
Pemodelan Regresi Logistik Biner Menggunakan Metode Momen Diperumum

Grace Oktavia Yusuf^{1*}, Andi Kresna Jaya², Nirwan Ilyas³
^{1,2,3}Departemen Statistika, Fakultas MIPA,
Universitas Hasanuddin, Makassar, 90245, Indonesia

* Corresponding author, email: grcoctaviayusuf@gmail.com

Abstract

Logistic regression is a regression model that is often used in modeling to data categories, but in determining the model sometime can not be solved in the normal way because the variable response is categorical following the Bernoulli distribution. So in determining the model needed a parameter estimation to get the information about the population parameters. The Generalized method of moments (GMM) is one method of the parameter estimation used to exploit the information on the condition of a population moment which is an extension of the moment method. By using the GMM parameter estimation, it is obtained that using the same moment condition with the generally moment method produces the same estimation with the moment method or OLS estimation. In estimating the logistic regression parameters, an algorithm is needed to solve its nonlinear form, so that it is used the iteration of the reweighted least square whose weight changes every time it is invented.

Keywords: Binary Logistic Regression, Generalized Method of Moments, Reweighted Least Square Iteration.

Abstrak

Regresi logistik merupakan model regresi yang sering digunakan dalam pemodelan data kategori, namun dalam menentukan modelnya terkadang tidak dapat diselesaikan dengan cara biasa dikarenakan variabel respon yang bersifat kategorikal mengikuti distribusi bernoulli. Sehingga dalam menentukan model diperlukan suatu estimasi parameter untuk mendapatkan informasi mengenai parameter populasi. Metode momen diperumum (Generalized method of moments/GMM) adalah salah satu metode estimasi parameter yang digunakan untuk mengeksploitasi informasi bentuk kondisi momen populasi yang merupakan perluasan dari metode momen. Dari penggunaan estimasi parameter GMM diperoleh bahwa dengan menggunakan kondisi momen yang sama dengan metode momen pada umumnya menghasilkan estimasi yang sama dengan metode momen ataupun dengan estimasi OLS. Dalam mengestimasi parameter regresi logistik pun diperlukan suatu algoritma untuk menyelesaikan bentuk nonlinearnya, sehingga digunakan iterasi Reweighted least square yang pembobotnya berubah setiap pengiterasian.

Kata Kunci: Regresi Logistik Biner, Metode Momen Diperumum, Iterasi *Reweighted Least Square*.

1. Pendahuluan

Regresi logistik adalah model regresi yang menggambarkan hubungan antara variabel bebas dengan variabel terikat biner yaitu variabel yang diasumsikan mempunyai 2 nilai (0 atau 1) yang seringkali diartikan gagal atau sukses. Variabel Y biasanya merupakan simbol dari variabel terikat, sedangkan variabel bebas disimbolkan dengan X. Pada regresi linear variabel terikatnya diasumsikan berdistribusi normal, sedangkan

untuk variabel terikat pada regresi logistik bersifat kategorikal, dimana variabel terikat tersebut mengikuti distribusi Bernoulli. Distribusi dari variabel terikat ini yang menjadi pembeda antara regresi logistik dengan regresi linear [1].

