i* pada kelas ke-*x*, $i = 1,\ldots,n$; $x = 1,\ldots,C$

Jika tiap kelas dapat diwakili oleh vektor mean $\boldsymbol{\mu}_x$ dan vektor variansi $\boldsymbol{\sigma}_x{}^2$ maka $f(\boldsymbol{y}|X = x)$ dapat ditulis dengan $f(\boldsymbol{y}|\boldsymbol{\mu}_x,\ \boldsymbol{\sigma}_x{}^2)$.

Jika $Y_1, Y_2, \ldots, Y_n$ merupakan variabel indikator yang saling bebas dan berdistribusi normal pada setiap kelas laten maka untuk tiap kelas $f(\boldsymbol{y}|\boldsymbol{\mu}_x,\ \boldsymbol{\sigma}_x{}^2)$ dapat dituliskan dengan

$$f(\boldsymbol{y}|\boldsymbol{\mu}_x,\ \boldsymbol{\sigma}_x{}^2) = \prod_{i=1}^{n} f_i(\boldsymbol{y}|\boldsymbol{\mu}_x,\ \boldsymbol{\sigma}_x{}^2)$$

$$= \left[\frac{1}{\prod_{i=1}^{n} \sigma_{ix}\sqrt{2\pi}}\right].exp\left[-\sum_{i=1}^{n} \frac{(y-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right]$$

Karena *X* adalah variabel laten yang bersifat kategorik yang dibentuk oleh variabel indikator kontinu $Y_1, Y_2, \ldots, Y_n$ maka pdf bersama dari ***y*** dapat dituliskan sebagai *mixture* distribusi dari ***y*** untuk setiap kelas dari variabel laten *X* atau dengan perkataan lain dapat dituliskan dengan

$$f(\boldsymbol{y}) = P(1)f(\boldsymbol{y}|\boldsymbol{\mu}_1, \boldsymbol{\sigma}_1{}^2) + \cdots + P(C)f(\boldsymbol{y}|\boldsymbol{\mu}_C, \boldsymbol{\sigma}_C{}^2)$$

$$= \sum_{x=1}^{C} P(x)f(\boldsymbol{y}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) \qquad (3.1)$$

dimana $P(x) > 0$ dan $\sum_{x=1}^{C} P(x) = 1$.

Persamaan (3.1) disebut *latent profile model* dimana $P(x)$, μ_{ix}, dan $\sigma_{ix}{}^2$ merupakan parameter dari model yang akan ditaksir. Dari persamaan (3.1) nilai $f(\boldsymbol{y})$ untuk individu ke-*k* dapat dituliskan sebagai:

$$f(\boldsymbol{y_k}) = \sum_{x=1}^{C} P(x)f(\boldsymbol{y}_k|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$$

3.2 Penaksiran Parameter dalam Model

Salah satu metode yang dapat digunakan untuk menaksir nilai parameter-parameter dalam *latent profile model* adalah metode *joint* maksimum likelihood. Misalkan terdapat *K* individu yang diamati. Karena antar variabel indikator dalam satu kelas saling bebas maka didapatkan fungsi likelihood untuk *latent profile model* adalah sebagai berikut:

$$L(\boldsymbol{y};\ \mathbf{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) = \prod_{k=1}^{K} f(\boldsymbol{y_k}; \boldsymbol{\theta})$$

$$= \prod_{k=1}^{K} [\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)] \quad (3.2)$$

dimana $\mathbf{P} = (P(1), \ldots, P(C))$, $\boldsymbol{\mu}_x = (\mu_{11}, \ldots, \mu_{n1}, \ldots, \mu_{1x}, \ldots, \mu_{nC})$, dan $\boldsymbol{\sigma}_x{}^2 = (\sigma_{11}{}^2, \ldots, \sigma_{n1}{}^2, \ldots, \sigma_{1C}{}^2, \ldots, \sigma_{nC}{}^2)$.

Prinsip dari metode ini adalah mencari taksiran dari $P(x)$, μ_{ix} dan $\sigma_{ix}{}^2$, $i = 1, \ldots, n$, $x = 1, \ldots, C$ yang secara bersama memaksimumkan fungsi likelihood. Mencari nilai taksiran parameter $P(x)$, μ_{ix} dan $\sigma_{ix}{}^2$ sebut $\hat{P}(x)$, $\hat{\mu}_{ix}$, dan $\hat{\sigma}_{ix}^2$ yang memaksimumkan bentuk logaritma dari fungsi likelihood $lnL(\boldsymbol{y}; \boldsymbol{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$ akan memberikan hasil yang sama dengan mencari nilai $\hat{P}(x)$, $\hat{\mu}_{ix}$, dan $\hat{\sigma}_{ix}^2$ yang memaksimumkan fungsi likelihood $L(\boldsymbol{y}; \boldsymbol{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$. Maka baik $L(\boldsymbol{y}; \boldsymbol{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$ atau $lnL(\boldsymbol{y}; \boldsymbol{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$ dapat digunakan untuk mencari nilai $\hat{P}(x)$, $\hat{\mu}_{ix}$, dan $\hat{\sigma}_{ix}^2$.

$$lnL(\boldsymbol{y};\ \boldsymbol{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) = ln\,(\prod_{k=1}^{K} f(\boldsymbol{y_k}; \boldsymbol{\theta}))$$

$$= ln(\prod_{k=1}^{K} \{\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)\})$$

$$= \sum_{k=1}^{K} ln[\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)] \quad (3.3)$$

Karena terdapat syarat bahwa

$$\sum_{x=1}^{C} P(x) = 1 \quad (3.4)$$

maka akan dicari nilai $\hat{P}(x)$, $\hat{\mu}_{ix}$, dan $\hat{\sigma}_{ix}^2$ yang memaksimumkan $ln\mathrm{L}(\mathbf{y}; \boldsymbol{P}, \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$ di bawah pembatasan (3.4). Dengan perkataan lain akan dicari nilai $\hat{P}(x)$, $\hat{\mu}_{ix}$, dan $\hat{\sigma}_{ix}^2$ yang memaksimumkan

$$Q = \ln L - \gamma\,[\sum_{x=1}^{C} (P(x) - 1)]$$

$$= \sum_{k=1}^{K} ln[\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)] - \gamma[\sum_{x=1}^{C} (P(x) - 1)] \quad (3.5)$$

dimana γ adalah pengali Lagrange. Karena diasumsikan setiap variabel indikator berdistribusi normal maka

$$f(\boldsymbol{y_k}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) = \left[\frac{1}{\prod_{i=1}^{n} \sigma_{ix}\sqrt{2\pi}}\right].exp\left[-\sum_{i=1}^{n} \frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right]$$

Taksiran maksimum likelihood bersama dari parameter $P(x)$, μ_{ix} dan $\sigma_{ix}{}^2$ diperoleh dengan menurunkan secara parsial fungsi Q pada persamaan (3.5) terhadap parameter $P(x)$, μ_{ix} dan $\sigma_{ix}{}^2$ kemudian disamakan dengan nol. Turunan parsial pertama dari Q terhadap $P(x)$, μ_{ix} dan $\sigma_{ix}{}^2$ adalah sebagai berikut:

$$\frac{\partial\, Q}{\partial\, P(x)} = \frac{\partial}{\partial\, P(x)}(\ln L - \gamma\,[\textstyle\sum_{x=1}^{C} P(x) - 1])$$

$$= \frac{\partial}{\partial\, P(x)} \sum_{k=1}^{K} ln[\textstyle\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)] - \gamma[\sum_{x=1}^{C} P(x) - 1]$$

$$= \sum_{k=1}^{K}\left(\frac{f(y_k|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)P(x)}\right) - \gamma = 0$$

$$= \sum_{k=1}^{K}\left(\frac{f(y_k|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(y_k)}\right) - \gamma = 0$$

$$\gamma = \sum_{k=1}^{K}\left(\frac{f(y_k|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(y_k)}\right) \qquad (3.6)$$

Kemudian kalikan kedua ruas pada persamaan (3.6) dengan $P(x)$ maka didapat

$$\gamma P(x) = \sum_{k=1}^{K}\left(\frac{f(y_k|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)P(x)}{f(\boldsymbol{y_k})}\right)$$

$$= \sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}, x; \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right), \qquad x = 1, \ldots, C \qquad (3.7)$$

dimana $f(\boldsymbol{y_k},\ x; \boldsymbol{\mu}_x,\ \boldsymbol{\sigma}_x{}^2)$ adalah pdf bersama dari n variabel indikator dan kelas laten x. Kemudian jumlahkan persamaan (3.7) untuk semua kelas laten, sehingga persamaan (3.7) menjadi

$$\gamma P(1) + \cdots + \gamma P(C) = \frac{f(\boldsymbol{y_1}, 1; \boldsymbol{\mu}_1, \boldsymbol{\sigma}_1{}^2) + \cdots + f(\boldsymbol{y}_1, C; \boldsymbol{\mu}_C, \boldsymbol{\sigma}_C{}^2)}{f(\boldsymbol{y_1})} + \cdots + \frac{f(\boldsymbol{y_K}, 1; \boldsymbol{\mu}_1, \boldsymbol{\sigma}_1{}^2) + \cdots + f(\boldsymbol{y_K}, C; \boldsymbol{\mu}_C, \boldsymbol{\sigma}_C{}^2)}{f(\boldsymbol{y_K})}$$

$$\gamma\big(P(1) + \cdots + P(C)\big) = \frac{\sum_{x=1}^{C} f(\boldsymbol{y_1}, x; \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_1})} + \cdots + \frac{\sum_{x=1}^{C} f(\boldsymbol{y_K}, x; \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_K})}$$

$$\gamma \sum_{x=1}^{C} P(x) = \sum_{k=1}^{K}\left[\frac{\sum_{x=1}^{C} f(\boldsymbol{y_k}, x; \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right]$$

$$\gamma.1 = \sum_{k=1}^{K}\left[\frac{f(\boldsymbol{y_k})}{f(\boldsymbol{y_k})}\right]$$

$$= \sum_{k=1}^{K} 1$$
$$= K \qquad (3.8)$$

Substitusikan persamaan (3.8) ke persamaan (3.7) menjadi

$$KP(x) = \sum_{k=1}^{K} \left(\frac{f(\boldsymbol{y_k}, x; \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})} \right)$$

sehingga didapat taksiran untuk $P(x)$ yaitu

$$\hat{P}(x) = \frac{\left[\sum_{k=1}^{K} \left(\frac{f(\boldsymbol{y_k}, x; \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{\mathrm{f}(\boldsymbol{y_k})} \right) \right]}{K} \qquad (3.9)$$

Karena $\hat{P}(x)$ merupakan fungsi dari $\boldsymbol{\mu}_x$ dan $\boldsymbol{\sigma}_x{}^2$ maka penyelesaian secara langsung sulit untuk dilakukan. Oleh karena itu dibutuhkan suatu algoritma untuk memperoleh nilai taksiran dari $P(x)$ yang akan dibahas pada sub bab selanjutnya.

