

BAB III
ESTIMASI PARAMETER PADA MODEL REGRESI
ORDINAL DUA LEVEL

Model untuk variabel respon berskala ordinal sangat diperlukan pada banyak bidang penelitian, karena seringkali subjek penelitian diklasifikasikan atau mungkin memberikan respon berskala ordinal. Sebagai contoh dalam bidang pendidikan, siswa mungkin memberikan respon untuk suatu item kuesioner: sangat tidak setuju, tidak setuju, netral, setuju, sangat setuju. Contoh lainnya dalam studi biomedis, subjek pasien diklasifikasikan terhadap suatu penyakit dalam kategori: parah, tidak terlalu parah, tidak menunjukkan gejala penyakit. Pada individu atau subjek amatan, seringkali subjek-subjek yang diamati bersarang dalam cluster-cluster (seperti siswa dalam sekolah, pasien dalam klinik, dll.), atau subjek diamati berulang kali pada periode waktu tertentu.

Pada bab 2 telah dibahas mengenai model regresi dua level untuk data hirarki atau bersarang dengan variabel respon kontinu y . Dalam bab ini akan dibahas mengenai pengembangan model regresi dua level untuk variabel respon berskala ordinal, biasa dikenal dengan model regresi ordinal dua level, serta membahas mengenai estimasi parameter pada model tersebut dengan menggunakan metode estimasi *maximum marginal likelihood*.

3.1 MODEL REGRESI ORDINAL DUA LEVEL

3.1.1 Konsep Threshold

Dalam menjelaskan model regresi dua level untuk variabel respon berskala ordinal, digunakan konsep *threshold*, yaitu diasumsikan bahwa variabel respon ordinal yang diketahui dibentuk dari suatu variabel laten kontinu yang tidak diketahui nilainya dengan memberikan nilai batas kategorik (*threshold*) tertentu.

Misalkan Y merupakan variabel ordinal dengan C kategori terurut, dimana $c = 1, \dots, C$. Maka, terdapat $C+1$ nilai *threshold* $\gamma_0, \gamma_1, \gamma_2, \dots, \gamma_C$, dengan $\gamma_0 = -\infty$, $\gamma_C = +\infty$, dan $\gamma_0 \leq \gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_C$. Nilai γ_c , $c = 1, \dots, C$, didefinisikan sebagai batas antara interval yang bersesuaian dengan kategori $c-1$ dan c . Untuk identifikasi, diasumsikan $\gamma_1 = 0$.

Berdasarkan konsep *threshold*, variabel respon laten kontinu (y) dan variabel respon ordinal (Y) didefinisikan sebagai berikut:

$$Y = \begin{cases} 1, & \text{jika } -\infty < y \leq \gamma_1 \\ 2, & \text{jika } \gamma_1 < y \leq \gamma_2 \\ \vdots & \\ C, & \text{jika } \gamma_{C-1} < y < \infty \end{cases} \quad (3.1.1.1)$$

Dengan perkataan lain, variabel ordinal Y akan berada pada kategori c ($Y = c$), $c = 1, 2, \dots, C$, jika y lebih besar dari nilai threshold γ_{c-1} , tapi tidak melebihi nilai threshold γ_c .

3.1.2 Model Regresi Ordinal Dua Level melalui Pendekatan Variabel

Laten Kontinu

Berdasarkan pendefinisian hubungan antara variabel respon ordinal Y dan variabel laten kontinu y pada subbab 3.1.1 di atas, selanjutnya ingin dilihat hubungan antara variabel respon laten kontinu y dengan variabel-variabel penjelas di setiap level. Misalkan notasi $i = 1, 2, \dots, N$ menyatakan unit-unit di level dua dan $k = 1, 2, \dots, n_i$ menyatakan unit-unit di level satu yang bersarang dalam tiap unit level-2. Secara umum, model regresi dua level untuk variabel respon laten kontinu y_{ik} dapat ditulis sebagai berikut:

- Untuk model level-1:

$$y_{ik} = \mathbf{x}'_{(1)ik} \mathbf{b}_i + \mathbf{w}'_{(1)ik} \boldsymbol{\alpha}_{(1)} + \varepsilon_{ik} \quad (3.1.2.2)$$

$$= \begin{pmatrix} 1 & x_{(1)1ik} & \cdots & x_{(1)Rik} \end{pmatrix} \begin{pmatrix} b_{0i} \\ b_{0i} \\ \vdots \\ b_{Ri} \end{pmatrix} + \begin{pmatrix} w_{(1)1ik} & w_{(1)2ik} & \cdots & w_{(1)Sik} \end{pmatrix} \begin{pmatrix} \alpha_{(1)1} \\ \alpha_{(1)2} \\ \vdots \\ \alpha_{(1)S} \end{pmatrix} + \varepsilon_{ik}$$

dengan,

y_{ik} = variabel respon laten kontinu untuk unit ke- k level-1 dalam unit ke- i level-2, $i = 1, \dots, N$, $k = 1, \dots, n_i$.

\mathbf{b}_i = vektor ukuran $(R+1) \times 1$ dari koefisien *random* yang tidak diketahui untuk unit ke- i level 2.

$\alpha_{(1)}$ = vektor ukuran $S \times 1$ yang berisi koefisien dari variabel penjelas level satu, diasumsikan mempunyai nilai yang *fixed* untuk setiap unit level dua.

$\mathbf{x}'_{(1)ik}$ = vektor ukuran $1 \times (R+1)$, berisi variabel penjelas *fixed* level satu yang mempunyai efek yang berbeda untuk setiap unit ke- i level dua.

$\mathbf{w}'_{(1)ik}$ = vektor ukuran $1 \times S$, berisi variabel penjelas *fixed* level satu yang mempunyai efek yang *fixed* untuk setiap unit level dua.

ε_{ik} = *error* dari model level-1 ke- ik , $i = 1, 2, \dots, N$; $k = 1, 2, \dots, n_i$.

- Untuk model level-2:

$$\mathbf{b}_i = \boldsymbol{\mu} + \mathbf{w}'_{(2)i} \boldsymbol{\alpha}_{(2)} + \boldsymbol{\delta}_i \quad (3.1.2.3)$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \\ \vdots \\ b_{Ri} \end{pmatrix} = \begin{pmatrix} \mu_0 \\ \mu_1 \\ \vdots \\ \mu_R \end{pmatrix} + \begin{pmatrix} w_{(2)1i} & \dots & w_{(2)Qi} \end{pmatrix} \begin{pmatrix} \alpha_{(2)r1} \\ \vdots \\ \alpha_{(2)rQ} \end{pmatrix} + \begin{pmatrix} \delta_{0i} \\ \delta_{1i} \\ \vdots \\ \delta_{Ri} \end{pmatrix}, \quad r = 0, 1, \dots, R$$

dengan,

$\boldsymbol{\mu}$ = vektor ukuran $(R+1) \times 1$ berisi *intercept* level-2

$\boldsymbol{\alpha}_{(2)}$ = vektor ukuran $Q \times 1$ berisi koefisien dari variabel penjelas level dua diasumsikan bernilai *fixed*.