Dalam analisis regresi, estimasi parameter merupakan tujuan penting untuk mendapatkan estimator. Estimasi yang baik adalah estimasi yang dapat menggambarkan kuantitas populasi melalui kuantitas sampel. Dalam menentukan penaksir terbaik sangat dipengaruhi oleh penggunaan metode. Penaksiran parameter dapat dilakukan dengan metode maksimum likelihood [2]. Selanjutnya Taurif (2014) menjelaskan bahwa metode momen merupakan salah satu metode yang digunakan untuk memperoleh suatu estimator untuk parameter β dengan ide dasar berupa menyamakan antara momen-momen populasi dengan momen-momen sampel. Dari metode momen tersebut diperkenalkan suatu estimator yang lain yaitu Metode Momen Diperumum atau *Generalized Method of Moments* (GMM). Metode Momen Diperumum (GMM) merupakan salah satu metode yang diperkenalkan oleh Hansen sebagai estimasi parameter yang meminimalkan bentuk kuadrat dari kondisi momen sampel yang terboboti. GMM merupakan metode penaksiran parameter perluasan dari metode momen, dengan cara menyamakan momen kondisi dari populasi dengan momen kondisi dari sampel [3]. Sebagai metode estimasi yang bersifat umum (generalisasi), metode GMM jauh lebih fleksibel karena hanya memerlukan beberapa asumsi yang disebut dengan *moment conditions* populasi yang dikembangkan dari asumsi model [4]. Kadir (2018) dan Kuan (2010) menyatakan bahwa estimasi parameter menggunakan GMM dan metode momen atau OLS menghasilkan hasil estimasi yang sama dengan penggunaan kondisi momen yang sama pula, namun GMM menghasilkan nilai standar eror yang lebih kecil dibandingkan dengan metode momen ataupun OLS, sehingga penggunaannya lebih optimal [5,6].

Pada penelitian kali ini akan digunakan estimasi GMM dalam mengestimasi parameter dari model regresi logistik. Namun, pada umumnya estimasi parameter suatu fungsi tidak dapat diselesaikan secara analitik jika diperoleh bentuk nonlinear, sehingga pada penelitian ini digunakan suatu algoritma untuk memperoleh nilai penduga pembobot (\hat{W}) yaitu algoritma *Iterative Reweighted Least Square* (IRLS). Dari penggunaan metode-metode tersebut akan diperoleh model regresi logistik yang tepat dengan penggunaan estimasi parameter GMM yang lebih fleksibel dan optimal, dan penyelesaian bentuk nonlinear-nya menggunakan *Iterative Reweighted Least Square* (IRLS).

2. Material dan Metode

2.1. Analisis Regresi Berganda

Analisis regresi berganda merupakan analisis untuk menjelaskan hubungan antara dua variabel atau lebih. Analisis regresi berganda dalam statistika merupakan salah satu metode yang bertujuan menentukan hubungan sebab-akibat antara satu variabel dengan

variabel-variabel lainnya, menelusuri pola hubungan pada model yang belum diketahui secara sempurna sebagai alat statistik secara eksploratif. Penerapan analisis regresi dapat secara luas pada berbagai bidang seperti kesehatan, ekonomi, teknik, manajemen, ilmu pertanian, sosial, dan sebagainya.

Bentuk umum model regresi berganda untuk p variabel bebas dengan notasi sebagai berikut:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (1)$$

dengan ε adalah error, $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ adalah parameter, Y adalah variabel terikat (respon) dan x_1, x_2, \dots, x_p adalah variabel bebas. Dalam matriks ditulis sebagai berikut [6]:

$$Y = X\beta + \varepsilon, \text{ dimana } \varepsilon \sim N(0, \sigma^2 I) \quad (2)$$

2.2. Regresi Logistik Biner

Model regresi logistik biner adalah suatu model yang menggambarkan hubungan antara beberapa variabel bebas dengan sebuah variabel respon (terikat) biner. Menurut Agresti (2002) variabel respon pada model regresi logistik dikatakan biner karena terdiri dari dua kategori yaitu 0 dan 1 [1]. Dengan bentuk spesifik dari model regresi logistik sebagai berikut:

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \\ \pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p)} \quad (3)$$

dengan variabel respon Y dituliskan sebagai $y = \pi(x) + \varepsilon$, dimana ε mempunyai salah satu dari kemungkinan dua nilai, yaitu apabila $y = 1$ maka $\varepsilon = 1 - \pi(x)$ dengan peluang $\pi(x)$ dan jika $y = 0$ maka $\varepsilon = -\pi(x)$ dengan peluang $1 - \pi(x)$. Sehingga distribusi erornya mempunyai mean sama dengan nol dan variansi ($\pi(x)(1 - \pi(x))$), yang mengikuti distribusi bernoulli.