Taksiran untuk μ_{ix} yaitu

$$\frac{\partial\, \mathrm{Q}}{\partial\ \mu_{ix}} = \frac{\partial}{\partial\ \mu_{ix}} (\ln L - \gamma\ [\sum_{x=1}^{C} P(x) - 1])$$
$$= \frac{\partial}{\partial\ \mu_{ix}} \sum_{k=1}^{K} ln[\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)] - \gamma[\sum_{x=1}^{C} P(x) - 1]$$
$$= \frac{\partial}{\partial\ \mu_{ix}} \sum_{k=1}^{K} ln[\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)] - 0$$
$$= \frac{\partial}{\partial\ \mu_{ix}} \sum_{k=1}^{K} ln[\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)]$$
$$= \sum_{k=1}^{K} \frac{P(x)[\partial f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) / \partial \mu_{ix}]}{\sum_{x=1}^{C} P(x) f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}$$
$$= \sum_{k=1}^{K} \frac{P(x)[\partial f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) / \partial \mu_{ix}]}{f(\boldsymbol{y_k})}$$
$$= P(x) \sum_{k=1}^{K} \left(\frac{1}{f(\boldsymbol{y_k})} \right) . \left(\frac{\partial f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{\partial \mu_{ix}} \right) = 0$$

Asumsikan bahwa $P(x) \neq 0$ maka selesaikan persamaan

$$\sum_{k=1}^{K} \left(\frac{1}{f(\boldsymbol{y_k})} \right) . \left(\frac{\partial f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{\partial \mu_{ix}} \right) = 0 \qquad (3.10)$$

dimana

$$\frac{\partial f(\boldsymbol{y_k} | \boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)}{\partial \mu_{ix}} = \frac{\partial}{\partial \mu_{ix}} \left(\frac{exp\left[-\sum_{i=1}^{n} \frac{(y_{ki} - \mu_{ix})^2}{2\sigma_{ix}{}^2} \right]}{[\prod_{i=1}^{n} \sigma_{ix} \sqrt{2\pi}]} \right)$$

$$= \frac{-2(-1)(y_{ki}-\mu_{ix})}{2\sigma_{ix}^2}\left(\frac{exp\left[-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right]}{\left[\prod_{i=1}^{n}\sigma_{ix}\sqrt{2\pi}\right]}\right)$$

$$= \frac{(y_{ki}-\mu_{ix})}{\sigma_{ix}^2}\, f(\boldsymbol{y_k}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2) \tag{3.11}$$

Substitusikan persamaan (3.11) ke dalam persamaan (3.10) sehingga didapat

$$\sum_{k=1}^{K}\left(\frac{1}{f(\boldsymbol{y_k})}\right).\left(\frac{(y_{ki}-\mu_{ix})}{\sigma_{ix}^2}\, f(\boldsymbol{y_k}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)\right) = 0 \tag{3.12}$$

Asumsikan bahwa σ_{ix}^2 berhingga sehingga persamaan (3.12) menjadi

$$\sum_{k=1}^{K}\left(\frac{(y_{ki}-\mu_{ix})f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right) = 0 \tag{3.13}$$

Selesaikan persamaan (3.13) untuk μ_{ix} sehingga didapat taksiran untuk μ_{ix} yaitu

$$\sum_{k=1}^{K}\left(\left[\frac{y_{ki}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right]-\left[\frac{\mu_{ix}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right]\right) = 0$$

$$\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2).y_{ki}}{f(\boldsymbol{y_k})}\right)-\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2).\mu_{ix}}{f(\boldsymbol{y_k})}\right) = 0$$

$$\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2).y_{ki}}{f(\boldsymbol{y_k})}\right) = \mu_{ix}\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right)$$

$$\hat{\mu}_{ix} = \frac{\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2).y_{ki}}{f(\boldsymbol{y_k})}\right)}{\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right)} \tag{3.14}$$

Karena $\hat{\mu}_{ix}$ juga merupakan fungsi dari $\boldsymbol{\mu}_x$ dan $\boldsymbol{\sigma}_x{}^2$ maka penyelesaian secara langsung sulit untuk dilakukan. Oleh karena itu dibutuhkan suatu algoritma untuk memperoleh nilai taksiran dari μ_{ix} yang akan dibahas pada subbab selanjutnya.

Selanjutnya taksiran untuk σ_{ix}^2 yaitu

$$\frac{\partial\, \mathrm{Q}}{\partial\sigma_{ix}^2} = \frac{\partial}{\partial\sigma_{ix}^2}\left(\ln L - \gamma\left[\sum_{x=1}^{C}P(x)-1\right]\right)$$

$$= \frac{\partial}{\partial\sigma_{ix}^2}\sum_{k=1}^{K} ln\left[\sum_{x=1}^{C}P(x)f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right]-\gamma\left[\sum_{x=1}^{C}P(x)-1\right]$$

$$= \frac{\partial}{\partial\sigma_{ix}^2}\sum_{k=1}^{K} ln\left[\sum_{x=1}^{C}P(x)f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right]-0$$

$$= \frac{\partial}{\partial\sigma_{ix}^2}\sum_{k=1}^{K} ln\left[\sum_{x=1}^{C}P(x)f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right]$$

$$= \sum_{k=1}^{K}\frac{P(x)\left[\partial f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)/\partial\sigma_{ix}^2\right]}{\sum_{x=1}^{C}P(x)f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}$$

$$= \sum_{k=1}^{K} \frac{P(x)[\partial f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)/\partial\sigma_{ix}^2]}{f(\boldsymbol{y_k})}$$

$$= P(x)\sum_{k=1}^{K}\left(\frac{1}{f(\boldsymbol{y_k})}\right).\left(\frac{\partial f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{\partial\sigma_{ix}^2}\right) = 0$$

Asumsikan bahwa $P(x) \neq 0$ maka selesaikan persamaan

$$\sum_{k=1}^{K}\left(\frac{1}{f(\boldsymbol{y_k})}\right).\left(\frac{\partial f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{\partial\sigma_{ix}^2}\right) = 0 \qquad (3.15)$$

dimana

$$\frac{\partial f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{\partial\sigma_{ix}^2} = \frac{\partial}{\partial\sigma_{ix}^2}\left(\frac{exp\left[-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right]}{\left[\prod_{i=1}^{n}\sigma_{ix}\sqrt{2\pi}\right]}\right)$$

$$= \frac{\partial}{\partial\sigma_{ix}^2}\left[\left(\frac{1}{\left[\prod_{i=1}^{n}\sigma_{ix}\sqrt{2\pi}\right]}\right).exp\left(-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right] \qquad (3.16)$$

Misalkan

$$u = \left(\frac{1}{\left[\prod_{i=1}^{n}\sigma_{ix}\sqrt{2\pi}\right]}\right) \qquad \text{dan} \qquad v = exp\left(-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)$$

turunan dari u dan v adalah sebagai berikut:

$$u' = \frac{\partial u}{\partial\sigma_{ix}{}^2} = \frac{\partial}{\partial\sigma_{ix}{}^2}\left(\frac{1}{\left[\prod_{i=1}^{n}\sigma_{ix}\sqrt{2\pi}\right]}\right)$$

$$= \frac{\partial}{\partial\sigma_{ix}{}^2}\left(\frac{(\sigma_{ix}{}^2)^{-1/2}}{\sqrt{2\pi}\left[\prod_{i=1}^{n-1}\sigma_{ix}\sqrt{2\pi}\right]}\right)$$

$$= \frac{-1/2(\sigma_{ix}{}^2)^{-3/2}}{\sqrt{2\pi}\left[\prod_{i=1}^{n-1}\sigma_{ix}\sqrt{2\pi}\right]}$$

$$= \frac{-1}{2\sigma_{ix}{}^2\sigma_{ix}\sqrt{2\pi}\left[\prod_{i=1}^{n-1}\sigma_{ix}\sqrt{2\pi}\right]}$$

$$= \frac{-1}{2\sigma_{ix}{}^2\left[\prod_{i=1}^{n}\sigma_{ix}\sqrt{2\pi}\right]}$$

dan

$$v' = \frac{\partial v}{\partial\sigma_{ix}{}^2} = \frac{\partial}{\partial\sigma_{ix}{}^2}\left[exp\left(-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right]$$

$$= \frac{2(y_{ki}-\mu_{ix})^2}{(2\sigma_{ix}{}^2)^2}\left[exp\left(-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right]$$

$$= \frac{(y_{ki}-\mu_{ix})^2}{2(\sigma_{ix}{}^2)^2}\left[exp\left(-\sum_{i=1}^{n}\frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right]$$

Maka persamaan (3.16) menjadi

$$\frac{\partial f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{\partial\sigma_{ix}^2} = \frac{\partial}{\partial\sigma_{ix}^2}\left[\left(\frac{1}{[\prod_{i=1}^n \sigma_{ix}\sqrt{2\pi}]}\right).exp\left(-\sum_{i=1}^n \frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right] = u'v + v'u$$

$$= \left[\frac{-1}{2\sigma_{ix}{}^2[\prod_{i=1}^n \sigma_{ix}\sqrt{2\pi}]}.exp\left(-\sum_{i=1}^n \frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right] +$$

$$\left[\frac{(y_{ki}-\mu_{ix})^2}{2(\sigma_{ix}{}^2)^2}\left[exp\left(-\sum_{i=1}^n \frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)\right].\frac{1}{[\prod_{i=1}^n \sigma_{ix}\sqrt{2\pi}]}\right]$$

$$= \left[\frac{-1}{2\sigma_{ix}{}^2}.\frac{exp\left(-\sum_{i=1}^n \frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)}{[\prod_{i=1}^n \sigma_{ix}\sqrt{2\pi}]}\right] +$$

$$\left[\frac{(y_{ki}-\mu_{ix})^2}{2(\sigma_{ix}{}^2)^2}.\frac{exp\left(-\sum_{i=1}^n \frac{(y_{ki}-\mu_{ix})^2}{2\sigma_{ix}{}^2}\right)}{[\prod_{i=1}^n \sigma_{ix}\sqrt{2\pi}]}\right]$$

$$= \left[\frac{-1}{2\sigma_{ix}{}^2}.f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right] + \left[\frac{(y_{ki}-\mu_{ix})^2}{2(\sigma_{ix}{}^2)^2}.f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right]$$

$$= f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\left[\frac{(y_{ki}-\mu_{ix})^2}{2(\sigma_{ix}{}^2)^2} - \frac{1}{2\sigma_{ix}{}^2}\right]$$

$$= \frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{2\sigma_{ix}{}^2}\left[\frac{(y_{ki}-\mu_{ix})^2}{\sigma_{ix}{}^2} - 1\right] \qquad (3.17)$$

Substitusikan persamaan (3.17) ke dalam persamaan (3.15) sehingga didapat

$$\sum_{k=1}^K \left(\frac{1}{f(\boldsymbol{y_k})}\right).\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{2\sigma_{ix}{}^2}\left[\frac{(y_{ki}-\mu_{ix})^2}{\sigma_{ix}{}^2} - 1\right]\right) = 0 \qquad (3.18)$$

Asumsikan bahwa σ_{ix}^2 berhingga sehingga persamaan (3.18) menjadi

$$\sum_{k=1}^K \left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right).\left(\frac{(y_{ki}-\mu_{ix})^2}{\sigma_{ix}{}^2} - 1\right) = 0 \qquad (3.19)$$

Selesaikan persamaan (3.19) untuk $\sigma_{ix}{}^2$ sehingga didapat taksiran untuk $\sigma_{ix}{}^2$yaitu

$$\sum_{k=1}^K \left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right).\left(\frac{(y_{ki}-\mu_{ix})^2}{\sigma_{ix}{}^2} - 1\right) = 0$$

$$\sum_{k=1}^K \left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right).\left(\frac{(y_{ki}-\mu_{ix})^2}{\sigma_{ix}{}^2} - \frac{\sigma_{ix}{}^2}{\sigma_{ix}{}^2}\right) = 0 \qquad (3.20)$$

Asumsikan bahwa σ_{ix}^2 berhingga sehingga persamaan (3.20) menjadi

$$\sum_{k=1}^K \left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right).((y_{ki} - \mu_{ix})^2 - \sigma_{ix}{}^2) = 0$$

$$\sum_{k=1}^{K}\left(\left[\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)(y_{ki}-\mu_{ix})^2}{f(\boldsymbol{y_k})}\right]-\left[\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\sigma_{ix}{}^2}{f(\boldsymbol{y_k})}\right]\right)=0$$

$$\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)(y_{ki}-\mu_{ix})^2}{f(\boldsymbol{y_k})}\right)-\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\sigma_{ix}{}^2}{f(\boldsymbol{y_k})}\right)=0$$

$$\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)(y_{ki}-\mu_{ix})^2}{f(\boldsymbol{y_k})}\right)=\sigma_{ix}{}^2\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right)$$

$$\hat{\sigma}_{ix}{}^2=\frac{\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2).(y_{ki}-\mu_{ix})^2}{f(\boldsymbol{y_k})}\right)}{\sum_{k=1}^{K}\left(\frac{f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)}{f(\boldsymbol{y_k})}\right)} \qquad (3.21)$$

Karena $\hat{\sigma}_{ix}{}^2$ juga merupakan fungsi dari $\boldsymbol{\mu}_x$ dan $\boldsymbol{\sigma}_x{}^2$ maka penyelesaian secara langsung sulit untuk dilakukan. Oleh karena itu dibutuhkan suatu algoritma untuk memperoleh nilai taksiran dari $\sigma_{ix}{}^2$ yang akan dibahas pada sub bab selanjutnya. Untuk menyelesaikan taksiran maksimum likelihood dari $P(x)$, μ_{ix}, dan $\sigma_{ix}{}^2$ akan digunakan algoritma EM (*Expectation-Maximization*).

3.3 Algoritma EM (*Expectation-Maximization*)

Seperti telah dijelaskan dalam bab sebelumnya, algoritma EM adalah suatu proses iteratif untuk menghitung taksiran maksimum likelihood yang dilakukan dengan 2 tahap, yaitu tahapan E-*step* dan M-*step*.