$\mathbf{w}'_{(2)i}$ = vektor ukuran $1 \times Q$ berisi *fixed* kovariat atau variabel penjelas *fixed* level-2 .

δ_i = vektor ukuran 1 x (R+1) berisi efek *random*.

Asumsi: - ε_{ik} berdistribusi iid $N(0, \sigma^2)$

- δ_i berdistribusi multivariat normal $N_{R+1}(\mathbf{0}, \Sigma_\delta)$

- Vektor δ_i dan residual level-1 ε_{ik} saling bebas

Model (3.1.2.2) dan (3.1.2.3) dapat digabung menjadi *combined model*

sebagai berikut:

$$\begin{aligned}
 y_{ik} &= \mathbf{x}'_{(1)ik} (\boldsymbol{\mu} + \mathbf{w}'_{(2)ik} \boldsymbol{\alpha}_{(2)} + \boldsymbol{\delta}_i) + \mathbf{w}'_{(1)ik} \boldsymbol{\alpha}_{(1)} + \varepsilon_{ik} \\
 &= \begin{pmatrix} 1 & \cdots & \mathbf{x}_{(1)Rik} \end{pmatrix} \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_R \end{pmatrix} + \begin{pmatrix} w_{(2)1i} & \cdots & w_{(2)Qi} \end{pmatrix} \begin{pmatrix} \alpha_{(2)r1} \\ \vdots \\ \alpha_{(2)rQ} \end{pmatrix} + \begin{pmatrix} \delta_{0i} \\ \vdots \\ \delta_{Ri} \end{pmatrix} + \begin{pmatrix} w_{(1)1ik} & \cdots & w_{(1)Sik} \end{pmatrix} \begin{pmatrix} \alpha_{(1)1} \\ \vdots \\ \alpha_{(1)S} \end{pmatrix} + \varepsilon_{ik} \\
 &= \begin{pmatrix} 1 & \cdots & \mathbf{x}_{(1)Rik} \end{pmatrix} \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_R \end{pmatrix} + \left[\mathbf{1} \begin{pmatrix} w_{(2)1i} & \cdots & w_{(2)Qi} \end{pmatrix} \cdots \mathbf{x}_{(1)Rik} \begin{pmatrix} w_{(2)1i} & \cdots & w_{(2)Qi} \end{pmatrix} \right] \begin{pmatrix} \alpha_{(2)r1} \\ \vdots \\ \alpha_{(2)rQ} \end{pmatrix} \\
 &\quad + \begin{pmatrix} 1 & \cdots & \mathbf{x}_{(1)Rik} \end{pmatrix} \begin{pmatrix} \delta_{0i} \\ \vdots \\ \delta_{Ri} \end{pmatrix} + \begin{pmatrix} w_{(1)1ik} & \cdots & w_{(1)Sik} \end{pmatrix} \begin{pmatrix} \alpha_{(1)1} \\ \vdots \\ \alpha_{(1)S} \end{pmatrix} + \varepsilon_{ik} \\
 &= \begin{pmatrix} 1 & \cdots & \mathbf{x}_{(1)Rik} \end{pmatrix} \begin{pmatrix} \mu_0 \\ \vdots \\ \mu_R \end{pmatrix} + \left[\mathbf{1} \cdot \mathbf{w}'_{(2)i} \cdots \mathbf{x}_{(1)Rik} \cdot \mathbf{w}'_{(2)i} \right] \begin{pmatrix} \alpha_{(2)r1} \\ \vdots \\ \alpha_{(2)rQ} \end{pmatrix} + \begin{pmatrix} 1 & \cdots & \mathbf{x}_{(1)Rik} \end{pmatrix} \begin{pmatrix} \delta_{0i} \\ \vdots \\ \delta_{Ri} \end{pmatrix} \\
 &\quad + \begin{pmatrix} w_{(1)1ik} & \cdots & w_{(1)Sik} \end{pmatrix} \begin{pmatrix} \alpha_{(1)1} \\ \vdots \\ \alpha_{(1)S} \end{pmatrix} + \varepsilon_{ik}
 \end{aligned}$$

$$\begin{aligned}
y_{ik} &= \mathbf{x}'_{(1)ik} \boldsymbol{\mu} + (\mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i}) \boldsymbol{\alpha}_{(2)} + \mathbf{x}'_{(1)ik} \boldsymbol{\delta}_i + \mathbf{w}'_{(1)ik} \boldsymbol{\alpha}_{(1)} + \varepsilon_{ik} \\
&= \mathbf{x}'_{(1)ik} (\boldsymbol{\mu} + \boldsymbol{\delta}_i) + (\mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i}) \boldsymbol{\alpha}_{(2)} + \mathbf{w}'_{(1)ik} \boldsymbol{\alpha}_{(1)} + \varepsilon_{ik} \\
&= \mathbf{x}'_{(1)ik} (\boldsymbol{\mu} + \boldsymbol{\delta}_i) + \begin{bmatrix} \mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i} & \vdots & \mathbf{w}'_{(1)ik} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{(2)} \\ \vdots \\ \boldsymbol{\alpha}_{(1)} \end{bmatrix} + \varepsilon_{ik}
\end{aligned} \tag{3.1.2.4}$$

dengan,

$\mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i}$ merupakan *kroncker product* dari vektor $\mathbf{x}'_{(1)ik}$ dan $\mathbf{w}'_{(2)i}$.

$\begin{bmatrix} \mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i} \\ \vdots \\ \mathbf{w}'_{(1)ik} \end{bmatrix}$ merupakan partisi vektor $\mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i}$ dan $\mathbf{w}'_{(1)ik}$.

$\begin{bmatrix} \boldsymbol{\alpha}_{(2)} \\ \vdots \\ \boldsymbol{\alpha}_{(1)} \end{bmatrix}$ merupakan partisi vektor $\boldsymbol{\alpha}_{(2)}$ dan $\boldsymbol{\alpha}_{(1)}$.

Untuk memperingkas penulisan, model regresi ordinal dua level di atas juga dapat dinyatakan dalam:

$$y_{ik} = \mathbf{x}'_{ik} \boldsymbol{\beta}_i + \mathbf{w}'_{ik} \boldsymbol{\alpha} + \varepsilon_{ik} \tag{3.1.2.5}$$

$$= (1 \quad x_{1ik} \quad \cdots \quad x_{Rik}) \begin{pmatrix} \beta_{0i} \\ \beta_{1i} \\ \vdots \\ \beta_{Ri} \end{pmatrix} + (w_{1ik} \quad w_{2ik} \quad \cdots \quad w_{Pik}) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_P \end{pmatrix} + \varepsilon_{ik}$$

dengan, \mathbf{x}'_{ik} = *design vector* untuk $R + 1$ efek *random*

\mathbf{w}'_{ik} = vektor ukuran $1 \times P$ berisi *fixed* kovariat.