2.3. Ordinary Least Square (OLS)

Salah satu metode yang digunakan untuk menaksir parameter model regresi linear adalah metode Ordinary Least Square (OLS). Metode ini bertujuan untuk meminimalkan jumlah kuadrat galat. Misalkan diberikan sampel Y , maka aturan main yang memungkinkan pemakaian sampel untuk mendapatkan taksiran dari β adalah dengan membuat $\varepsilon = Y - X\beta$ sekecil mungkin. Maka perlu memilih parameter β sehingga S seminimal mungkin dengan S adalah:

$$S = \varepsilon' \varepsilon \\ = Y'Y - 2\beta' X'Y + \beta' X'X\beta$$

Untuk meminimumkan fungsi S dapat diperoleh dengan melakukan turunan parsial pertama S terhadap β

$$\begin{aligned} \frac{\partial(S)}{\partial\beta} \Big|_{\beta = \hat{\beta}} &= 0 \\ &= -\mathbf{X}'\mathbf{Y}' + \mathbf{X}'\mathbf{X}\hat{\beta} \end{aligned}$$

Selanjutnya diperoleh penaksir $\hat{\beta}$ yaitu:

$$\begin{aligned} \mathbf{X}'\mathbf{X}\hat{\beta} &= \mathbf{X}'\mathbf{Y}' \\ \hat{\beta}_{OLS} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y}) \end{aligned} \tag{4}$$

2.4. Metode Momen

Metode momen merupakan salah satu metode yang digunakan untuk memperoleh suatu estimator untuk parameter β dengan ide dasar berupa menyamakan antara momen-momen populasi dengan momen-momen sampel. Menurut Chausse (2010) metode momen memerlukan beberapa asumsi tentang apa yang disebut momen kondisi (*moment condition*), dengan kata lain faktor penting dalam asumsi metode momen adalah momen kondisi populasi yang dikembangkan dari asumsi model [4]:

1. $var(\varepsilon_i) = \sigma^2$
2. $E[x_i(y_i - x_i\beta)] = E[x_i\varepsilon_i] = 0$
3. $E[\varepsilon_i\varepsilon_j] = 0 ; i \neq j$

Penaksiran parameter model regression digunakan asumsi pada poin (2) yaitu:

$$\begin{aligned} g(\beta) &= \frac{1}{n} \sum_{i=1}^n x_i(y_i - x_i\beta) = 0 \\ &= \frac{1}{n} (\mathbf{X}'(\mathbf{Y} - \mathbf{X}\beta)) = 0 \end{aligned}$$

Selanjutnya parameter $\hat{\beta}$ dapat diperoleh dengan menyelesaikan fungsi $g(\beta) = 0$ yaitu:

$$\begin{aligned} \frac{1}{n} (\mathbf{X}'(\mathbf{Y} - \mathbf{X}\beta)) &= 0 \\ \frac{1}{n} \mathbf{X}'\mathbf{Y} - \frac{1}{n} \mathbf{X}'\mathbf{X}\beta &= 0 \\ \frac{1}{n} \mathbf{X}'\mathbf{Y} &= \frac{1}{n} \mathbf{X}'\mathbf{X}\beta \\ \hat{\beta}_{MM} &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y}) = \hat{\beta}_{OLS} \end{aligned}$$

2.5. Metode Momen Diperumum

Metode momen diperumum (GMM) merupakan metode penaksiran parameter perluasan dari metode momen. GMM menyamakan momen kondisi dari populasi dengan momen kondisi dari sampel. Metode GMM didapat dengan cara meminimumkan jumlah kuadrat terboboti dari momen kondisi sampel. Matriks pembobot \hat{W} merupakan suatu matrik simetris definit positif yang bukan fungsi dari β dengan notasi sebagai berikut:

$$J(\beta) = \|\bar{g}(\beta)\|_{\hat{W}}^2 = \bar{g}(\beta)' \hat{W} \bar{g}(\beta) \tag{5}$$

Dalam pengaplikasiannya GMM menjadi lebih optimal dari penggunaan teknik estimasi yang lain karena merupakan bentuk diperumum dari metode momen. Sehingga estimasi GMM dapat diaplikasikan pada bentuk momen yang disebut “*moment condition*” dengan teknik estimasi yang lainnya.