1. E-*step* (langkah ekspektasi)

Seperti telah dijelaskan dalam Bab 2, dalam E-*step* akan dicari $E\left[ln[f(x,y,\boldsymbol{\theta}_t)]\middle|y,\widehat{\boldsymbol{\theta}}_{t-1}\right]=\sum_x ln[f(x,y,\boldsymbol{\theta}_t)]f\left(x\middle|y,\widehat{\boldsymbol{\theta}}_{t-1}\right)$. Dalam *latent profile model*, tahapan E-*step* dilakukan untuk mencari ekspektasi dari $ln[\prod_{k=1}^{K}P(x)f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)]$ untuk setiap kelas-*x* dari variabel laten *X*.

$$E\left[ln[f(x,y,\boldsymbol{\theta}_t)]\middle|y,\widehat{\boldsymbol{\theta}}_{t-1}\right]=$$

$$E\left[ln\left[\prod_{k=1}^{K}\left(P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right)\right]\middle|y_k,\hat{P}(x)^{(t-1)},\hat{\mu}_{ix}{}^{(t-1)},\hat{\sigma}_{ix}{}^{2(t-1)}\right]$$

dimana:

$$f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)=\left[\frac{1}{\prod_{i=1}^{n}\sigma_{ix}{}^{(t)}\sqrt{2\pi}}\right].exp\left[-\sum_{i=1}^{n}\frac{\left(y_{ki}-\mu_{ix}{}^{(t)}\right)^2}{2\sigma_{ix}{}^{2(t)}}\right]$$

$P(x)^{(t)}$ adalah nilai $P(x)$ pada iterasi ke-t

$\mu_{ix}{}^{(t)}$ adalah nilai μ_{ix} pada iterasi ke-t

$\sigma_{ix}{}^{2(t)}$ adalah nilai $\sigma_{ix}{}^2$ pada iterasi ke-t

$$E\ \left[ln\left[\prod_{k=1}^{K}\left(P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right)\right]\middle|y_k,\hat{P}(x)^{(t-1)},\hat{\mu}_{ix}{}^{(t-1)},\hat{\sigma}_{ix}{}^{2(t-1)}\right]$$
$$=\ \sum_{x=1}^{C}\left[ln\left[\prod_{k=1}^{K}\left(P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right)\right]\right.$$
$$\left..Pr\left(x\middle|\boldsymbol{y_k},\hat{P}(x)^{(t-1)},\hat{\mu}_{ix}{}^{(t-1)},\hat{\sigma}_{ix}{}^{2(t-1)}\right)\right] \qquad (3.22)$$

dimana

$$ln\left[\prod_{k=1}^{K}\left(P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right)\right]=\ \sum_{k=1}^{K}ln\left[P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right]$$
$$=\sum_{k=1}^{K}\left[lnP(x)^{(t)}+lnf(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right]$$
$$=\ \sum_{k=1}^{K}lnP(x)^{(t)}+\sum_{k=1}^{K}ln\left[\frac{exp\left[-\sum_{i=1}^{n}\frac{\left(y_{ki}-\mu_{ix}{}^{(t)}\right)^2}{2\sigma_{ix}{}^{2(t)}}\right]}{\prod_{i=1}^{n}\sigma_{ix}{}^{(t)}\sqrt{2\pi}}\right]$$
$$=\ \sum_{k=1}^{K}lnP(x)^{(t)}+\ \sum_{k=1}^{K}\left[-\left(\sum_{i=1}^{n}\frac{\left(y_{ki}-\mu_{ix}{}^{(t)}\right)^2}{2\sigma_{ix}{}^{2(t)}}\right)-\left(\frac{n}{2}ln(2\pi)\right)-\right.$$
$$\left.\left(\sum_{i=1}^{n}ln\sigma_{ix}{}^{(t)}\right)\right] \qquad (3.23)$$

Sebut $Pr\left(x\middle|\boldsymbol{y_k},\hat{P}(x)^{(t-1)},\hat{\mu}_{ix}{}^{(t-1)},\hat{\sigma}_{ix}{}^{2(t-1)}\right)=\ T_{x,k}^{(t-1)}$, x = 1,…, C ; k = 1, …, K maka berdasarkan Teorema Bayes

$$Pr\left(x\middle|\boldsymbol{y_k},\hat{P}(x)^{(t-1)},\hat{\mu}_{ix}{}^{(t-1)},\hat{\sigma}_{ix}{}^{2(t-1)}\right)=\ \frac{\Pr(X=x).f\left(\boldsymbol{y_k}\middle|\boldsymbol{\mu_x},\boldsymbol{\sigma_x}{}^2\right)}{\sum_{x=1}^{C}\Pr(X=x).f\left(\boldsymbol{y_k}\middle|\boldsymbol{\mu_x},\boldsymbol{\sigma_x}{}^2\right)}$$
$$T_{x,k}^{(t-1)}=\ \frac{\hat{P}(x)^{(t-1)}.f\left(\boldsymbol{y_k}\middle|\hat{\boldsymbol{\mu}}_{x,}{}^{(t-1)}\hat{\boldsymbol{\sigma}}_x{}^{2(t-1)}\right)}{\sum_{x=1}^{C}\hat{P}(x)^{(t-1)}.f\left(y_k\middle|\hat{\boldsymbol{\mu}}_{x,}{}^{(t-1)}\hat{\boldsymbol{\sigma}}_x{}^{2(t-1)}\right)} \qquad (3.24)$$

Substitusikan persamaan (3.23) dan persamaan (3.24), ke dalam persamaan (3.22) menjadi

$$E\ \left[ln\left[\prod_{k=1}^{K}\left(P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,\boldsymbol{\sigma}_x{}^2)\right)\right]\middle|y_k,\hat{P}(x)^{(t-1)},\hat{\mu}_{ix}{}^{(t-1)},\hat{\sigma}_{ix}{}^{2(t-1)}\right]$$
$$=\ \sum_{x=1}^{C}\left[\left(\sum_{k=1}^{K}lnP(x)^{(t)}+\ \sum_{k=1}^{K}\left[-\left(\sum_{i=1}^{n}\frac{\left(y_{ki}-\mu_{ix}{}^{(t)}\right)^2}{2\sigma_{ix}{}^{2(t)}}\right)-\right.\right.\right.$$
$$\left.\left.\left.\left(\frac{n}{2}ln(2\pi)\right)-\left(\sum_{i=1}^{n}ln\sigma_{ix}{}^{(t)}\right)\right]\right).\left(T_{x,k}^{(t-1)}\right)\right] \qquad (3.25)$$

t = 1,2,…

2. M-*step* (langkah maksimisasi)

Setelah melakukan E-*step*, langkah selanjutnya adalah melakukan M-*step*, dimana pada proses ini akan dicari nilai taksiran untuk $P(x)^{(t)}$, ${\mu_{ix}}^{(t)}$, dan ${\sigma_{ix}}^{2(t)}$ yang memaksimumkan

$$E\ \left[ln\left[\prod_{k=1}^{K}\left(P(x)^{(t)}f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,{\boldsymbol{\sigma}_x}^2)\right)\right]\Big|y_k,\hat{P}(x)^{(t-1)},{\hat{\mu}_{ix}}^{(t-1)},{\hat{\sigma}_{ix}}^{2(t-1)}\right]$$

$$= \sum_{x=1}^{C}\left[\left(\sum_{k=1}^{K}lnP(x)^{(t)} + \sum_{k=1}^{K}\left[-\left(\sum_{i=1}^{n}\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2}{2{\sigma_{ix}}^{2(t)}}\right)-\right.\right.\right.$$

$$\left.\left.\left.\left(\frac{n}{2}ln(2\pi)\right)-\left(\sum_{i=1}^{n}ln{\sigma_{ix}}^{(t)}\right)\right]\right).\left(T_{x,k}^{(t-1)}\right)\right]$$

yang didapat pada E-*step*.

Untuk mencari nilai taksiran dari $P(x)^{(t)}$, ${\mu_{ix}}^{(t)}$, dan ${\sigma_{ix}}^{2(t)}$ yang memaksimumkan persamaan (3.25) diperoleh dengan menurunkan secara parsial persamaan (3.25) terhadap parameter $P(x)^{(t)}$, ${\mu_{ix}}^{(t)}$, dan ${\sigma_{ix}}^{2(t)}$ kemudian disamakan dengan nol. Karena terdapat syarat bahwa $\sum_{x=1}^{C}P(x) = 1$ maka akan dicari nilai $P(x)^{(t)}$, ${\mu_{ix}}^{(t)}$, dan ${\sigma_{ix}}^{2(t)}$ yang memaksimumkan

$$I = E\ \left[ln\left[\prod_{k=1}^{K}\left(P(x)f(\boldsymbol{y_k}|\boldsymbol{\mu}_x,{\boldsymbol{\sigma}_x}^2)\right)\right]\Big|y_k,\hat{P}(x)^{(t-1)},{\hat{\mu}_{ix}}^{(t-1)},{\hat{\sigma}_{ix}}^{2(t-1)}\right]$$

$$-\gamma\left[\sum_{x=1}^{C}P(x)-1\right]$$

$$= \sum_{x=1}^{C}\left[\left(\sum_{k=1}^{K}lnP(x)^{(t)} + \sum_{k=1}^{K}\left[-\left(\sum_{i=1}^{n}\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2}{2{\sigma_{ix}}^{2(t)}}\right)-\right.\right.\right.$$

$$\left.\left.\left.\left(\frac{n}{2}ln(2\pi)\right)-\left(\sum_{i=1}^{n}ln{\sigma_{ix}}^{(t)}\right)\right]\right).\left(T_{x,k}^{(t-1)}\right)\right]-\gamma\left[\sum_{x=1}^{C}P(x)-1\right]$$

Taksiran dari $P(x)^{(t)}$ didapat dengan menurunkan *I* terhadap $P(x)^{(t)}$ yaitu

$$\frac{\partial I}{\partial P(x)^{(t)}} = \frac{\partial}{\partial P(x)^{(t)}}\left(\sum_{x=1}^{C}\sum_{k=1}^{K}\left(T_{x,k}^{(t-1)}\right)lnP(x)^{(t)}-\gamma\left[\sum_{x=1}^{C}P(x)-1\right]\right)$$

$$\frac{\partial I}{\partial P(1)^{(t)}} = \frac{\partial}{\partial P(1)^{(t)}}\left(\sum_{x=1}^{C}\sum_{k=1}^{K}\left(T_{x,k}^{(t-1)}\right)lnP(x)^{(t)}-\gamma\left[\sum_{x=1}^{C}P(x)-1\right]\right)$$

$$= \frac{\sum_{k=1}^{K}\left(T_{1,k}^{(t-1)}\right)}{P(1)^{(t)}}-\gamma = 0$$

$$P(1)^{(t)} = \frac{\sum_{k=1}^{K}\left(T_{1,k}^{(t-1)}\right)}{\gamma}$$

$$\vdots$$

$$P(C)^{(t)} = \frac{\sum_{k=1}^{K}\left(T_{C,k}^{(t-1)}\right)}{\gamma}$$

Perhatikan persamaan berikut

$$\sum_{x=1}^{C} P(x) = 1$$

$$P(1) + \cdots + P(C) = 1$$

$$\frac{\sum_{k=1}^{K}\left(T_{1,k}^{(t-1)}\right)}{\gamma} + \cdots + \frac{\sum_{k=1}^{K}\left(T_{C,k}^{(t-1)}\right)}{\gamma} = 1$$

$$\frac{\sum_{k=1}^{K}\left(T_{1,k}^{(t-1)} + \cdots + T_{C,k}^{(t-1)}\right)}{\gamma} = 1$$

$$\frac{\sum_{k=1}^{K} 1}{\gamma} = 1$$

$$\frac{K}{\gamma} = 1$$

$$\gamma = K$$

Sehingga didapat taksiran untuk $P(x)^{(t)}$ yaitu

$$\hat{P}(x)^{(t)} = \frac{\sum_{k=1}^{K}\left(T_{x,k}^{(t-1)}\right)}{K} \quad (3.26)$$

Taksiran untuk ${\mu_{ix}}^{(t)}$ yaitu

$$\frac{\partial I}{\partial {\mu_{ix}}^{(t)}} = \frac{\partial}{\partial {\mu_{ix}}^{(t)}}\left(\sum_{x=1}^{C}\sum_{k=1}^{K}\left[-\sum_{i=1}^{n}\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2}{2{\sigma_{ix}}^{2(t)}}\right]\left[T_{x,k}^{(t-1)}\right]\right) = 0$$

$$= \sum_{k=1}^{K}\left[-\frac{2(-1)\left(y_{ki}-{\mu_{ix}}^{(t)}\right)}{2{\sigma_{ix}}^{2(t)}}\right]\left[T_{x,k}^{(t-1)}\right] = 0$$

$$= \sum_{k=1}^{K}\left[\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)}{{\sigma_{ix}}^{2(t)}}\right]\left[T_{x,k}^{(t-1)}\right] = 0$$