$\boldsymbol{\beta}_i$ = vektor ukuran $(R+1) \times 1$ dari efek *random* yang tidak diketahui

untuk unit ke- i level-2, $i = 1, \dots, N$, diasumsikan berdistribusi

$$N_{R+1}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\boldsymbol{\delta}})$$

α = vektor ukuran $P \times 1$ dari *fixed* koefisien regresi yang tidak diketahui

ε_{ik} = residual model level-1 ke-ik, $i = 1, 2, \dots, N$; $k = 1, 2, \dots, n_i$, diasumsikan berdistribusi iid $N(0, \sigma^2)$

Hubungan antara model (3.1.2.4) dan (3.1.2.5) adalah :

$$\mathbf{x}'_{ik} = \mathbf{x}'_{(1)ik}$$

$$\boldsymbol{\beta}_i = \boldsymbol{\mu} + \boldsymbol{\delta}_i$$

$$\mathbf{w}'_{ik} = [\mathbf{x}'_{(1)ik} \otimes \mathbf{w}'_{(2)i} \vdots \mathbf{w}'_{(1)ik}]$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}'_{(2)} \\ \vdots \\ \boldsymbol{\alpha}'_{(1)} \end{bmatrix} = [\boldsymbol{\alpha}'_{(2)} \vdots \boldsymbol{\alpha}'_{(1)}]$$

Dari model (3.1.1.1) dan (3.1.2.5), maka dapat dicari probabilitas bahwa variabel respon kategorik Y_{ik} akan berada pada kategori c ($c = 1, \dots, C$), bersyarat pada $\boldsymbol{\beta}$ dan $\boldsymbol{\alpha}$, adalah sebagai berikut:

$$\begin{aligned} \Pr(Y_{ik} = c | \boldsymbol{\beta}, \boldsymbol{\alpha}) &= \Pr(\gamma_{c-1} < y_{ik} < \gamma_c) \\ &= \Pr(\gamma_{c-1} < \mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha} + \varepsilon_{ik} < \gamma_c) \\ &= \Pr(\gamma_{c-1} - (\mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha}) < \varepsilon_{ik} < \gamma_c - (\mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha})) \\ &= \Pr\left(\frac{\gamma_{c-1} - (\mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha})}{\sigma} < \frac{\varepsilon_{ik}}{\sigma} < \frac{\gamma_c - (\mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha})}{\sigma}\right) \\ &= \Pr\left(\frac{\varepsilon_{ik}}{\sigma} < \frac{\gamma_c - (\mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha})}{\sigma}\right) - \Pr\left(\frac{\varepsilon_{ik}}{\sigma} < \frac{\gamma_{c-1} - (\mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha})}{\sigma}\right) \end{aligned}$$

Karena ε_{ik} diasumsikan berdistribusi normal, maka pertidaksamaan di atas dapat dituliskan:

$$\Pr(Y_{ik} = c | \boldsymbol{\beta}, \boldsymbol{\alpha}) = \Phi\left[\frac{\gamma_c - z_{ik}}{\sigma}\right] - \Phi\left[\frac{\gamma_{c-1} - z_{ik}}{\sigma}\right] \quad (3.1.2.6)$$

$$\text{di mana, } z_{ik} = \mathbf{x}'_{ik}\boldsymbol{\beta}_i + \mathbf{w}'_{ik}\boldsymbol{\alpha} \quad (3.1.2.7)$$

$\Phi(\cdot)$ menyatakan fungsi distribusi kumulatif normal standar.

Tanpa menghilangkan sifat keumuman, jika diasumsikan bahwa $\sigma = 1$, maka model (3.1.2.6) menjadi:

$$\Pr(Y_{ik} = c | \boldsymbol{\beta}, \boldsymbol{\alpha}) = \Phi[\gamma_c - z_{ik}] - \Phi[\gamma_{c-1} - z_{ik}] \quad (3.1.2.8)$$

3.2 ESTIMASI PARAMETER PADA MODEL REGRESI ORDINAL DUA LEVEL

3.2.1 Metode Estimasi *Maximum Marginal Likelihood*

Misalkan \mathbf{Y}_i , $i = 1, \dots, N$ menyatakan vektor dari variabel respon ordinal dari unit ke- i level-2, untuk n_i unit level-1 yang bersarang di dalamnya.

$$\mathbf{Y}_i = [Y_{i1} \quad Y_{i2} \quad \dots \quad Y_{in_i}]$$

Karena variabel respon Y ordinal, maka variabel Y juga dapat dinyatakan sebagai barisan dari variabel *dummy* (McKelvey and Zavoina, 1975).

Dengan perkataan lain, dapat didefinisikan

$$d_{ikc} = \begin{cases} 1 & \text{jika } Y_{ik} = c \\ 0 & \text{jika } Y_{ik} \neq c \end{cases}$$

untuk $c = 1, 2, \dots, C$. Jadi, vektor \mathbf{Y}_i di atas dapat ditulis kembali menjadi:

$$\begin{aligned} \mathbf{Y}_i &= \left[d_{i11} \ d_{i12} \ \dots \ d_{i1C} \ ; \ d_{i21} \ d_{i22} \ \dots \ d_{i2C} \ ; \ \dots \ ; \ d_{in_1} \ d_{in_2} \ \dots \ d_{in_C} \right] \\ &= \left[\mathbf{d}_{i1c} \ \mathbf{d}_{i2c} \ \dots \ \mathbf{d}_{in_c} \right] \end{aligned}$$