2.6. GMM Interpretasi dari Estimasi OLS

Untuk model regresi linear adalah sebagai berikut:

$$Y = X\beta + \varepsilon \quad (6)$$

Dengan mengikuti asumsi momen kondisi populasi, yaitu:

$$E[X\varepsilon] = E[X(Y - X\beta)] = 0 \quad (7)$$

Selanjutnya diperoleh estimasi parameter sebagai berikut:

$$\hat{\beta}_{MM} = (X'X)^{-1}(X'Y) = \hat{\beta}_{OLS} \quad (8)$$

Kuan (2010) menjelaskan bahwa kondisi ini sama dengan momen kondisi pertama untuk estimasi OLS. Karena hal tersebut penaksir OLS yang diselesaikan dari kondisi momen sampel diatas dapat dianggap sebagai penaksir GMM, dengan penerapan teori asimtotik untuk estimasi GMM [6], yaitu:

1. $W = \Omega^{-1}(\beta)$ dengan $\Omega(\beta) = \sigma_{\varepsilon}^2 E(\mathbf{X}\mathbf{X}') = \sigma_{\varepsilon}^2 Q$ (9)
2. $\frac{\partial J(\beta)}{\partial \beta} = -E(\mathbf{X}\mathbf{X}') = -Q$

2.7. Iterasi Reweighted Least Square

Karena fungsi logistik merupakan fungsi nonlinear maka proses perhitungannya dapat didekati dengan metode *Iteratively Reweighted Least Square* (IRLS). Pada iterasi ini pembobot berubah setiap iterasi, adapun penduga WLS adalah sebagai berikut:

$$\hat{\beta}_{WLS} = [(X')W(X)]^{-1}(X')W(Y) \quad (10)$$

Dengan melakukan pendugaan parameter dilakukan pengiterasian pada iterasi ke-t dalam proses iterasi $t = 0, 1, 2, \dots$ adalah sebagai berikut:

$$\hat{\beta}_{t+1} = \hat{\beta}_t + \left(\frac{\partial^2 J(\beta)}{\partial^2 \beta}\right)^{-1} \left(-E\left(\frac{\partial J(\beta)}{\partial \beta}\right)\right) \quad (11)$$

dengan:

- $\hat{\beta}_t$: vektor untuk β pada iterasi ke-t
- $\frac{\partial^2 J(\beta)}{\partial^2 \beta}$: turunan kedua dari fungsi kriteria $J(\beta)$
- $\frac{\partial J(\beta)}{\partial \beta}$: turunan pertama dari fungsi kriteria $J(\beta)$

Proses iterasi berhenti sampai diperoleh penduga parameter yang konvergen yaitu $\sum_{i=1}^n |e^{g(x)}|$ konvergen dimana $\sum_{i=1}^n |\varepsilon_i^m - \varepsilon_i^{m-1}| < 0.001$ [7].

2.8. Uji Signifikansi Parameter

Dalam pembentukan model regresi perlu diketahui apakah parameter yang telah diestimasi mempengaruhi variabel respon model. Untuk mengetahui pengaruh variabel

bebas terhadap variabel terikat secara bersama-sama di dalam model, dapat menggunakan Uji Likelihood Ratio. Hipotesisnya adalah sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$ (tidak ada pengaruh variabel bebas secara simultan terhadap variabel terikat).

H_1 : Minimal ada satu $\beta_j \neq 0$ (ada pengaruh paling sedikit satu variabel bebas terhadap variabel terikat).