Asumsikan bahwa ${\sigma_{ix}}^{2(t)}$ berhingga sehingga didapat taksiran untuk ${\mu_{ix}}^{(t)}$ yaitu

$$\sum_{k=1}^{K}\left[\left(y_{ki} - {\mu_{ix}}^{(t)}\right)\right]\left[T_{x,k}^{(t-1)}\right] = 0$$

$$\sum_{k=1}^{K}\left[y_{ki}.T_{x,k}^{(t-1)}\right] - \left[{\mu_{ix}}^{(t)}.T_{x,k}^{(t-1)}\right] = 0$$

$$\sum_{k=1}^{K}\left[y_{ki}.T_{x,k}^{(t-1)}\right] = {\mu_{ix}}^{(t)}\sum_{k=1}^{K} T_{x,k}^{(t-1)}$$

$${\hat{\mu}_{ix}}^{(t)} = \frac{\sum_{k=1}^{K}\left[y_{ki}.T_{x,k}^{(t-1)}\right]}{\sum_{k=1}^{K} T_{x,k}^{(t-1)}} \quad (3.27)$$

Selanjutnya taksiran untuk ${\sigma_{ix}}^{2(t)}$ yaitu

$$\frac{\partial I}{\partial {\sigma_{ix}}^{2(t)}} = \frac{\partial}{\partial {\sigma_{ix}}^{2(t)}}\left(\sum_{x=1}^{C}\sum_{k=1}^{K}\left[-\sum_{i=1}^{n}\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2}{2{\sigma_{ix}}^{2(t)}} - \frac{1}{2}ln\left({\sigma_{ix}}^{2(t)}\right)\right] T_{x,k}^{(t-1)}\right) = 0$$

$$= \sum_{k=1}^{K}\left[\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2}{2\left({\sigma_{ix}}^{2(t)}\right)^2} - \frac{1}{2{\sigma_{ix}}^{2(t)}}\right]\left[T_{x,k}^{(t-1)}\right] = 0$$

$$= \sum_{k=1}^{K}\left[\frac{\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2 - {\sigma_{ix}}^{2(t)}}{2\left({\sigma_{ix}}^{2(t)}\right)^2}\right]\left[T_{x,k}^{(t-1)}\right] = 0$$

Asumsikan bahwa ${\sigma_{ix}}^{2(t)}$ berhingga sehingga didapat taksiran untuk ${\sigma_{ix}}^{2(t)}$ yaitu

$$\sum_{k=1}^{K}\left[\left(y_{ki} - {\mu_{ix}}^{(t)}\right)^2 - {\sigma_{ix}}^{2(t)}\right]\left[T_{x,k}^{(t-1)}\right] = 0$$

$$\sum_{k=1}^{K}\left[\left(y_{ki} - {\mu_{ix}}^{(t)}\right)^2 T_{x,k}^{(t-1)}\right] - \left[{\sigma_{ix}}^{2(t)} T_{x,k}^{(t-1)}\right] = 0$$

$$\sum_{k=1}^{K}\left[\left(y_{ki} - {\mu_{ix}}^{(t)}\right)^2 T_{x,k}^{(t-1)}\right] = {\sigma_{ix}}^{2(t)}\sum_{k=1}^{K} T_{x,k}^{(t-1)}$$

$$\hat{\sigma}_{ix}^{\ 2(t)} = \frac{\sum_{k=1}^{K}\left[\left(y_{ki}-{\mu_{ix}}^{(t)}\right)^2 T_{x,k}^{(t-1)}\right]}{\sum_{k=1}^{K} T_{x,k}^{(t-1)}} \quad (3.28)$$

Proses E-*step* dan M-*step* ini akan dilakukan terus secara iteratif sampai didapatkan suatu taksiran untuk $P(x)$, μ_{ix}, dan ${\sigma_{ix}}^2$ yang konvergen atau didapatkan $\left|\hat{P}(x)^{(t)} - \hat{P}(x)^{(t-1)}\right|$, $\left|{\hat{\mu}_{ix}}^{(t)} - {\hat{\mu}_{ix}}^{(t-1)}\right|$ dan $\left|{\hat{\sigma}_{ix}}^{2(t)} - {\hat{\sigma}_{ix}}^{2(t-1)}\right|$, *i* = 1,…, *n*; *x* = 1,…, *C*, yang cukup kecil.

3.4 Uji Kecocokan Model

Dalam pembentukan suatu model diperlukan uji kecocokan model. Dalam pembentukan *latent profile model*, uji kecocokan model dilakukan untuk setiap banyak kelas *C* > 1, kemudian menentukan banyak kelas optimal yang memberikan model yang cocok . Uji kecocokan model pada tugas akhir ini akan dilakukan dengan uji rasio likelihood.

Akan diuji hipotesis sebagai berikut:

H_0 : Model cocok untuk suatu *C* tertentu ; *C* = 2, 3, ….

H_1 : tidak demikian

Model cocok untuk suatu *C* tertentu jika dapat dibentuk *C* buah kelas laten dimana variabel indikator di setiap kelas laten saling bebas dan berdistribusi normal atau dengan perkataan lain model cocok jika:

$$f(\boldsymbol{y_k}) = \sum_{x=1}^{C} \hat{P}(x) f(\boldsymbol{y}_k | \boldsymbol{\hat{\mu}}_x, \boldsymbol{\hat{\sigma}}_x{}^2)$$

$$= \sum_{x=1}^{C} \frac{\hat{P}(x) exp\left[-\sum_{i=1}^{n} \frac{(y_{ki}-\hat{\mu}_{ix})^2}{2\hat{\sigma}_{ix}{}^2}\right]}{\prod_{i=1}^{n} \hat{\sigma}_{ix}\sqrt{2\pi}}$$

Hipotesis H_1 menyatakan tidak perlu dibentuk kelas laten sehingga model dapat dituliskan sebagai berikut:

$$f(\boldsymbol{y_k}) = \prod_{i=1}^{n} f_i(\boldsymbol{y_k}) = \frac{exp\left[-\sum_{i=1}^{n} \frac{(y_{ki}-\hat{\mu}_i)^2}{2\hat{\sigma}_i{}^2}\right]}{\prod_{i=1}^{n} \hat{\sigma}_i\sqrt{2\pi}}$$

Definisikan fungsi likelihood sebagai berikut:

$$L_0 = \prod_{k=1}^{K}\left[\sum_{x=1}^{C} \hat{P}(x) f(\boldsymbol{y_k} | \boldsymbol{\hat{\mu}}_x, \boldsymbol{\hat{\sigma}}_x{}^2)\right]$$

$$= \prod_{k=1}^{K}\left[\sum_{x=1}^{C} \frac{\hat{P}(x) exp\left[-\sum_{i=1}^{n} \frac{(y_{ki}-\hat{\mu}_{ix})^2}{2\hat{\sigma}_{ix}{}^2}\right]}{\prod_{i=1}^{n} \hat{\sigma}_{ix}\sqrt{2\pi}}\right]$$

$$L_1 = \prod_{k=1}^{K}\left[\prod_{i=1}^{n} f_i(\boldsymbol{y_k})\right]$$

$$= \prod_{k=1}^{K}\left[\frac{exp\left[-\sum_{i=1}^{n} \frac{(y_{ki}-\hat{\mu}_i)^2}{2\hat{\sigma}_i{}^2}\right]}{\prod_{i=1}^{n} \hat{\sigma}_i\sqrt{2\pi}}\right]$$

maka rasio likelihoodnya adalah

$$\lambda = \frac{L_0}{L_1}$$

$$= \prod_{k=1}^{K}\left[\frac{\sum_{x=1}^{C} \hat{P}(x) f(\boldsymbol{y}_k | \boldsymbol{\hat{\mu}}_x, \boldsymbol{\hat{\sigma}}_x{}^2)}{\prod_{i=1}^{n} f_i(\boldsymbol{y_k})}\right]$$

Statistik uji yang digunakan dalam uji kecocokan *latent profile model* adalah

$$L^2 = -2ln\lambda$$

$$= -2ln\left[\prod_{k=1}^{K}\left[\frac{\sum_{x=1}^{C} \hat{P}(x) f(\boldsymbol{y_k} | \boldsymbol{\hat{\mu}}_x, \boldsymbol{\hat{\sigma}}_x{}^2)}{\prod_{i=1}^{n} f_i(\boldsymbol{y_k})}\right]\right]$$

$$= 2ln\left[\prod_{k=1}^{K}\left[\frac{\prod_{i=1}^{n} f_i(\boldsymbol{y_k})}{\sum_{x=1}^{C} \hat{P}(x) f(\boldsymbol{y_k} | \boldsymbol{\hat{\mu}}_x, \boldsymbol{\hat{\sigma}}_x{}^2)}\right]\right]$$

$$= 2\sum_{k=1}^{K}\left[ln\left[\frac{\prod_{i=1}^{n} f_i(\boldsymbol{y_k})}{\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}\right]\right]$$

$$= 2\sum_{k=1}^{K}\left[ln\left[\prod_{i=1}^{n} f_i(\boldsymbol{y_k})\right] - ln\left[\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)\right]\right]$$

$$= 2\sum_{k=1}^{K}\left[ln\left[\frac{exp\left[-\sum_{i=1}^{n}\frac{(y_{ki}-\hat{\mu}_i)^2}{2\hat{\sigma}_i{}^2}\right]}{\prod_{i=1}^{n}\hat{\sigma}_i\sqrt{2\pi}}\right] - ln\left[\sum_{x=1}^{C}\frac{\hat{P}(x)exp\left[-\sum_{i=1}^{n}\frac{(y_{ki}-\hat{\mu}_{ix})^2}{2\hat{\sigma}_{ix}{}^2}\right]}{\prod_{i=1}^{n}\hat{\sigma}_{ix}\sqrt{2\pi}}\right]\right]$$

L^2 akan berdistribusi χ_r^2, dimana $r = r_2 - r_1 = (Kn - 2n) - (Kn - 2nx - x) = (2nx + x) - 2n$. Dimana r_1 adalah derajat bebas untuk model dibawah H_0 dan r_2 adalah derajat bebas untuk model dibawah H_1. H_0 ditolak jika $L^2 > \chi_r^2$.

3.5 Menentukan Banyak Kelas Laten Optimal

Untuk memilih *latent profile model* terbaik, diantara yang cocok, atau dengan kata lain menentukan banyak kelas optimal yang cukup menjelaskan hubungan diantara variabel-variabel indikatornya adalah dengan menggunakan uji rasio likelihood. Akan diuji hipotesis sebagai berikut:

H_0 : Model dengan C kelas

H_1 : Model dengan $C + 1$ kelas

Pada pengujian hipotesis tersebut, ingin dilihat apakah penambahan kelas laten pada model signifikan.

Definisikan fungsi likelihood sebagai berikut:

$$L_0 = \prod_{k=1}^{K}\left[\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)\right]$$

$$= \prod_{k=1}^{K}\left[\sum_{x=1}^{C}\frac{\hat{P}(x)exp\left[-\sum_{i=1}^{n}\frac{(y_{ki}-\hat{\mu}_{ix})^2}{2\hat{\sigma}_{ix}{}^2}\right]}{\prod_{i=1}^{n}\hat{\sigma}_{ix}\sqrt{2\pi}}\right]$$

$$L_1 = \prod_{k=1}^{K}\left[\sum_{x=1}^{C+1}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)\right]$$

$$= \prod_{k=1}^{K}\left[\sum_{x=1}^{C+1}\frac{\hat{P}(x)exp\left[-\sum_{i=1}^{n}\frac{(y_{ki}-\hat{\mu}_{ix})^2}{2\hat{\sigma}_{ix}{}^2}\right]}{\prod_{i=1}^{n}\hat{\sigma}_{ix}\sqrt{2\pi}}\right]$$

maka rasio likelihoodnya adalah

$$\lambda = \frac{L_0}{L_1}$$

$$= \frac{\prod_{k=1}^{K}\left[\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)\right]}{\prod_{k=1}^{K}\left[\sum_{x=1}^{C+1}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)\right]}$$

$$= \prod_{k=1}^{K}\left[\frac{\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}{\sum_{x=1}^{C+1}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}\right]$$

Statistik uji yang digunakan dalam uji kecocokan *latent profile model* adalah

$$L^2 = -2ln\lambda$$

$$= -2ln\left[\prod_{k=1}^{K}\left[\frac{\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}{\sum_{x=1}^{C+1}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}\right]\right]$$

$$= 2ln\left[\prod_{k=1}^{K}\left[\frac{\sum_{x=1}^{C+1}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}{\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}\right]\right]$$

$$= 2\sum_{k=1}^{K} ln\left[\frac{\sum_{x=1}^{C+1}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}{\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\boldsymbol{\hat{\mu}}_x,\boldsymbol{\hat{\sigma}}_x{}^2)}\right]$$

L^2 akan berdistribusi χ_r^2, dimana r = banyaknya parameter yang ditaksir pada model dengan C + 1 kelas – banyaknya parameter yang ditaksir pada model dengan C kelas = $(2nx + x + 2n + 1) - (2nx + x) = 2n + 1$. H_0 ditolak jika $L^2 > \chi_r^2$. Jika H_0 ditolak maka penambahan kelas laten pada *latent profile model* signifikan dan banyaknya kelas optimal setidaknya ada C + 1 kelas. Uji diteruskan sampai penambahan kelas laten tidak signifikan lagi atau dengan perkataan lain H_0 tidak ditolak dan banyak kelas laten optimal ada sebanyak C.