Dari persamaan (3.1.2.8), maka probabilitas bersama dari n_i variabel respon level-1 dalam sembarang unit ke- i level-2, $i = 1, \dots, N$, bersyarat pada $\boldsymbol{\beta}$ dan $\boldsymbol{\alpha}$, adalah sebagai berikut:

$$\begin{aligned} \ell(\mathbf{Y}_i | \boldsymbol{\beta}, \boldsymbol{\alpha}) &= \prod_{k=1}^{n_i} \frac{\left(\sum_{c=1}^C d_{ikc} \right)!}{d_{ik1}! d_{ik2}! \dots d_{ikC}!} \left[\Pr(Y_{ik} = 1 | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ik1}} \cdot \left[\Pr(Y_{ik} = 2 | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ik2}} \cdot \dots \cdot \left[\Pr(Y_{ik} = C | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ikC}} \\ &= \prod_{k=1}^{n_i} \frac{1!}{d_{ik1}! d_{ik2}! \dots d_{ikC}!} \left[\Pr(Y_{ik} = 1 | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ik1}} \cdot \left[\Pr(Y_{ik} = 2 | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ik2}} \cdot \dots \cdot \left[\Pr(Y_{ik} = C | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ikC}} \\ &= \prod_{k=1}^{n_i} \left[\Pr(Y_{ik} = 1 | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ik1}} \cdot \left[\Pr(Y_{ik} = 2 | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ik2}} \cdot \dots \cdot \left[\Pr(Y_{ik} = C | \boldsymbol{\beta}, \boldsymbol{\alpha}) \right]^{d_{ikC}} \\ &= \prod_{k=1}^{n_i} \left[\Phi(\gamma_1 - z_{ik}) - \Phi(\gamma_0 - z_{ik}) \right]^{d_{ik1}} \cdot \left[\Phi(\gamma_2 - z_{ik}) - \Phi(\gamma_1 - z_{ik}) \right]^{d_{ik2}} \cdot \dots \cdot \left[\Phi(\gamma_C - z_{ik}) - \Phi(\gamma_{C-1} - z_{ik}) \right]^{d_{ikC}} \end{aligned}$$

Secara umum, persamaan di atas dapat ditulis:

$$\ell(\mathbf{Y}_i | \boldsymbol{\beta}, \boldsymbol{\alpha}) = \prod_{k=1}^{n_i} \prod_{c=1}^C \left[\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik}) \right]^{d_{ikc}} \quad (3.2.1.9)$$

Persamaan (3.2.1.9) merupakan fungsi *likelihood* bersyarat untuk sembarang unit level dua ke- i .

P.d.f *marginal* dari \mathbf{Y}_i dinyatakan sebagai berikut :

$$h(\mathbf{Y}_i) = \int_{\boldsymbol{\beta}} \ell(\mathbf{Y}_i | \boldsymbol{\beta}, \boldsymbol{\alpha}) g(\boldsymbol{\beta}) d\boldsymbol{\beta} \quad (3.2.1.10)$$

di mana, $g(\boldsymbol{\beta})$ merupakan p.d.f dari vektor $\boldsymbol{\beta}$.

Umumnya fungsi distribusi dari efek *random* diasumsikan merupakan fungsi distribusi normal standar (Johnson *and* Albert, 1999). Oleh karena itu, vektor $\boldsymbol{\beta}$ yang berdistribusi $N_{R+1}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\delta)$ akan distandarisasi menjadi vektor $\boldsymbol{\theta}$ yang berdistribusi multivariat normal standar $N_{R+1}(\mathbf{0}, \mathbf{I})$ dengan cara melakukan transformasi berikut:

$$\boldsymbol{\beta} = \mathbf{T}\boldsymbol{\theta} + \boldsymbol{\mu}$$

$$\mathbf{T}\boldsymbol{\theta} = \boldsymbol{\beta} - \boldsymbol{\mu}$$

$$\boldsymbol{\theta} = \mathbf{T}^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu})$$

dimana $\mathbf{T}\mathbf{T}' = \boldsymbol{\Sigma}_\delta$ adalah dekomposisi *Cholesky* dari $\boldsymbol{\Sigma}_\delta$.

$$\mathbf{T} = \begin{pmatrix} \sigma_{\delta_0} & 0 & 0 & 0 \\ \frac{\sigma_{\delta_1\delta_0}}{\sigma_{\delta_0}} & \sqrt{\frac{\sigma_{\delta_1}^2\sigma_{\delta_0}^2 - \sigma_{\delta_1\delta_0}^2}{\sigma_{\delta_0}^2}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\sigma_{\delta_R\delta_0}}{\sigma_{\delta_0}} & \frac{\sigma_{\delta_R\delta_1} - \left(\frac{\sigma_{\delta_R\delta_0}\sigma_{\delta_1\delta_0}}{\sigma_{\delta_0}^2}\right)}{\sqrt{\frac{(\sigma_{\delta_1}^2)(\sigma_{\delta_0}^2) - \sigma_{\delta_1\delta_0}^2}{\sigma_{\delta_0}^2}}} & \dots & \sqrt{\frac{(\sigma_{\delta_1}^2)(\sigma_{\delta_0}^2) - (\sigma_{\delta_R\delta_0}^2 + \sigma_{\delta_R\delta_1}^2 + \dots + \sigma_{\delta_R\delta_{R-1}}^2)}{\sigma_{\delta_0}^2}} \end{pmatrix}$$

Dapat ditunjukkan bahwa vektor $\boldsymbol{\theta}$ berdistribusi multivariat normal standar $N_{R+1}(\mathbf{0}, \mathbf{I})$. (Bukti dapat dilihat di lampiran 3)

Berdasarkan transformasi ini, maka persamaan (3.1.2.7) menjadi

$$z_{ik} = \mathbf{x}'_{ik} (\mathbf{T}\boldsymbol{\theta} + \boldsymbol{\mu}) + \mathbf{w}'_{ik} \boldsymbol{\alpha} \quad (3.2.1.11)$$

dan p.d.f *marginal* pada (3.2.1.10) menjadi:

$$\begin{aligned}
 h(\mathbf{Y}_i) &= \int_{\boldsymbol{\theta}} \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
 &= \int_{\boldsymbol{\theta}} \left(\prod_{k=1}^{n_i} \prod_{c=1}^C [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})]^{d_{ikc}} \right) g(\boldsymbol{\theta}) d\boldsymbol{\theta}
 \end{aligned} \tag{3.2.1.12}$$

di mana, $g(\boldsymbol{\theta}) =$ p.d.f multivariat normal standar.

Konsekuensi dari mentransformasi dari $\boldsymbol{\beta}$ ke $\boldsymbol{\theta}$ yaitu bahwa parameter dalam model yang akan diestimasi adalah faktor *cholesky* \mathbf{T} dan bukan matriks kovariansi $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$.