Untuk $j = 1, 2, \dots, p$. Statistik uji yang digunakan adalah:

$$G^2 = -2 \ln \left[\frac{\binom{n_1}{n}^{n_1} \binom{n_0}{n}^{n_0}}{\prod_{i=1}^n [\hat{\pi}_i]^{y_i} [1 - \hat{\pi}_i]^{1 - y_i}} \right] \quad (12)$$

Dengan :

$$n_1 = \sum_{i=1}^n y_i, n_0 = \sum_{i=1}^n (1 - y_i), n = n_0 + n_1$$

Statistik G^2 ini mengikuti distribusi *Chi-Square* dengan derajat bebas p sehingga hipotesis ditolak jika $p - value < \alpha$ atau $G^2 > \chi_{(a,p)}^2$, yang berarti variabel bebas X secara bersama-sama mempengaruhi variabel terikat Y.

3. Hasil dan Diskusi

Metode Momen Diperumum didapat dengan cara meminimumkan jumlah kuadrat terboboti dari momen kondisi sampel, dengan suatu matriks bobot (W) yang merupakan suatu matriks definit positif dengan notasi pada persamaan (5) sebagai berikut:

$$J(\beta) = \|\bar{g}(\beta)\|_W^2 = \bar{g}(\beta)' W \bar{g}(\beta)$$

Dengan mengacu pada model regresi logistik biner yaitu:

$$Y = \frac{\exp(\mathbf{X}\beta)}{1 + \exp(\mathbf{X}\beta)} + \varepsilon_i = \pi(x) + \varepsilon$$

Kondisi momen populasi:

$$E[\mathbf{X}\varepsilon] = 0$$

Dengan mensubstitusikan bentuk galat

$$E[\mathbf{X}(Y - \pi(x))] = 0$$

Sehingga kondisi momen sampel menjadi

$$\left[n^{-1} \sum_{i=1}^n x_i (y_i - \pi(x_i)) \right] = \left[n^{-1} \sum_{i=1}^n g_i(\beta) \right] = \bar{g}(\beta) \quad (13)$$

Untuk mengestimasi parameter dengan menggunakan GMM, dilakukan dengan cara meminimumkan fungsi kriteria berikut:

$$\begin{aligned} \frac{\partial J(\beta)}{\partial \beta} \Big|_{\beta = \hat{\beta}} &= 0 \\ \frac{\partial \bar{g}(\beta)' W \bar{g}(\beta)}{\partial \beta} \Big|_{\beta = \hat{\beta}} &= 0 \\ 2 \frac{\partial \bar{g}(\beta)'}{\partial \beta} \Big|_{\beta = \hat{\beta}} W \bar{g}(\hat{\beta}) &= 0 \end{aligned} \quad (14)$$

Dengan teori asimtotik untuk estimasi GMM diketahui pada persamaan (9) dimana:

$$W = \sigma_\varepsilon^2 E(\mathbf{X}\mathbf{X}') = \sigma_\varepsilon^2 \frac{1}{n} \mathbf{X}\mathbf{X}' \quad (15)$$

Dengan demikian jika persamaan (13) dan (15) disubstitusikan ke persamaan (14) diperoleh:

$$\begin{aligned}
 & 2 \frac{\partial \bar{g}(\beta)'}{\partial \beta} \Big|_{\beta=\hat{\beta}} W \bar{g}(\hat{\beta}) = 0 \\
 & 2 \left(\frac{\partial}{\partial \beta} \Big|_{\beta=\hat{\beta}} [\mathbf{XY} - \mathbf{X}\boldsymbol{\pi}(\mathbf{x})] \right) \left(\sigma_{\varepsilon}^2 \frac{1}{n} \mathbf{XX}' \right)^{-1} (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0 \\
 & 2 \left(\frac{\partial}{\partial \beta} \Big|_{\beta=\hat{\beta}} \left[\mathbf{XY} - \frac{\mathbf{X}(\exp(\mathbf{X}\hat{\boldsymbol{\beta}}))}{1+\exp(\mathbf{X}\hat{\boldsymbol{\beta}})} \right] \right) \left(\sigma_{\varepsilon}^2 \frac{1}{n} \mathbf{XX}' \right)^{-1} (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0 \\
 & -2 \left(\mathbf{XX}' \frac{(\exp(\mathbf{X}\hat{\boldsymbol{\beta}}))}{1+\exp(\mathbf{X}\hat{\boldsymbol{\beta}})} \frac{1}{1+\exp(\mathbf{X}\hat{\boldsymbol{\beta}})} \right) \left(\frac{1}{n} \mathbf{XX}' \right)^{-1} (\sigma_{\varepsilon}^2)^{-1} (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0 \\
 & -2 \left(\mathbf{XX}'(\hat{\boldsymbol{\pi}}(\mathbf{x}))(\mathbf{1} - \hat{\boldsymbol{\pi}}(\mathbf{x})) \right) \left(\frac{1}{n} \mathbf{XX}' \right)^{-1} (\sigma_{\varepsilon}^2)^{-1} (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0
 \end{aligned}$$

Diketahui varians dari regresi logistik adalah $\sigma_{\varepsilon}^2 = \text{diag}[(\hat{\boldsymbol{\pi}}(\mathbf{x}))(\mathbf{1} - \hat{\boldsymbol{\pi}}(\mathbf{x}))]$, sehingga:

$$\begin{aligned}
 & -2 \left(\mathbf{XX}'(\hat{\boldsymbol{\pi}}(\mathbf{x}))(\mathbf{1} - \hat{\boldsymbol{\pi}}(\mathbf{x})) \right) (\mathbf{XX}')^{-1} \left((\hat{\boldsymbol{\pi}}(\mathbf{x}))(\mathbf{1} - \hat{\boldsymbol{\pi}}(\mathbf{x})) \right)^{-1} (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0 \\
 & -2 (\mathbf{XX}')(\hat{\boldsymbol{\pi}}(\mathbf{x}))(\mathbf{1} - \hat{\boldsymbol{\pi}}(\mathbf{x})) (\mathbf{XX}')^{-1} \left((\hat{\boldsymbol{\pi}}(\mathbf{x}))(\mathbf{1} - \hat{\boldsymbol{\pi}}(\mathbf{x})) \right)^{-1} (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0 \\
 & -2 (\mathbf{XY} - \mathbf{X}\hat{\boldsymbol{\pi}}(\mathbf{x})) = 0 \quad \text{atau} \quad -2 \left(\mathbf{XY} - \mathbf{X} \frac{(\exp(\mathbf{X}\hat{\boldsymbol{\beta}}))}{1+\exp(\mathbf{X}\hat{\boldsymbol{\beta}})} \right) = 0 \tag{16}
 \end{aligned}$$

Dengan berdasar pada momen kondisi dan penduga parameter menggunakan GMM, akan dilakukan metode iterasi untuk menyelesaikan bentuk nonlinear dari regresi logistik biner. Berdasarkan persamaan 11 perlu dilakukan turunan kedua untuk memenuhi persamaan tersebut. Maka dengan melakukan turunan kedua pada fungsi kriteria $J(\beta)$ diperoleh:

$$\begin{aligned}
 \frac{\partial^2 J(\beta)}{\partial^2 \beta} \Big|_{\beta=\hat{\beta}} &= \frac{\partial}{\partial \beta} \Big|_{\beta=\hat{\beta}} \left[-2 \sum_{i=1}^n \left(\frac{-x_i(\exp(x_i\hat{\beta}))}{1+\exp(x_i\hat{\beta})} + y_i x_i \right) \right] \\
 &= 2 \sum_{i=1}^n \left(\frac{\partial}{\partial \beta} \Big|_{\beta=\hat{\beta}} \left(\frac{x_i(\exp(x_i\hat{\beta}))}{1+\exp(x_i\hat{\beta})} \right) \right)
 \end{aligned}$$