BAB 4
CONTOH APLIKASI PEMBENTUKAN KELAS LATEN PADA VARIABEL LATEN "TINGKAT MENGATUR DIRI SENDIRI"

Sebagai contoh terapan dalam tugas akhir ini, akan dibentuk kelas-kelas dari variabel laten tingkat mengatur diri sendiri berdasarkan variabel indikator tingkat ketaktergantungan, skor tanggung jawab dan tingkat ketenangan yang bersifat kontinu. Dalam hal ini variabel laten "tingkat mengatur diri sendiri" dalam bentuk kategorik akan lebih mudah untuk diinterpretasikan. Penulis mengangkat contoh ini karena kemampuan mengatur diri sendiri sangat diperlukan mahasiswa baru khususnya dalam proses penyesuaian diri dari masa SMA ke perguruan tinggi.

4.1 Sumber Data

Analisis data dilakukan berdasarkan seluruh mahasiswa baru departemen matematika FMIPA UI angkatan 2010. Pengukuran setiap variabel indikator dilakukan berdasarkan skala likert dengan jarak yang sama. Sharma (1996) mengatakan bahwa pengukuran dengan skala likert akan menghasilkan skala interval jika perbedaan kategori yang berurutan sama.

4.2 Analisis Data

Berdasarkan data yang diperoleh penulis, akan dibentuk kelas untuk variabel laten tingkat mengatur diri sendiri berdasarkan variabel indikatornya yaitu

1. Y_1 = tingkat ketaktergantungan
2. Y_2 = skor tanggung jawab
3. Y_3 = tingkat ketenangan

Data diberikan pada lampiran 8. Pada kasus ini, secara umum *latent profile model* dapat dituliskan sebagai:

$$f(y) = \sum_{x=1}^{C} P(x) f(\boldsymbol{y}|\boldsymbol{\mu}_x, \boldsymbol{\sigma}_x{}^2)$$

untuk suatu *C* tertentu, *C* = 1, 2,…

Dengan menggunakan software Matlab diperoleh nilai taksiran parameter *P*(*x*), μ_{ix} dan $\sigma_{ix}{}^2$ untuk model dengan banyak kelas *C* tertentu. Untuk setiap model dengan banyak kelas = *C* tersebut akan dilakukan pengujian hipotesis untuk melihat kecocokan model menggunakan uji rasio likelihood dengan statistik uji L^2. Selanjutnya akan dipilih *latent profile model* terbaik yang akan dipilih menggunakan uji rasio likelihood dengan statistik uji L^2.

Berikut adalah nilai taksiran parameter *P*(*x*), μ_{ix} dan $\sigma_{ix}{}^2$ untuk model dengan banyak kelas *C* tertentu.

Tabel 4.1 Nilai Taksiran untuk Parameter *P*(*x*) untuk Model dengan Banyak Kelas *C* Tertentu.

	P(1)	*P*(2)	*P*(3)	*P*(4)	…
C = 1	1				
C = 2	0.61	0.39			
C = 3	0.35	0.23	0.42		
C = 4	0.0.24	0.42	0.1	0.24	
⋮					

Tabel 4.2 Nilai Taksiran untuk Parameter μ_{ix} untuk Model dengan Banyak Kelas *C* Tertentu.

		μ_{i1}	μ_{i2}	μ_{i3}	μ_{i4}	…
C = 1	*i* =1	17.6087				
	i =2	22.5507				
	i =3	17.3043				
C = 2	*i* =1	17.6531	17.5379			
	i =2	23.0847	21.6999			
	i =3	19.4135	13.9437			

C = 3	i =1	17.5719	14.6879	19.2363		
	i =2	21.599	24.2508	22.3967		
	i =3	13.6985	20.3798	18.5615		
C = 4	i =1	14.6489	19.2488	15.0016	18.7363	
	i =2	23.6134	22.421	25.1392	20.7011	
	i =3	16.6508	18.6988	23.1392	13.2016	
⋮						

Tabel 4.3 Nilai Taksiran untuk Parameter σ_{ix}^2 untuk Model dengan Banyak Kelas *C* Tertentu.

		σ_{i1}^2	σ_{i2}^2	σ_{i3}^2	σ_{i4}^2	…
C = 1	i =1	7.095				
	i =2	6.104				
	i =3	11.45				
C = 2	i =1	7.35	6.413			
	i =2	4.7646	6.8306			
	i =3	4.971	2.9611			
C = 3	i =1	5.5354	2.7172	3.2007		
	i =2	7.312	4.3196	3.5433		
	i =3	2.7254	8.4822	2.4406		
C = 4	i =1	2.618	3.1761	3.4262	2.1237	
	i =2	4.5598	3.5809	2.4122	5.6825	
	i =3	3.1845	2.3316	3.5584	2.6764	
⋮						

Setelah diperoleh nilai taksiran dari tiap parameter, untuk setiap *latent profile model* dengan banyak kelas = C, C = 1, 2, …, akan dilakukan pengujian hipotesis untuk setiap *latent profile model* dengan banyak kelas $C > 1$ menggunakan uji rasio likelihood sebagai berikut:

Hipotesis:

H_0 : Model cocok untuk suatu C tertentu ; C = 2, 3, ….

H_1 : tidak demikian

Tingkat signifikansi: α = 0.05

Statistik uji:

$$L^2 = 2\sum_{k=1}^{69}\left[ln\left[\frac{\prod_{i=1}^{3} f_i(\boldsymbol{y_k})}{\sum_{x=1}^{C}\hat{P}(x)f(\boldsymbol{y_k}|\hat{\boldsymbol{\mu}}_x,\hat{\boldsymbol{\sigma}}_x{}^2)}\right]\right]$$

Aturan keputusan: H_0 ditolak jika $L^2 > \chi^2_{r,0.05}$; $r = 7x - 6$.

Dari analisis data didapat bahwa model yang cocok adalah untuk $C = 2$.

Setelah model cocok, selanjutnya adalah memilih *latent profile model* terbaik, diantara yang cocok, atau dengan kata lain menentukan banyak kelas optimal yang cukup menjelaskan hubungan diantara variabel-variabel indikatornya dengan menggunakan uji rasio likelihood sebagai berikut:

Hipotesis:

H_0 : Model dengan 2 kelas

H_1 : Model dengan 3 kelas

Tingkat signifikansi: α = 0.05

Statistik uji:

$$L^2 = 2\sum_{k=1}^{69} ln\left[\frac{\sum_{x=1}^{3}\hat{P}(x)f(\boldsymbol{y_k}|\hat{\boldsymbol{\mu}}_x,\hat{\boldsymbol{\sigma}}_x{}^2)}{\sum_{x=1}^{2}\hat{P}(x)f(\boldsymbol{y_k}|\hat{\boldsymbol{\mu}}_x,\hat{\boldsymbol{\sigma}}_x{}^2)}\right]$$

Aturan keputusan: H_0 ditolak jika $L^2 > \chi^2_{r,0.05}$; $r = 7$.

Dari hasil perhitungan didapat bahwa nilai $L^2 = 3.572 < \chi^2_{7,0.05} = 14.1$ sehingga H_0 tidak ditolak.

Kesimpulan: banyaknya kelas untuk variabel laten tingkat mengatur diri sendiri adalah dua kelas.

Sehingga, *latent profile model* pada kasus ini adalah sebagai berikut:

$$f(y) = \sum_{x=1}^{2}\hat{P}(x)f\left(\boldsymbol{y_k}\middle|\hat{\boldsymbol{\mu}}_x,\hat{\boldsymbol{\sigma}}_x{}^2\right)$$

$$= 0.61\, f\left(\boldsymbol{y_k}\middle|\hat{\boldsymbol{\mu}}_1,\hat{\boldsymbol{\sigma}}_1{}^2\right) + 0.39\, f\left(\boldsymbol{y_k}\middle|\hat{\boldsymbol{\mu}}_2,\hat{\boldsymbol{\sigma}}_2{}^2\right)$$

dengan

$$\boldsymbol{\mu}_1 = \begin{bmatrix} 17.6531 \\ 23.0847 \\ 19.4135 \end{bmatrix}, \quad \boldsymbol{\mu}_2 = \begin{bmatrix} 17.5379 \\ 21.6999 \\ 13.9437 \end{bmatrix}$$

dan

$$\boldsymbol{\sigma}_1^2 = \begin{bmatrix} 7.35 \\ 4.7646 \\ 4.971 \end{bmatrix}, \quad \boldsymbol{\sigma}_2^2 = \begin{bmatrix} 6.413 \\ 6.8306 \\ 2.9611 \end{bmatrix}$$

Jadi, dari analisis diatas dapat disimpulkan bahwa tingkat mengatur diri sendiri dari mahasiswa baru Departemen Matematika FMIPA UI dapat dikelompokkan menjadi dua kelas yaitu:

1. Mahasiswa yang masuk dalam kelas 1 mempunyai profil:
   - mean tingkat ketaktergantungan sebesar 17.6531 dan variansi 7.35
   - mean skor tanggung jawab sebesar 23.0847 dan variansi 4.7646
   - mean tingkat ketenangan sebesar 19.4135 dan variansi 4.971

   dengan probabilitas 0.61.
2. Mahasiswa yang masuk dalam kelas 2 mempunyai profil:
   - mean tingkat ketaktergantungan sebesar 17.5379 dan variansi 6.413
   - mean skor tanggung jawab sebesar 21.6999 dan variansi 6.8306
   - mean tingkat ketenangan sebesar 13.9437 dan variansi 2.9611

   dengan probabilitas 0.39.

BAB 5
PENUTUP

5.1 Kesimpulan

Dalam tugas akhir ini dapat disimpulkan bahwa:

1) Pembentukan kelas-kelas dari variabel laten kategorik berdasarkan variabel indikator kontinu dapat dilakukan dengan *latent profile model*.
2) Penentuan banyak kelas laten optimal untuk variabel laten dapat dicari dengan menggunakan uji rasio likelihood berdasarkan *latent profile model* yang melibatkan mean dari variabel indikator, variansi dari variabel indikator dan probabilitas suatu individu masuk pada kelas laten tertentu.

   Mean dari variabel indikator, variansi dari variabel indikator dan probabilitas suatu individu masuk pada kelas laten tertentu ditaksir dengan metode *joint* maksimum likelihood dan diselesaikan dengan algoritma EM (*Expectation-Maximization*).

5.2 Saran

- Tugas akhir ini dapat dilanjutkan untuk pengelompokkan individu-individu ke dalam kelas laten yang terbentuk.
- Tugas akhir ini dapat dilanjutkan untuk melibatkan sifat keordinalan dari kelas laten.

DAFTAR PUSTAKA

Borman, Sean. 2004. *The EM Algorithm A Short Tutorial*.

Dunmur, AP. & D. M. Titterington. 1998. *Parameter Estimation in Latent Profile Models*. Computational Statistics and Data Analysis 27, pp. 371-388.

Hogg, Robert V. & Aleen T. Craig. 1995. *Introduction to Mathematical Statistics Fifth Edition*. New Jersey: Prentice-Hall International. Inc.

Klugman, Stuart A. Harry H. Panjer. & Gordon T. Wilmot. 2004. *Loss Models From Data to Decisions Second Edition.*

Novianti. 2008. *Latent Class Model*. Depok: Universitas Indonesia.