Misalkan $\nu(\mathbf{T})$ adalah vektor ukuran $(R+1)(R+2)/2 \times 1$ yang berisi elemen-elemen di diagonal utama dan elemen-elemen di bawah diagonal utama dari matriks segitiga bawah \mathbf{T} . Untuk mempermudah notasi, misalkan $\nu(\mathbf{T}) = [t_{00} \ t_{10} \ \dots \ t_{RR}]'$, dengan t_{RR} adalah elemen ke $(R+1)(R+2)/2$. Jadi, berdasarkan pendefinisian vektor $\nu(\mathbf{T})$ di atas, dan berdasarkan persamaan (3.2.1.11) dan (3.2.1.12), terlihat bahwa parameter-parameter dalam model yang akan diestimasi adalah γ_c ($c = 2, \dots, C-1$),

$$\boldsymbol{\mu} = [\mu_0 \ \mu_1 \ \dots \ \mu_R]', \quad \boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_P]', \quad \text{dan} \quad \nu(\mathbf{T}) = [t_{00} \ t_{10} \ \dots \ t_{RR}]'$$

Untuk mencari estimasi atau taksiran parameter-parameter tersebut, diperlukan suatu metode estimasi. Dalam tugas akhir ini, metode estimasi yang diperlukan adalah metode estimasi *Maximum Marginal Likelihood* (*MMLE*). Prinsip dari metode ini adalah mencari taksiran parameter γ_c , $\boldsymbol{\alpha}$, $\boldsymbol{\mu}$, dan $\nu(\mathbf{T})$ yang memaksimumkan fungsi *marginal likelihood*. Fungsi

marginal likelihood untuk keseluruhan N unit level dua diperoleh sebagai hasil kali dari p.d.f marginal setiap unit level dua, atau dapat ditulis:

$$\begin{aligned} L &= \prod_{i=1}^N h(\mathbf{Y}_i) \\ &= \prod_{i=1}^N \left[\int_{\theta} \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \right] \end{aligned} \quad (3.2.1.13)$$

Untuk mempermudah perhitungan dalam mendapatkan taksiran *maximum marginal likelihood* dari parameter γ_c , α , μ , dan $\nu(\mathbf{T})$ digunakan bentuk logaritma dari fungsi *marginal likelihood* pada persamaan (3.2.1.13), yaitu:

$$\begin{aligned} \log L &= \log \left(\prod_{i=1}^N h(\mathbf{Y}_i) \right) \\ &= \log(h(\mathbf{Y}_1) \cdot h(\mathbf{Y}_2) \cdots h(\mathbf{Y}_N)) \\ &= \log h(\mathbf{Y}_1) + \log h(\mathbf{Y}_2) + \cdots + \log h(\mathbf{Y}_N) \\ &= \sum_{i=1}^N \log h(\mathbf{Y}_i) \end{aligned} \quad (3.2.1.14)$$

Persamaan (3.2.1.14) disebut juga fungsi *marginal log-likelihood*.

Taksiran *maximum marginal likelihood* dari parameter γ_c , α , μ dan $\nu(\mathbf{T})$ diperoleh dari turunan parsial pertama fungsi *marginal log-likelihood* terhadap parameter γ_c , α , μ , dan $\nu(\mathbf{T})$, dan kemudian menyamakannya dengan nol.

Misalkan ,

$$\phi(\gamma_c - z_{ik}) = \frac{1}{\sqrt{2\pi}} e^{-(\gamma_c - z_{ik})^2/2}$$

dan,

$$\delta_{cc'} = \begin{cases} 1 & \text{jika } c = c' \\ 0 & \text{jika } c \neq c' \end{cases}$$

Maka,

$$\begin{aligned} \bullet \quad \frac{\partial}{\partial \gamma_{c'}} \Phi(\gamma_c - z_{ik}) &= \phi(\gamma_c - z_{ik}) \left[\frac{\partial}{\partial \gamma_{c'}} (\gamma_c - z_{ik}) \right] \\ &= \phi(\gamma_c - z_{ik}) \left[\frac{\partial}{\partial \gamma_{c'}} (\gamma_c - (\mathbf{x}'_{ik} (\mathbf{T}\boldsymbol{\theta} + \boldsymbol{\mu}) + \mathbf{w}'_{ik} \boldsymbol{\alpha})) \right] \\ &= \phi(\gamma_c - z_{ik}) \frac{\partial \gamma_c}{\partial \gamma_{c'}} \\ &= \phi(\gamma_c - z_{ik}) \delta_{cc'}, \quad c, c' = 2, \dots, C-1 \end{aligned} \quad (3.2.1.15)$$

$$\begin{aligned} \bullet \quad \frac{\partial}{\partial \mu_{r'}} \Phi(\gamma_c - z_{ik}) &= -\phi(\gamma_c - z_{ik}) \left[\frac{\partial}{\partial \mu_{r'}} (-z_{ik}) \right] \\ &= -\phi(\gamma_c - z_{ik}) \frac{\partial (\mathbf{x}'_{ik} (\mathbf{T}\boldsymbol{\theta} + \boldsymbol{\mu}) + \mathbf{w}'_{ik} \boldsymbol{\alpha})}{\partial \mu_{r'}} \\ &= -\phi(\gamma_c - z_{ik}) \frac{\partial (\mathbf{x}'_{ik} \boldsymbol{\mu})}{\partial \mu_{r'}} = -\phi(\gamma_c - z_{ik}) \frac{\partial \left(\sum_{r=0}^R x_{rik} \mu_r \right)}{\partial \mu_{r'}} \\ &= -\phi(\gamma_c - z_{ik}) \sum_{r=1}^R x_{jik} \frac{\partial \mu_r}{\partial \mu_{r'}} \text{, bernilai 1 jika } r = r' \\ &= -\phi(\gamma_c - z_{ik}) x_{r'ik}, \quad r' = 0, 1, \dots, R \end{aligned} \quad (3.2.1.16)$$

$$\begin{aligned}
\bullet \frac{\partial}{\partial \alpha_{p'}} \Phi(\gamma_c - z_{ik}) &= -\phi(\gamma_c - z_{ik}) \frac{\partial(\mathbf{x}'_{ik}(\mathbf{T}\boldsymbol{\theta} + \boldsymbol{\mu}) + \mathbf{w}'_{ik}\boldsymbol{\alpha})}{\partial \alpha_{p'}} \\
&= -\phi(\gamma_c - z_{ik}) \frac{\partial(\mathbf{w}'_{ik}\boldsymbol{\alpha})}{\partial \alpha_{p'}} = -\phi(\gamma_c - z_{ik}) \frac{\partial\left(\sum_{p=1}^P w_{pik} \alpha_p\right)}{\partial \alpha_{p'}} \\
&= -\phi(\gamma_c - z_{ik}) \sum_{p=1}^P w_{pik} \frac{\partial \alpha_p}{\partial \alpha_{p'}}, \quad \text{bernilai 1 jika } p = p' \\
&= -\phi(\gamma_c - z_{ik}) w_{p'ik}, \quad p' = 1, 2, \dots, P
\end{aligned} \tag{3.2.1.17}$$