Selanjutnya menyelesaikan turunan keduanya, digunakan aturan fungsi sebagai berikut:

Misalkan: $u = x_i(\exp(x_i\hat{\beta}))$ dan $v = 1 + \exp(x_i\hat{\beta})$ maka

$u' = x_i^2(\exp(x_i\hat{\beta}))$ dan $v' = x(\exp(x_i\hat{\beta}))$ sehingga

$$\begin{aligned}
 \frac{\partial^2 J(\beta)}{\partial^2 \beta} \Big|_{\beta=\hat{\beta}} &= - \sum_{i=1}^n \left(\frac{u'v - uv'}{v^2} \right) \\
 &= 2 \sum_{i=1}^n \left(\frac{x_i^2(\exp(x_i\hat{\beta}))(1+\exp(x_i\hat{\beta})) - x_i(\exp(x_i\hat{\beta}))(x_i(\exp(x_i\hat{\beta})))}{(1+\exp(x_i\hat{\beta}))^2} \right) \\
 &= 2 \sum_{i=1}^n \left(\frac{x_i^2(\exp(x_i\hat{\beta})) + x_i^2(\exp(x_i\hat{\beta}))\exp(x_i\hat{\beta}) - x_i(\exp(x_i\hat{\beta}))x_i(\exp(x_i\hat{\beta}))}{(1+\exp(x_i\hat{\beta}))^2} \right)
 \end{aligned}$$

$$\begin{aligned}
 &= 2 \sum_{i=1}^n \left(\frac{x_i^2 (\exp(x_i \hat{\beta}))}{(1 + \exp(x_i \hat{\beta}))^2} \right) \\
 &= 2 \sum_{i=1}^n \left(x_i^2 \frac{\exp(x_i \hat{\beta})}{1 + \exp(x_i \hat{\beta})} \frac{1}{1 + \exp(x_i \hat{\beta})} \right) \\
 &= 2 \sum_{i=1}^n \left(x_i^2 \hat{\pi}(x_i) (1 - \hat{\pi}(x_i)) \right) \\
 &= 2 \sum_{i=1}^n (x_i^2 \widehat{W}_i)
 \end{aligned}$$

Atau dalam bentuk matriks dapat ditulis:

$$\left. \frac{\partial^2 J(\beta)}{\partial^2 \beta} \right|_{\beta=\hat{\beta}} = 2(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}) \tag{17}$$

dengan $\widehat{\mathbf{W}} = \text{diag}[\hat{\pi}(\mathbf{x})(1 - \hat{\pi}(\mathbf{x}))]$

Berdasarkan persamaan (16) dan (17), maka persamaan (11) dapat dituliskan sebagai berikut:

$$\begin{aligned}
 \hat{\beta}_{t+1} &= \hat{\beta}_t + \left(\frac{\partial^2 J(\beta)}{\partial^2 \beta} \right)^{-1} \left(-E \left(\frac{\partial J(\beta)}{\partial \beta} \right) \right) \\
 &= \hat{\beta}_t + \left(2(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}) \right)^{-1} \left(-E(-2\mathbf{X}'(\mathbf{Y} - \hat{\pi}(\mathbf{x}))) \right) \\
 &= \hat{\beta}_t + (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} (\mathbf{X}'(\mathbf{Y} - \hat{\pi}(\mathbf{x})))
 \end{aligned}$$

Untuk memperoleh pendugaan parameter $\hat{\beta}_j$ maka $\frac{\partial \bar{g}(\beta)}{\partial \beta} = 0$. Karena $\hat{\pi}(\mathbf{x})$ merupakan fungsi nonlinear maka penyelesaian fungsi tersebut didekati dengan menggunakan metode IRLS. Dengan penduga WLS yaitu:

$$\hat{\beta}_{WLS} = (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} (\mathbf{X}'\widehat{\mathbf{W}}\widehat{\mathbf{Y}})$$