Persadanta, Pintanugra. 2008. *Item Response Model*. Depok: Universitas Indonesia.

Roussas, George G. 1997. *A Course in Mathematical Statistics Second Edition*. San Diego: Academic Press.

Takane, Y. 1976. *A Statistical Procedure for The Latent Profile Model*. Japanese Psychological Research Vol. 18, No. 2, pp. 82-90.

Wade, Tracey D, PhD. Ross D. Crosby, PhD. & Nicholas G. Martin. 2006. *Use of Latent Profile Analysis to Identify Eating Disorder Phenotypes in an Adult Australian Twin Cohort*. Arch Gen Psychiatry.

Welling, Max. *EM-Algorithm*. Pasadena: California Institute of Technology.

Lampiran 1 Menunjukkan, $\boldsymbol{\theta}$ memaksimumkan $L(\boldsymbol{\theta}) \leftrightarrow \boldsymbol{\theta}$ memaksimumkan $lnL(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, \theta_{2, \ldots,} \theta_m)$

Bukti:

($\rightarrow$) karena $\boldsymbol{\theta}$ memaksimumkan $L(\boldsymbol{\theta})$ maka

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_1} = 0,$$

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_2} = 0,$$

$$\vdots$$

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_m} = 0,$$

$$\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2} < 0, \quad i = 1,\ldots, m,$$

$$D = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2} . \frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_j}^2} - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2 > 0, \quad i = 1,\ldots, m, \quad j = 1,\ldots,m, \quad i \neq j$$

Akan ditunjukkan $\boldsymbol{\theta}$ juga memaksimumkan $lnL(\boldsymbol{\theta})$ yaitu

- $\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_1} = 0$

  $$\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_1} = \frac{1}{L(\boldsymbol{\theta})} . \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_1} = \frac{1}{L(\boldsymbol{\theta})} . 0 = 0$$

  $$\therefore \frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_1} = 0$$

- $\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_2} = 0$

  $$\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_2} = \frac{1}{L(\boldsymbol{\theta})} . \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_2} = \frac{1}{L(\boldsymbol{\theta})} . 0 = 0$$

  $$\therefore \frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_2} = 0$$

  $$\vdots$$

- $\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_m} = 0$

  $$\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_m} = \frac{1}{L(\boldsymbol{\theta})} . \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_m} = \frac{1}{L(\boldsymbol{\theta})} . 0 = 0$$

  $$\therefore \frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_m} = 0$$

- $\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2} < 0, \ i = 1,\ldots, m$

  $$\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2} = \frac{\partial}{\partial \theta_i}\left(\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_i}\right) = \frac{\partial}{\partial \theta_i}\left(\frac{1}{L(\boldsymbol{\theta})} . \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}\right)$$

(lanjutan)

$$= \frac{\partial}{\partial\theta_i}\left(\frac{1}{L(\boldsymbol{\theta})}\right).\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i} + \frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{1}{L(\boldsymbol{\theta})} = \frac{\partial}{\partial\theta_i}\left(\frac{1}{L(\boldsymbol{\theta})}\right).0 + \frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{1}{L(\boldsymbol{\theta})}$$

$$= \frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{1}{L(\boldsymbol{\theta})} < 0.\frac{1}{L(\boldsymbol{\theta})} = 0$$

$$\therefore \ \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_i}^2} < 0, \ i = 1,\ldots, m$$

- $D = \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_j}^2} - \left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2 > 0, \ i = 1,\ldots, m, \quad j = 1,\ldots,m, \quad i \neq j$

  dimana

$$\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_i}^2} = \frac{\partial}{\partial\theta_i}\left(\frac{\partial lnL(\boldsymbol{\theta})}{\partial\theta_i}\right) = \frac{\partial}{\partial\theta_i}\left(\frac{\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}}{L(\boldsymbol{\theta})}\right), \quad i = 1,\ldots, m$$

$$= \frac{\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.L(\boldsymbol{\theta}) - \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}}{L(\boldsymbol{\theta})^2}$$

$$\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_j}^2} = \frac{\partial}{\partial\theta_j}\left(\frac{\partial lnL(\boldsymbol{\theta})}{\partial\theta_j}\right) = \frac{\partial}{\partial\theta_j}\left(\frac{\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}}{L(\boldsymbol{\theta})}\right), \quad j = 1,\ldots, m$$

$$= \frac{\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2}.L(\boldsymbol{\theta}) - \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}}{L(\boldsymbol{\theta})^2}$$

$$\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i} = \frac{\partial}{\partial\theta_j}\left(\frac{\partial lnL(\boldsymbol{\theta})}{\partial\theta_i}\right) = \frac{\partial}{\partial\theta_j}\left(\frac{\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}}{L(\boldsymbol{\theta})}\right), \quad i = 1,\ldots, m, \ j = 1,\ldots, m, \ i \neq j$$

$$= \frac{\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}.L(\boldsymbol{\theta}) - \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}}{L(\boldsymbol{\theta})^2}$$

$$D = \left[\frac{\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.L(\boldsymbol{\theta}) - \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}}{L(\boldsymbol{\theta})^2}\right].\left[\frac{\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2}.L(\boldsymbol{\theta}) - \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}}{L(\boldsymbol{\theta})^2}\right]$$

$$- \left[\frac{\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}.L(\boldsymbol{\theta}) - \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}}{L(\boldsymbol{\theta})^2}\right]^2, \quad i = 1,\ldots, m, \ j = 1,\ldots, m, \ i \neq j$$

$$= \frac{1}{L(\boldsymbol{\theta})^4}\left[\left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2}.L(\boldsymbol{\theta})^2\right) + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}\right)^2 \right.$$

$$\left. - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\left(\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}\right)^2.L(\boldsymbol{\theta})\right) - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2}.\left(\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}\right)^2.L(\boldsymbol{\theta})\right)\right]$$

(lanjutan)

$$-\frac{1}{L(\boldsymbol{\theta})^4}\left[\left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2 . L(\boldsymbol{\theta})^2 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}\right)^2\right.$$

$$\left.-2\left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}.L(\boldsymbol{\theta}).\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}\right)\right]$$

$$= \frac{L(\boldsymbol{\theta})^2}{L(\boldsymbol{\theta})^4}\left[\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2} - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2\right]$$

$$+\frac{L(\boldsymbol{\theta})}{L(\boldsymbol{\theta})^4}\left[2.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i} - \frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}\left(\frac{\partial L(\theta)}{\partial\theta_j}\right)^2 - \frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2}.\left(\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}\right)^2\right]$$

$$= \frac{1}{L(\boldsymbol{\theta})^2}\left[\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2} - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2\right] + \frac{1}{L(\boldsymbol{\theta})^3}\left[2.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}.0.0\right.$$

$$\left.-\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.(0)^2 - \frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2}.(0)^2\right]$$

$$= \frac{1}{L(\boldsymbol{\theta})^2}\left[\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial{\theta_j}^2} - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2\right] > \frac{1}{L(\boldsymbol{\theta})^2}.0 = 0$$

$\therefore D = \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_j}^2} - \left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2 > 0$, $i = 1,\ldots, m$, $j = 1,\ldots, m$, $i \neq j$

Dengan perkataan lain terbukti bahwa $\boldsymbol{\theta}$ juga memaksimumkan $lnL(\boldsymbol{\theta})$.

($\leftarrow$) karena $\boldsymbol{\theta}$ memaksimumkan $lnL(\boldsymbol{\theta})$ maka

$$\frac{\partial lnL(\boldsymbol{\theta})}{\partial\theta_1} = 0,$$

$$\frac{\partial lnL(\boldsymbol{\theta})}{\partial\theta_2} = 0,$$

$$\vdots$$

$$\frac{\partial lnL(\boldsymbol{\theta})}{\partial\theta_m} = 0$$

$$\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_i}^2} < 0, \quad i = 1,\ldots, m$$

$$D = \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial{\theta_j}^2} - \left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}\right)^2 > 0, \quad i = 1,\ldots, m,\ j = 1,\ldots, m,\ i \neq j$$

Akan ditunjukkan $\boldsymbol{\theta}$ juga memaksimumkan $L(\boldsymbol{\theta})$ yaitu:

- $\frac{\partial L(\boldsymbol{\theta})}{\partial\theta_1} = 0$

(lanjutan)

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_1} = L(\boldsymbol{\theta}).\frac{1}{L(\boldsymbol{\theta})}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_1} = L(\boldsymbol{\theta}).\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_1} = L(\boldsymbol{\theta}).0 = 0$$

$\therefore\ \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_1} = 0$

- $\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_2} = 0$

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_2} = L(\boldsymbol{\theta}).\frac{1}{L(\boldsymbol{\theta})}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_2} = L(\boldsymbol{\theta}).\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_2} = L(\boldsymbol{\theta}).0 = 0$$

$\therefore\ \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_2} = 0$

$\vdots$

- $\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_m} = 0$

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_m} = L(\boldsymbol{\theta}).\frac{1}{L(\boldsymbol{\theta})}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_m} = L(\boldsymbol{\theta}).\frac{\partial lnL(\boldsymbol{\theta})}{\partial \theta_m} = L(\boldsymbol{\theta}).0 = 0$$

$\therefore\ \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_m} = 0$

- $\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2} < 0, \quad i = 1,\ldots, m$

$$\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2} = \frac{1}{L(\boldsymbol{\theta})}.\left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.L(\boldsymbol{\theta})^2 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}\right)^2\right]$$

$$= \frac{1}{L(\boldsymbol{\theta})}.\left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.L(\boldsymbol{\theta})^2 + 0\right]$$

$$= \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.L(\boldsymbol{\theta}) < 0.L(\boldsymbol{\theta}) = 0$$

$\therefore\ \frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2} < 0, \quad i = 1,\ldots, m$

- $D = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_j}^2} - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2 > 0, \quad i = 1,\ldots, m, \quad j = 1,\ldots,m, \quad i \neq j$

dimana

$$\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2} = \frac{\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.L(\boldsymbol{\theta})^2 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}\right)^2}{L(\boldsymbol{\theta})}, \quad i = 1,\ldots, m$$

$$\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_j}^2} = \frac{\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2}.L(\boldsymbol{\theta})^2 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}\right)^2}{L(\boldsymbol{\theta})}, \quad j = 1,\ldots, m$$

$$\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i} = \frac{\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}.L(\boldsymbol{\theta})^2 + \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}}{L(\boldsymbol{\theta})}, \quad i = 1,\ldots, m, \quad j = 1,\ldots,m, \quad i \neq j$$

(lanjutan)

$$D = \left[\frac{\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.L(\boldsymbol{\theta})^2 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}\right)^2}{L(\boldsymbol{\theta})}\right] . \left[\frac{\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2}.L(\boldsymbol{\theta})^2 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}\right)^2}{L(\boldsymbol{\theta})}\right]$$

$$- \left[\frac{\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}.L(\boldsymbol{\theta})^2 + \frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}}{L(\boldsymbol{\theta})}\right]^2, \quad i = 1,\dots, m, \quad j = 1,\dots,m, \quad i \neq j$$

$$= \frac{1}{L(\boldsymbol{\theta})^2}\left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2}.L(\boldsymbol{\theta})^4 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}\right)^2\right.$$

$$\left. + \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}\right)^2.L(\boldsymbol{\theta})^2 + \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2}.\left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}\right)^2.L(\boldsymbol{\theta})^2\right]$$

$$- \frac{1}{L(\boldsymbol{\theta})^2}\left[\left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2.L(\boldsymbol{\theta})^4 + \left(\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}\right)^2\right.$$

$$\left. + 2\left(\frac{\partial^2 lnL(\theta)}{\partial \theta_j \partial \theta_i}.L(\boldsymbol{\theta})^2.\frac{\partial L(\theta)}{\partial \theta_j}.\frac{\partial L(\theta)}{\partial \theta_i}\right)\right]$$

$$= \frac{L(\boldsymbol{\theta})^4}{L(\boldsymbol{\theta})^2}\left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2} - \left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2\right]$$

$$+ \frac{L(\boldsymbol{\theta})^2}{L(\boldsymbol{\theta})^2}\left[\frac{\partial^2 lnL(\theta)}{\partial {\theta_i}^2}.\left(\frac{\partial L(\theta)}{\partial \theta_j}\right)^2 + \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2}.\left(\frac{\partial L(\theta)}{\partial \theta_i}\right)^2\right.$$

$$\left. - 2\left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_j}.\frac{\partial L(\boldsymbol{\theta})}{\partial \theta_i}\right)\right]$$

$$= L(\boldsymbol{\theta})^2\left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2} - \left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2\right] + \left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.(0)^2 + \right.$$

$$\left. \frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2}.(0)^2 - 2.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}.0.0\right]$$

$$= L(\boldsymbol{\theta})^2\left[\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial {\theta_j}^2} - \left(\frac{\partial^2 lnL(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2\right] > L(\boldsymbol{\theta})^2.0 = 0$$

$$\therefore\ D = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_i}^2}.\frac{\partial^2 L(\boldsymbol{\theta})}{\partial {\theta_j}^2} - \left(\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \theta_j \partial \theta_i}\right)^2 > 0, \quad i = 1,\dots, m, \quad j = 1,\dots,m, \quad i \neq j$$

Dengan perkataan lain terbukti bahwa $\boldsymbol{\theta}$ juga memaksimumkan $L(\boldsymbol{\theta})$. Jadi, terbukti bahwa $\boldsymbol{\theta}$ memaksimumkan $L(\boldsymbol{\theta}) \leftrightarrow \boldsymbol{\theta}$ memaksimumkan $lnL(\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, \theta_{2, \dots,} \theta_m)$.