$$\begin{aligned}
\bullet \frac{\partial}{\partial t_m} \Phi(\gamma_c - z_{ik}) &= -\phi(\gamma_c - z_{ik}) \frac{\partial(\mathbf{x}'_{ik} \mathbf{T}\boldsymbol{\theta})}{\partial t_m} \\
&= -\phi(\gamma_c - z_{ik}) \frac{\partial}{\partial t_m} \left(1 \quad x_{1ik} \quad \dots \quad x_{Rik} \right) \begin{pmatrix} t_{00} & 0 & \dots & 0 \\ t_{10} & t_{11} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ t_{R0} & t_{R1} & \dots & t_{RR} \end{pmatrix} \begin{pmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_R \end{pmatrix} \\
&= -\phi(\gamma_c - z_{ik}) \frac{\partial}{\partial t_m} (t_{00}\theta_0 + x_{1ik}(t_{10}\theta_0 + t_{11}\theta_1) + \dots + x_{Rik}(t_{R0}\theta_0 + t_{R1}\theta_1 + \dots + t_{RR}\theta_R)) \\
&= -\phi(\gamma_c - z_{ik}) \frac{\partial}{\partial t_m} (t_{00}\theta_0 + t_{10}x_{1ik}\theta_0 + \dots + t_{R0}x_{Rik}\theta_0 + t_{11}x_{1ik}\theta_1 + \dots + t_{R1}x_{1ik}\theta_1 + \dots + t_{RR}x_{Rik}\theta_R) \\
&= -\phi(\gamma_c - z_{ik}) \theta_h x_{rik}, \quad 0 \leq h \leq r \leq R
\end{aligned} \tag{3.2.1.18}$$

Selanjutnya, dari persamaan (3.2.1.9) dan dari sifat perkalian logaritma, diperoleh:

$$\begin{aligned}
\log \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) &= \log \left[\prod_{k=1}^{n_i} \prod_{c=1}^C \left[\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik}) \right]^{d_{ikc}} \right] \\
&= \log \left[\left(\left[\Phi(\gamma_1 - z_{i1}) - \Phi(\gamma_0 - z_{i1}) \right]^{d_{i11}} \cdots \left[\Phi(\gamma_1 - z_{in_i}) - \Phi(\gamma_0 - z_{in_i}) \right]^{d_{in_i1}} \right) \right. \\
&\quad \cdot \left(\left[\Phi(\gamma_2 - z_{i1}) - \Phi(\gamma_2 - z_{i1}) \right]^{d_{i12}} \cdots \left[\Phi(\gamma_2 - z_{in_i}) - \Phi(\gamma_2 - z_{in_i}) \right]^{d_{in_i2}} \right) \cdots \\
&\quad \left. \left(\left[\Phi(\gamma_C - z_{i1}) - \Phi(\gamma_{C-1} - z_{i1}) \right]^{d_{i1C}} \cdots \left[\Phi(\gamma_C - z_{in_i}) - \Phi(\gamma_{C-1} - z_{in_i}) \right]^{d_{in_iC}} \right) \right] \\
&= \left(d_{i11} \log \left[\Phi(\gamma_1 - z_{i1}) - \Phi(\gamma_0 - z_{i1}) \right] + \cdots + d_{in_i1} \left[\Phi(\gamma_1 - z_{in_i}) - \Phi(\gamma_0 - z_{in_i}) \right] \right) \\
&\quad + \left(d_{i12} \left[\log \Phi(\gamma_2 - z_{i1}) - \Phi(\gamma_1 - z_{i1}) \right] + \cdots + d_{in_i2} \left[\Phi(\gamma_2 - z_{in_i}) - \Phi(\gamma_1 - z_{in_i}) \right] \right) + \cdots \\
&\quad + \left(d_{i1C} \log \left[\Phi(\gamma_C - z_{i1}) - \Phi(\gamma_{C-1} - z_{i1}) \right] + \cdots + d_{in_iC} \left[\Phi(\gamma_C - z_{in_i}) - \Phi(\gamma_{C-1} - z_{in_i}) \right] \right) \\
&= \sum_{k=1}^{n_i} \sum_{C=1}^C d_{ikc} \log \left[\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik}) \right] \tag{3.2.1.19}
\end{aligned}$$

Persamaan (3.2.1.15), (3.2.1.16), (3.2.1.17), (3.2.1.18), dan (3.2.19) dapat digunakan untuk mencari turunan parsial pertama fungsi log L terhadap parameter-parameter $\gamma_{c'}$ ($c' = 2, \dots, C-1$), $\mu_{r'}$ ($r' = 0, 1, \dots, R$), $\alpha_{p'}$ ($p' = 1, \dots, P$), dan t_{rh} ($0 \leq h \leq r \leq R$). Jadi, berdasarkan persamaan-persamaan tersebut, diperoleh turunan parsial pertama fungsi *marginal log likelihood* terhadap parameter-parameter di atas adalah:

$$\begin{aligned}
\bullet \frac{\partial \log L}{\partial \gamma_{c'}} &= \sum_{i=1}^N \frac{\partial}{\partial \gamma_{c'}} (\log h(\mathbf{Y}_i)) \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \frac{\partial (h(\mathbf{Y}_i))}{\partial \gamma_{c'}} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\frac{\partial}{\partial \gamma_{c'}} [\ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha})] \right) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\frac{\partial}{\partial \gamma_{c'}} [\log \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha})] \right) \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\frac{\partial}{\partial \gamma_{c'}} \left[\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})] \right] \right) \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \frac{\partial}{\partial \gamma_{c'}} (\log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})]) \right) \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\frac{\partial}{\partial \gamma_{c'}} [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})]}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\frac{\partial}{\partial \gamma_{c'}} [\Phi(\gamma_c - z_{ik})] - \frac{\partial}{\partial \gamma_{c'}} [\Phi(\gamma_{c-1} - z_{ik})]}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\boldsymbol{\theta}} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_c - z_{ik}) \delta_{\alpha_{c'}} - \phi(\gamma_{c-1} - z_{ik}) \delta_{\alpha_{c-1, c'}}}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \boldsymbol{\theta}, \boldsymbol{\alpha}) g(\boldsymbol{\theta}) d\boldsymbol{\theta} = 0
\end{aligned}$$

(3.2.1.20)

Untuk $c' = 2, \dots, C-1$.

$$\begin{aligned}
\bullet \frac{\partial \log L}{\partial \mu_r} &= \sum_{i=1}^N \frac{\partial}{\partial \mu_r} (\log h(\mathbf{Y}_i)) \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \frac{\partial (h(\mathbf{Y}_i))}{\partial \mu_r} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \frac{\partial}{\partial \mu_r} [\ell(\mathbf{Y}_i | \theta, \alpha)] g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\frac{\partial}{\partial \mu_r} [\log \ell(\mathbf{Y}_i | \theta, \alpha)] \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\frac{\partial}{\partial \mu_r} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})] \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \frac{\partial}{\partial \mu_r} (\log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})]) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\frac{\partial}{\partial \mu_r} ([\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})])}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{-\phi(\gamma_c - z_{ik}) x_{rik} - (-\phi(\gamma_{c-1} - z_{ik}) x_{rik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_{c-1} - z_{ik}) - \phi(\gamma_c - z_{ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) x_{rik} \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta = 0
\end{aligned}$$

(3.2.1.21)

Untuk $r' = 0, 1, \dots, R$.

$$\begin{aligned}
\bullet \frac{\partial \log L}{\partial \alpha_{p'}} &= \sum_{i=1}^N \frac{\partial}{\partial \alpha_{p'}} (\log h(\mathbf{Y}_i)) \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \frac{\partial(h(\mathbf{Y}_i))}{\partial \alpha_{p'}} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \frac{\partial}{\partial \alpha_{p'}} [\ell(\mathbf{Y}_i | \theta, \alpha)] g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\frac{\partial}{\partial \alpha_{p'}} [\log \ell(\mathbf{Y}_i | \theta, \alpha)] \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\frac{\partial}{\partial \alpha_{p'}} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})] \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \frac{\partial}{\partial \alpha_{p'}} (\log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})]) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\partial}{\partial \alpha_{p'}} \left([\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})] \right) \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{-\phi(\gamma_c - z_{ik}) w_{p'ik} - (-\phi(\gamma_{c-1} - z_{ik}) w_{p'ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_{c-1} - z_{ik}) - \phi(\gamma_c - z_{ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) w_{p'ik} \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta = 0
\end{aligned}$$

(3.2.1.22)

Untuk $p' = 1, 2, \dots, P$.

$$\begin{aligned}
\bullet \frac{\partial \log L}{\partial t_h} &= \sum_{i=1}^N \frac{\partial}{\partial t_h} (\log h(\mathbf{Y}_i)) \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \frac{\partial(h(\mathbf{Y}_i))}{\partial t_h} \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \frac{\partial}{\partial t_h} [\ell(\mathbf{Y}_i | \theta, \alpha)] g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\frac{\partial}{\partial t_h} [\log \ell(\mathbf{Y}_i | \theta, \alpha)] \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\frac{\partial}{\partial t_h} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})] \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \frac{\partial}{\partial t_h} (\log [\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})]) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\frac{\partial}{\partial t_h} ([\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})])}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{(\phi(\gamma_{c-1} - z_{ik}) \theta_h x_{rik}) - (\phi(\gamma_c - z_{ik}) \theta_h x_{rik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta \\
&= \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \int_{\theta} \left(\sum_{k=1}^{n_i} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_{c-1} - z_{ik}) - \phi(\gamma_c - z_{ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \theta_h x_{rik} \right) \ell(\mathbf{Y}_i | \theta, \alpha) g(\theta) d\theta = 0
\end{aligned}$$

(3.2.1.23)

Untuk $0 \leq h \leq r \leq R$.

3.2.2 Metode Fisher Scoring

Persamaan-persamaan (3.2.1.20), (3.2.1.21), (3.2.1.22), dan (3.2.1.23) disebut persamaan *marginal likelihood*. Terlihat bahwa persamaan-persamaan *marginal likelihood* di atas tidak linier dalam $\gamma_{c'}$, $\alpha_{p'}$, $\mu_{r'}$, dan t_{rh} , ($c' = 2, \dots, C-1$; $r' = 0, 1, \dots, R$; $p' = 1, \dots, P$; $0 \leq h \leq r \leq R$), maka untuk mencari taksiran dari $\hat{\gamma}_{c'}$, $\hat{\alpha}_{p'}$, $\hat{\mu}_{r'}$, dan \hat{t}_{rh} digunakan suatu metode numerik. Metode numerik yang digunakan dalam skripsi ini adalah metode *Fisher Scoring*.

Dalam statistik, metode *Fisher Scoring* adalah bentuk lain dari metode *Newton-Raphson* yang digunakan untuk menyelesaikan persamaan *maximum likelihood* secara numerik.

Misalkan Θ adalah vektor dari parameter-parameter yang tidak diketahui, yaitu:

$$\Theta = (\mu_0, \dots, \mu_R \quad \alpha_1, \dots, \alpha_p \quad \gamma_2, \dots, \gamma_{C-1} \quad t_{00}, \dots, t_{RR})'$$

Metode *Newton-Raphson* adalah suatu metode untuk mencari estimasi parameter secara numerik menggunakan algoritma:

$$\hat{\Theta}_{m+1} = \hat{\Theta}_m - [H(\hat{\Theta}_m)]^{-1} U(\hat{\Theta}_m), \quad m = 0, 1, \dots$$

dengan:

$\hat{\Theta}_m$ = taksiran dari parameter Θ pada iterasi ke- m

$U(\hat{\Theta}^{(m)})$ = vektor turunan parsial pertama dari log L dihitung pada $\Theta = \hat{\Theta}^{(m)}$.

$H(\hat{\Theta}^{(m)})$ = matriks turunan parsial kedua dari log L dihitung pada $\Theta = \hat{\Theta}^{(m)}$.

Bukti dari algoritma ini diberikan di lampiran 4.

Karakteristik metode *Fisher scoring* adalah penggantian matriks Hessian $H(\hat{\Theta}_m)$ pada algoritma *Newton Raphson* dengan ekspektasi dari matriks Hessian, $E[H(\hat{\Theta}_m)]$.

Dalam model regresi ordinal dua level, penggunaan metode *Fisher Scoring* untuk mendapatkan taksiran $\hat{\gamma}_c$, $\hat{\alpha}_p$, $\hat{\mu}_r$, dan \hat{t}_{rh} dilakukan dengan tahapan-tahapan berikut:

1. Pilih taksiran awal dari parameter Θ , yaitu

$$\hat{\Theta}^{(0)} = \left(\hat{\mu}_0^{(0)}, \dots, \hat{\mu}_R^{(0)} \quad \hat{\alpha}_1^{(0)}, \dots, \hat{\alpha}_P^{(0)} \quad \hat{\gamma}_2^{(0)}, \dots, \hat{\gamma}_{C-1}^{(0)} \quad \hat{t}_{00}^{(0)}, \dots, \hat{t}_{RR}^{(0)} \right)', \text{ dan anggap}$$

parameter-parameter tersebut diketahui.

2. Tentukan taksiran dari Θ pada iterasi ke- $m+1$ ($m = 0, 1, \dots$), yaitu $\hat{\Theta}_{m+1}$ secara iteratif menggunakan formula:

$$\hat{\Theta}^{(m+1)} = \hat{\Theta}^{(m)} - \left[E[H(\hat{\Theta}^{(m)})] \right]^{-1} U(\hat{\Theta}^{(m)}), \quad m = 0, 1, \dots \quad (3.2.2.24)$$

di mana ekspektasi dari matriks turunan parsial kedua diberikan oleh:

$$E[H(\hat{\Theta}^{(m)})] = - \sum_{i=1}^N h^{-2}(\mathbf{Y}_i) \frac{\partial h(\mathbf{Y}_i)}{\partial \hat{\Theta}^{(m)}} \left(\frac{\partial h(\mathbf{Y}_i)}{\partial \hat{\Theta}^{(m)}} \right)'$$

3. Jika $\hat{\Theta}^{m+1} \approx \hat{\Theta}^{(m)}$ (misalkan $\|\hat{\Theta}^{(m+1)} - \hat{\Theta}^{(m)}\| < 10^{-4}$), maka hentikan proses iterasi dan kemudian ambil $\hat{\Theta}^{(m+1)}$ sebagai taksiran dari Θ .