Lampiran 2   Menunjukkan bahwa – log (x) adalah fungsi konveks

Suatu fungsi dikatakan fungsi konveks jika:

$\forall\, x_1, x_2 \in$ domain (f) → untuk $\lambda \in [0,1]$

Berlaku $\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)$

Misalkan $f(x) = -\log(x)$,

Maka $f'(x) = -\frac{1}{x}$

$f''(x) = \frac{1}{x^2}$

Karena x^2 selalu bernilai positif, maka $f''(x) = \frac{1}{x^2} > 0$.

Berdasarkan teorema: "Jika $f(x)$ dapat diturunkan 2 kali dan $f''(x) \geq 0$ maka $f(x)$ adalah fungsi konveks",

Jadi, dapat dinyatakan bahwa – log(x) adalah fungsi konveks.

Lampiran 3 Membuktikan pertidaksamaan Jensen

"Jika f adalah fungsi konveks, dan untuk $\lambda_i \in [0,1]$ sedemikian sehingga $\sum_{i=1}^{n} \lambda_i = 1$, maka berlaku $\sum_{i=1}^{n} \lambda_i f(x_i) \geq f(\sum_{i=1}^{n} \lambda_i x_i)$."

Bukti:

Misalkan f fungsi konveks, maka berdasarkan definisi fungsi konveks,

$\forall\, x_1, x_2 \in$ domain (f) → untuk $\lambda \in [0,1]$ berlaku $\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)$

Pembuktian pertidaksamaan Jensen akan dilakukan dengan induksi matematika.

- Akan dibuktikan untuk $n = 2$ pertidaksamaan Jensen benar.

  Adib: Jika f adalah fungsi konveks dan untuk $\lambda_i \in [0,1]$ sedemikian sehingga $\sum_{i=1}^{2} \lambda_i = 1$, maka berlaku $\sum_{i=1}^{2} \lambda_i f(x_i) \geq f(\sum_{i=1}^{2} \lambda_i x_i)$

  Bukti:

$$
\begin{aligned}
\sum_{i=1}^{2} \lambda_i f(x_i) &= \lambda_1 f(x_1) + \lambda_2 f(x_2) \\
&= \lambda_1 f(x_1) + (1-\lambda_1) f(x_2) \quad \leftarrow \text{karena } \textstyle\sum_{i=1}^{2} \lambda_i = 1 \\
&\geq f(\lambda_1 x_1 + (1-\lambda_1) x_2) \quad \leftarrow \text{karena } f \text{ fungsi konveks} \\
&\geq f(\textstyle\sum_{i=1}^{2} \lambda_i x_i)
\end{aligned}
$$

  Jadi, terbukti bahwa untuk $n = 2$ pertidaksamaan Jensen benar.

- Misalkan pertidaksamaan Jensen benar untuk n.

  Berlaku:

  Jika f adalah fungsi konveks dan untuk $\lambda_i \in [0,1]$ sedemikian sehingga $\sum_{i=1}^{n} \lambda_i = 1$, maka berlaku

$$\sum_{i=1}^{n} \lambda_i f(x_i) \geq f(\textstyle\sum_{i=1}^{n} \lambda_i x_i)$$

$$\lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n) \geq f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n)$$

- Akan dibuktikan untuk $n + 1$ pertidaksamaan Jensen juga benar.

  Adib: Jika f fungsi konveks dan untuk $\lambda_i \in [0,1]$ sedemikian sehingga $\sum_{i=1}^{n+1} \lambda_i = 1$, maka:

$$\sum_{i=1}^{n+1} \lambda_i f(x_i) \geq f(\textstyle\sum_{i=1}^{n+1} \lambda_i x_i)$$

$$\lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n) + \lambda_{n+1} f(x_{n+1})$$

(lanjutan)

$$\geq \ f(\lambda_1 x_1 + \lambda_2 x_2 + \ \dots + \lambda_n x_n + \ \lambda_{n+1} x_{n+1})$$

Bukti:

Karena

$$\sum_{i=1}^{n+1} \lambda_i = 1$$

$$\sum_{i=1}^{n} \lambda_i \ = 1 - \ \lambda_{n+1}$$

$$= \frac{\sum_{i=1}^{n} \lambda_i}{1- \lambda_{n+1}} = 1$$

$$= \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_n}{1- \lambda_{n+1}} = 1$$

$$= \frac{\lambda_1}{1- \lambda_{n+1}} + \frac{\lambda_2}{1- \lambda_{n+1}} + \ \dots + \frac{\lambda_n}{1- \lambda_{n+1}} = 1$$

maka,

$$\frac{\lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)}{1- \lambda_{n+1}} = \frac{\lambda_1}{1- \lambda_{n+1}} f(x_1) + \ \dots + \frac{\lambda_n}{1- \lambda_{n+1}} f(x_n)$$

$$\geq \ f\left(\frac{\lambda_1}{1- \lambda_{n+1}} x_1 + \ \dots + \frac{\lambda_n}{1- \lambda_{n+1}} x_n\right)$$

Jadi,

$$\frac{\lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)}{1- \lambda_{n+1}} \ \geq \ f\left(\frac{\lambda_1}{1- \lambda_{n+1}} x_1 + \ \dots + \frac{\lambda_n}{1- \lambda_{n+1}} x_n\right)$$

Kalikan kedua ruas dengan $(1 - \ \lambda_{n+1})$ didapat:

$$\lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) \geq (1 - \ \lambda_{n+1}) f\left(\frac{\lambda_1}{1- \lambda_{n+1}} x_1 + \ \dots + \frac{\lambda_n}{1- \lambda_{n+1}} x_n\right)$$

Tambahkan kedua ruas dengan $\lambda_{n+1} f(x_{n+1})$ didapat:

$$\lambda_1 f(x_1) + \cdots + \lambda_n f(x_n) + \lambda_{n+1} f(x_{n+1}) \geq (1 - \ \lambda_{n+1}) f\left(\frac{\lambda_1}{1- \lambda_{n+1}} x_1 + \ \dots + \frac{\lambda_n}{1- \lambda_{n+1}} x_n\right) + \lambda_{n+1} f(x_{n+1})$$

atau

$$\sum_{i=1}^{n+1} \lambda_i f(x_i) \ \geq (1 - \ \lambda_{n+1}) \left(f\left(\sum_{i=1}^{n} \frac{\lambda_i}{1- \lambda_{n+1}} x_i \right) \right) + \lambda_{n+1} f(x_{n+1})$$

Karena $(1 - \ \lambda_{n+1}) + \ \lambda_{n+1} = 1$ dan f adalah fungsi konveks, maka berdasarkan definisi fungsi konveks, didapat:

(lanjutan)

$$\sum_{i=1}^{n+1} \lambda_i\, f(x_i) \ \geq (1-\ \lambda_{n+1}) f\left(\sum_{i=1}^{n} \frac{\lambda_i}{1-\ \lambda_{n+1}} x_i\right) + \lambda_{n+1}\, f(x_{n+1})$$

$$\geq f\left((1-\ \lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1-\ \lambda_{n+1}} x_i + \lambda_{n+1}\, x_{n+1}\right)$$

$$= f\left(\sum_{i=1}^{n} (1-\ \lambda_{n+1}) \frac{\lambda_i}{1-\ \lambda_{n+1}} x_i + \lambda_{n+1}\, x_{n+1}\right)$$

$$= f\left(\sum_{i=1}^{n} \lambda_i x_i + \lambda_{n+1}\, x_{n+1}\right) = \ f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right)$$

atau

$$\sum_{i=1}^{n+1} \lambda_i\, f(x_i) \ \geq \ f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right)$$

Jadi, terbukti bahwa untuk (n + 1) pertidaksamaan Jensen benar. Karena untuk n = 2, n, dan (n+1) pertidaksamaan Jensen terbukti benar, maka dapat disimpulkan bahwa pertidaksamaan Jensen benar untuk setiap n.

"$\forall\, n$, jika f adalah fungsi konveks, dan untuk $\lambda_i \in [0,1]$ sedemikian sehingga $\sum_{i=1}^{n} \lambda_i = 1$, maka berlaku $\sum_{i=1}^{n} \lambda_i\, f(x_i) \geq f(\sum_{i=1}^{n} \lambda_i x_i)$."

Lampiran 4 Membuktikan $E\,[u(x)] \geq u\,(E[x])$

Jika $u(x)$ adalah sebarang fungsi dari variabel random X dan $u(x)$ adalah fungsi konveks, maka berlaku:

$$E\,[u(x)] \geq u\,(E[X])$$

Bukti:

Misalkan X adalah variabel random diskret dan $f(x)$ adalah pdf dari X.

Misalkan terdapat n observasi, maka $\sum_{i=1}^{n} f(x_i) = 1$ dan

$$E[u(x)] = \sum_{i=1}^{n} u(x_i) f(x_i)$$

Jika $u(x)$ adalah fungsi konveks, maka berdasarkan pertidaksamaan Jensen,

$E\,[u(x)]$ dapat dituliskan menjadi:

$$\begin{aligned} E[u(x)] &= \sum_{i=1}^{n} u(x_i) f(x_i) \\ &\geq u\left[\sum_{i=1}^{n} x_i f(x_i)\right] \\ &= u(E[x_i]) \end{aligned}$$

Jadi, $E[u(x)] \geq u(E[x_i])$.

Jadi, terbukti bahwa jika $u(x)$ adalah sembarang fungsi dari variabel random X dan $u(x)$ adalah fungsi konveks, maka berlaku:

$$E\,[u(x)] \geq u\,(E[X])$$

Lampiran 5 Iterasi algoritma EM akan meningkatkan nilai $l(\boldsymbol{y}, \theta)$, pada setiap iterasinya

Bukti:

Misalkan $q(x)$ adalah suatu pdf sembarang dari X, dimana $\sum_x q(x) = 1$. Maka persamaan (2.6) dapat dituliskan sebagai:

$$l(\boldsymbol{y}, \theta) = \log [f(\boldsymbol{y};\theta)]$$
$$= \sum_x q(x) \log [f(\boldsymbol{y};\theta)]$$
$$= \sum_x q(x) \log [\frac{f(\boldsymbol{y},x;\theta)}{f(x|\boldsymbol{y},\theta)} \cdot \frac{q(x)}{q(x)}]$$
$$= \sum_x q(x) \left(\log [f(\boldsymbol{y},x;\theta)] - \log[q(x)] + \log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right]\right)$$
$$= \sum_x q(x) \log [f(x,\boldsymbol{y};\theta)] - \sum_x q(x) \log[q(x)] + \sum_x q(x) \log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right]$$

definisikan:

$$Q(q||f_{joint}) = \sum_x q(x) \log [f(x,\boldsymbol{y};\theta)]$$
$$H(q||q) = -\sum_x q(x) \log[q(x)]$$
$$KL(q||f_{post}) = \sum_x q(x) \log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right]$$

Jadi, persamaan (2.6) dapat dituliskan kembali menjadi:

$$l(\boldsymbol{y},\theta) = Q(q||f_{joint}) + H(q||q) + KL(q||f_{post})$$

Dapat dibuktikan bahwa KL bersifat:

a) $KL(q||f_{post}) \geq 0 \quad \forall\, q$

b) $\exists\, f$ sedemikian sehingga $KL(f||f_{post}) = 0$

Bukti:

a) Akan dibuktikan $KL(q||f_{post}) \geq 0$

$$KL(q||f_{post}) = \sum_x q(x) \log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right]$$
$$= \sum_x q(x) \left[-\log\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right]$$

(lanjutan)

Berdasarkan pertidaksamaan Jensen, untuk f suatu fungsi konveks $E\,[f\,(x)] \geq f\,(E[x])$, dan karena $-\log\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]$ adalah suatu fungsi konveks, maka berlaku:

$$E\left[-\log\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right] \geq -\log\left[E\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right]$$

$$\sum_{x} q(x)\left[-\log\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right] \geq -\log\left[\sum_{x} q(x)\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right]$$

$KL\,(q||f)$ dapat dituliskan sebagai berikut:

$$KL\,(q||f) = \sum_{x} q(x)\log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right] = \sum_{x} q(x)\left[-\log\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right]$$

$$\geq -\log\left[\sum_{x} q(x)\left[\frac{f(x|\boldsymbol{y},\theta)}{q(x)}\right]\right]$$

$$= -\log\left[\sum_{x} f(x|\boldsymbol{y},\theta)\right] = -\log[1]$$

$$= 0$$

Jadi, $KL\,(q||f) \geq 0,\ \forall$ q.

b) Akan dibuktikan $\exists\, f$ sedemikian sehingga $KL\left(f||f_{post}\right) = 0$

Misal pilih $q(x) = f(x|\boldsymbol{y},\theta)$, maka $\sum_{x} f(x|\boldsymbol{y},\theta) = 1$, dan:

$$\log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right] = log\left[\frac{f(x|\boldsymbol{y},\theta)}{f(x|\boldsymbol{y},\theta)}\right]$$

$$= \log[1]$$

$$= 0$$

Maka $KL\left(q||f_{post}\right)$ dapat dituliskan sebagai berikut:

$$KL\left(q||f_{post}\right) = \sum_{x} q(x)\log\left[\frac{q(x)}{f(x|\boldsymbol{y},\theta)}\right]$$

$$= \sum_{x} f(x|\boldsymbol{y},\theta)\,.\,0$$

$$= 0$$

Jadi, untuk $q(x) = f(x|\boldsymbol{y},\theta)$, nilai $KL\left(q||f_{post}\right) = 0$.

Misal $q(x) = f\left(x\middle|\boldsymbol{y},\hat{\theta}_{t-1}\right)$, maka $KL = 0$. Kemudian, sebut:

$$Q_{\hat{\theta}_{t-1}} = \sum_{x} f\left(x\middle|\boldsymbol{y},\hat{\theta}_{t-1}\right)\log\left[f(x,\boldsymbol{y},\hat{\theta}_{t})\right]$$

(lanjutan)

$$H_{\hat{\theta}_{t-1}} = -\sum_x f\big(x\big|\boldsymbol{y},\hat{\theta}_{t-1}\big)\log\big[f\big(x\big|\boldsymbol{y},\hat{\theta}_{t-1}\big)\big]$$

dengan mensubstitusikan $Q_{\hat{\theta}_{t-1}}$ dan $H_{\hat{\theta}_{t-1}}$ ke dalam persamaan (2.6), maka berlaku:

$$\begin{aligned} l(\boldsymbol{y},\theta_t) &= Q_{\hat{\theta}_{t-1}} + H_{\hat{\theta}_{t-1}} + 0 \\ &= Q_{\hat{\theta}_{t-1}} + H_{\hat{\theta}_{t-1}} \end{aligned}$$

dimana $H_{\hat{\theta}_{t-1}}$ tidak bergantung pada θ_t.

Misalkan $\hat{\theta}_t$ adalah taksiran θ yang memaksimumkan $Q_{\theta_{t-1}}$, maka:

$$\begin{aligned} l(\boldsymbol{y},\theta_t) &= Q_{\hat{\theta}_{t-1}} + H_{\hat{\theta}_{t-1}} \\ &\le Q_{\hat{\theta}_t} + H_{\hat{\theta}_t} \\ &= l(\boldsymbol{y},\theta_{t+1}) - KL_{\theta_t} \\ &= l(\boldsymbol{y},\theta_{t+1}) - 0 \\ &= l(\boldsymbol{y},\theta_{t+1}) \end{aligned}$$

Jadi, $l(\boldsymbol{y},\theta_t) \le l(\boldsymbol{y},\theta_{t+1})$.

Jadi, terbukti bahwa dengan menggunakan algoritma EM akan didapatkan taksiran θ yang memaksimumkan fungsi likelihood dari $\boldsymbol{Y}$.

Lampiran 6 Teorema Bayes

Misalkan B_1, B_2,…, B_T merupakan partisi dari ruang sampel S, $B_i \neq \emptyset$, i = 1, 2, …, T yang bersifat:

- $B_1 \cup B_2 \cup \ldots \cup B_T = S$
- $B_i \cap B_j = \emptyset$, untuk $i \neq j$

Misalkan A adalah sembarang kejadian yang merupakan himpunan bagian S, yang besifat $P(A) \neq 0$. Kejadian A dapat dipandang sebagai gabungan kejadian-kejadian $B_1 \cap A$, $B_2 \cap A$, …, $B_T \cap A$ yang saling terpisah satu sama lain.

$$A = (B_1 \cap A) \cup (B_2 \cap A) \cup \ldots \cup (B_T \cap A)$$

Karena $B_1 \cap A$, $B_2 \cap A$, …, $B_T \cap A$ merupakan himpunan-himpunan yang saling lepas, maka probabilitas kejadian A dapat dituliskan sebagai berikut:

$$\begin{aligned} P(A) &= P[(B_1 \cap A) \cup (B_2 \cap A) \cup \ldots \cup (B_T \cap A)] \\ &= P(B_1 \cap A) + P(B_2 \cap A) + \ldots + P(B_T \cap A) \\ &= P(B_1)\, P(A|B_1) + P(B_2)\, P(A|B_2) + \ldots + P(B_T)\, P(A|B_T) \\ &= \textstyle\sum_{i=1}^{T} P(B_i) P(A|B_i) \end{aligned}$$

Kemudian, menurut definisi peluang bersyarat diketahui bahwa:

$$\begin{aligned} P(B_r|A) &= \frac{P(B_r \cap A)}{P(A)} \\ &= \frac{P(B_r)P(A|B_r)}{P(A)} \\ &= \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{T} P(B_i)P(A|B_i)} \end{aligned}$$

Jadi, dapat dinyatakan jika terdapat kejadian-kejadian B_1, B_2, …, B_T merupakan partisi dari ruang sampel S, $B_i \neq \emptyset$, i = 1, 2, …, T, maka untuk sembarang kejadian A, dengan $P(A) \neq 0$, berlaku:

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

untuk r = 1, 2, …, T.

Lampiran 7 *Source code*

```
clc; clear;
P0{1} = 1; m0{1} = [17;23;17]; sigma0{1} = [7;6;11];
P0{2} = [0.5 0.5]; m0{2} = [17 19;28 17;26 11]; sigma0{2} = [7 6;
7 5; 5 3];
P0{3} = [0.4 0.2 0.4]; m0{3} = [14 17 22; 27 28 19; 14 26 20];
sigma0{3} = [7 5 4; 6 3 4; 4 8 3];
P0{4} = [0.3 0.2 0.2 0.3]; m0{4} = [12 23 21 18; 25 28 19 15; 14
17 23 13]; sigma0{4} = [3 3 5 3; 5 4 8 5; 4 2 4 3];
P0{5} = [0.2 0.3 0.2 0.2 0.1]; m0{5} = [18 16 19 21 11; 15 24 29
19 19;13 12 21 23 20]; sigma0{5} = [2 3 5 3 2; 6 4 2 2 5; 2 2 8 3
7];

y = xlsread('data.xls');
[K b] = size(y);

syms 'P1' 'P2' 'P3' 'P4' 'P5';
syms 'm11' 'm21' 'm31';
syms 'm12' 'm22' 'm32';
syms 'm13' 'm23' 'm33';
syms 'm14' 'm24' 'm34';
syms 'm15' 'm25' 'm35';
syms 's11' 's21' 's31';
syms 's12' 's22' 's32';
syms 's13' 's23' 's33';
syms 's14' 's24' 's34';
syms 's15' 's25' 's35';

sP = [P1 P2 P3 P4 P5];
sm = [m11 m12 m13 m14 m15; m21 m22 m23 m24 m25; m31 m32 m33 m34
m35];
ss = [s11 s12 s13 s14 s15; s21 s22 s23 s24 s25; s31 s32 s33 s34
s35];

%M - Step
for c = 1 : 5
    disp(c);
```

(lanjutan)

```
    flag = true;
    nP = [];    nm = [];    ns = [];
    while(flag)
        P = P0{c}; m = m0{c}; sigma = sigma0{c};
        I1 = I(c, K, y, P, m, sigma);
        for j = 1 : c
            nP(j) = solve(diff(I1, sP(j)), sP(j));
            for l = 1 : 3
                nm(l,j) = solve(diff(I1, sm(l,j)), sm(l,j));
                ns(l,j) = subs(solve(diff(I1, ss(l,j)), ss(l,j)),
sm(l,j), nm(l,j));
            end
        end
        if norm(nP - P0{c}) <= 10^-4
            flag = false;
        end
        P0{c} = nP;
        m0{c} = nm;
        sigma0{c} = ns;
    end
end

display('selesai');

function hasil = I(c, K, y, P, m, sigma)

syms 'P1' 'P2' 'P3' 'P4' 'P5';
syms 'm11' 'm21' 'm31';
syms 'm12' 'm22' 'm32';
syms 'm13' 'm23' 'm33';
syms 'm14' 'm24' 'm34';
syms 'm15' 'm25' 'm35';
syms 's11' 's21' 's31';
syms 's12' 's22' 's32';
syms 's13' 's23' 's33';
syms 's14' 's24' 's34';
```

(lanjutan)

```
syms 's15' 's25' 's35';

sP = [P1 P2 P3 P4 P5];
sm = [m11 m12 m13 m14 m15; m21 m22 m23 m24 m25; m31 m32 m33 m34
m35];
ss = [s11 s12 s13 s14 s15; s21 s22 s23 s24 s25; s31 s32 s33 s34
s35];

hasil = 0;
for x = 1 : c
    for k = 1 : K
        nT = T(c, x, k, y, P, m, sigma);
        h1 = nT*log(sP(x));
        h2 = nT*(-sum((y(k,:)'-sm(:,x)).^2 ./ (2*ss(:,x))));
        h3 = nT*(3/2 * log(2*pi) - 1/2 * sum(log(ss(:,x))));
        hasil = hasil + h1 + h2 + h3;
    end
end
hasil = hasil - K*(sum(sP) - 1);

function hasil = T(c, x, k, y, P, m, sigma)

h1 = P(x) * f(y(k,:), m(:,x), sigma(:,x));
h2 = 0;
for x1 = 1 : c
    h2 = h2 + (P(x1) * f(y(k,:), m(:,x1), sigma(:,x1)));
end
hasil = h1/h2;

function hasil = f(yk, mx, sigmax)

h1 = exp(-sum((yk - mx').^2));
h2 = prod(sqrt(2*pi*sigmax.^2));
hasil = h1/h2;
```

Lampiran 8 Data respon mahasiswa baru Departemen Matematika FMIPA UI angkatan 2010

No	Tingkat Ketaktergantungan	Skor Tanggung Jawab	Tingkat Ketenangan
1	16	28	17
2	17	22	18
3	17	22	18
4	19	23	14
5	12	26	21
6	21	23	18
7	19	21	11
8	19	23	18
9	18	26	23
10	20	22	12
11	19	23	17
12	22	19	12
13	13	25	25
14	22	19	20
15	15	22	25
16	18	15	13
17	19	25	17
18	20	22	19
19	17	23	14
20	16	25	15
21	17	19	16
22	20	22	14
23	17	20	20
24	17	22	20
25	20	23	17
26	17	22	18
27	20	23	15
28	21	25	20
29	21	23	20
30	17	19	19
31	21	21	16
32	16	22	17
33	18	24	11
34	19	18	15
35	20	25	16

36	16	20	14
37	19	22	18
38	23	22	18
39	22	24	19
40	17	24	15
41	20	20	20
42	18	24	17
43	19	25	20
44	12	25	14
45	20	20	11
46	16	25	19
47	14	20	20
48	19	23	12
49	13	25	19
50	16	24	18
51	21	26	22
52	13	21	15
53	15	26	20
54	16	27	25
55	16	23	20
56	19	23	18
57	19	17	11
58	17	24	16
59	15	21	16
60	18	23	17
61	13	22	18
62	18	19	21
63	17	22	15
64	16	24	23
65	20	20	21
66	13	22	17
67	14	27	14
68	18	21	14
69	13	23	16