Dari langkah kedua pada metode iteratif *Fisher scoring* di atas terlihat bahwa pada setiap iterasi, diperlukan penghitungan dari persamaan *marginal likelihood*. Persamaan *marginal likelihood* tersebut sulit untuk dihitung karena mengandung bentuk integral. Oleh karena itu, untuk mencari solusi dari persamaan *marginal likelihood* tersebut, integrasi secara numerik dapat digunakan pada ruang $\boldsymbol{\theta} = (\theta_0, \theta_2, \dots, \theta_R)$. Dalam tugas akhir ini, digunakan *Gauss-Hermite Quadrature* untuk mengaproksimasi integral di atas.

Dalam *Gauss-Hermite Quadrature*, integral diaproksimasi oleh suatu penjumlahan dari Q titik quadrature (*quadrature point*) tertentu untuk setiap dimensi dari integrasi sehingga untuk ruang $\boldsymbol{\theta}$ dengan $R+1$ dimensi, akan terdapat sumasi dari Q^{R+1} titik. Jadi, pada setiap iterasi algoritma *Fisher scoring* dan untuk setiap unit level-2, solusi persamaan *marginal likelihood* diperoleh dari sumasi Q^{R+1} *quadrature point*.

Untuk pdf *Univariate normal standard*, *quadrature point* dan bobotnya dinotasikan dengan B_q dan $A(B_q)$, $q = 1, \dots, Q$. Untuk pdf *multivariate normal standar* dengan $R+1$ -dimensi, $R+1$ -dimensi dari *quadrature point* dinotasikan dengan $\mathbf{B}'_q = (B_{q0}, B_{q1}, \dots, B_{qR})$, dan bobot yang bersesuaian dengan *quadrature point* tersebut diberikan oleh:

$$A(\mathbf{B}'_q) = \prod_{r=0}^R A(B_{qr}).$$

Probabilitas bersyarat $\ell(\mathbf{Y}_j | \boldsymbol{\theta}, \boldsymbol{\alpha})$ diperoleh dengan mensubstitusikan vektor dari efek random $\boldsymbol{\theta}$ dengan vektor dimensi $R+1$ dari *quadrature point*

\mathbf{B}_q , dan p.d.f *multivariate normal standard* $g(\boldsymbol{\theta})$ disubstitusi oleh bobot dari *quadrature point* $A(\mathbf{B}_q)$. Kemudian pdf *marginal* untuk setiap unit level-2 diaproksimasi dengan:

$$\begin{aligned} h(\mathbf{Y}_i) &\approx \sum_q^{Q^{R+1}} \ell(\mathbf{Y}_i | \mathbf{B}_q, \boldsymbol{\alpha}) A(\mathbf{B}_q) \\ &\approx \sum_{q=1}^Q \sum_{q=1}^Q \cdots \sum_{q=1}^Q \ell(\mathbf{Y}_i | \mathbf{B}_q, \boldsymbol{\alpha}) A(\mathbf{B}_{q0}) A(\mathbf{B}_{q1}) \cdots A(\mathbf{B}_{qR}) \end{aligned} \quad (3.2.2.25)$$

Dengan mensubstitusikan persamaan (3.2.2.25) ke dalam persamaan (3.2.1.20), (3.2.1.21), (3.2.1.22), dan (3.2.1.23) maka diperoleh.

- $\frac{\partial \log L}{\partial \gamma_{c'}} \approx \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \sum_{q=1}^{Q^{R+1}} \sum_{k=1}^{n_j} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_c - z_{ik}) \delta_{cc'} - \phi(\gamma_{c-1} - z_{ik}) \delta_{c-1,c'}}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) \ell(\mathbf{Y}_i | \mathbf{B}_q, \boldsymbol{\alpha}) A(\mathbf{B}_q)$
Untuk $c' = 2, \dots, C-1$
- $\frac{\partial \log L}{\partial \mu_{r'}} \approx \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \sum_{q=1}^{Q^{R+1}} \sum_{k=1}^{n_j} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_{c-1} - z_{ik}) - \phi(\gamma_c - z_{ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) x_{r'ik} \ell(\mathbf{Y}_i | \mathbf{B}_q, \boldsymbol{\alpha}) A(\mathbf{B}_q)$
Untuk $r' = 0, 1, \dots, R$
- $\frac{\partial \log L}{\partial \alpha_{p'}} \approx \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \sum_{q=1}^{Q^{R+1}} \sum_{k=1}^{n_j} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_{c-1} - z_{ik}) - \phi(\gamma_c - z_{ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) w_{p'ik} \ell(\mathbf{Y}_i | \mathbf{B}_q, \boldsymbol{\alpha}) A(\mathbf{B}_q)$
Untuk $p' = 1, 2, \dots, P$
- $\frac{\partial \log L}{\partial t_{rh}} \approx \sum_{i=1}^N h^{-1}(\mathbf{Y}_i) \sum_{q=1}^{Q^{R+1}} \sum_{k=1}^{n_j} \sum_{c=1}^C d_{ikc} \left(\frac{\phi(\gamma_{c-1} - z_{ik}) - \phi(\gamma_c - z_{ik})}{\Phi(\gamma_c - z_{ik}) - \Phi(\gamma_{c-1} - z_{ik})} \right) B_{qh} x_{nk} \ell(\mathbf{Y}_i | \mathbf{B}_q, \boldsymbol{\alpha}) A(\mathbf{B}_q)$
Untuk $0 \leq h \leq r \leq R$

Dari ke empat persamaan di atas dan berdasarkan persamaan (3.2.2.24), pada setiap iterasi metode *Fisher scoring*, penghitungan dari turunan parsial pertama dan ekspektasi dari matriks turunan parsial kedua

diproses dengan menjumlahkan unit-unit level dua dan titik-titik *quadrature*.

Dalam penjumlahan terhadap Q^{R+1} *quadrature point*, vektor efek random θ

disubstitusi oleh vektor dari *quadrature point* \mathbf{B}_q , dan p.d.f *multivariate*

normal standard $g(\theta)$ disubstitusi oleh bobot dari *quadrature point* $A(\mathbf{B}_q)$.

Dengan menjumlahkan unit-unit level dua dan titik-titik *quadrature* tersebut,

taksiran parameter pada iterasi ke- m , $m= 0,1,2,\dots$ diperoleh dari persamaan

(3.2.2.24), keseluruhan prosedur tersebut diulang sampai konvergen.