Maka persamaan dapat ditulis kembali menjadi:

$$\begin{aligned}
 \hat{\beta}_{t+1} &= \hat{\beta}_t + (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} (\mathbf{X}'(\mathbf{Y} - \hat{\pi}(\mathbf{x}))) \\
 &= (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\hat{\beta}_t) + (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} (\mathbf{X}'(\mathbf{Y} - \hat{\pi}(\mathbf{x}))) \\
 &= (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} [(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\hat{\beta}_t) + (\mathbf{X}'(\mathbf{Y} - \hat{\pi}(\mathbf{x})))] \\
 &= (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} [(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\hat{\beta}_t) + (\mathbf{X}'\widehat{\mathbf{W}}\widehat{\mathbf{W}}^{-1}(\mathbf{Y} - \hat{\pi}(\mathbf{x})))] \\
 &= (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} \mathbf{X}'\widehat{\mathbf{W}} [(\mathbf{X}\hat{\beta}_t) + (\widehat{\mathbf{W}}^{-1}(\mathbf{Y} - \hat{\pi}(\mathbf{x})))] \\
 &= (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} \mathbf{X}'\widehat{\mathbf{W}} \left[\ln \frac{\hat{\pi}(\mathbf{x})}{1 - \hat{\pi}(\mathbf{x})} + (\widehat{\mathbf{W}}^{-1}(\mathbf{Y} - \hat{\pi}(\mathbf{x}))) \right] \\
 &= (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} \mathbf{X}'\widehat{\mathbf{W}}\hat{\mathbf{z}}
 \end{aligned}$$

Dengan $\hat{\mathbf{z}} = \left[\ln \frac{\hat{\pi}(\mathbf{x})}{1 - \hat{\pi}(\mathbf{x})} + (\widehat{\mathbf{W}}^{-1}(\mathbf{Y} - \hat{\pi}(\mathbf{x}))) \right]$

Dari iterasi tersebut akan diperoleh penduga GMM untuk β dan W yang dinotasikan $\hat{\beta}_{GMM}$ dan \widehat{W} , yaitu:

$$\hat{\beta}_{GMM} = \hat{\beta} = (\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X})^{-1} \mathbf{X}'\widehat{\mathbf{W}}\hat{\mathbf{z}}$$

dengan $\widehat{W} = \text{diag}[\widehat{\pi}(x)(1 - \widehat{\pi}(x))]$

4. Kesimpulan

Pendugaan parameter pada model regresi logistik dengan metode GMM dengan menggunakan kondisi momen yang sama dengan metode momen menghasilkan hasil estimasi parameter yang sama pula. Sehingga estimasi GMM menjadi lebih fleksibel penggunaannya dengan berpacu pada momen kondisi yang digunakan. Kemudian pada pendugaan parameter dengan model regresi logistik tidak dapat diselesaikan secara analitik sehingga dibutuhkan suatu algoritma iterasi untuk menyelesaikan fungsi tersebut. Sehingga dalam jurnal ini untuk menduga estimasi parameter pada regresi logistik biner digunakan estimasi GMM dan pengiterasian IRLS, dan menghasilkan estimasi sebagai berikut:

$$\widehat{\beta}_{GMM} = \widehat{\beta} = (\mathbf{X}'\widehat{W}\mathbf{X})^{-1}\mathbf{X}'\widehat{W}\widehat{z}$$

Daftar Pustaka

- [1] Agresti, A. *Categorical Data Analysis*. New York: A John Wiley and Sons Inc, 2002.
- [2] Islamiyati, A. Estimasi Parameter Model Regresi Logistik Biner Komponen Utama Non Linear dengan Maksimum Likelihood. *Jurnal Matematika, Statistika dan Komputasi*, Vol 11, No. 2: 122-128, 2015.
- [3] Taurif, M. *Estimation Of Generalized Method Of Moment In Logistic*. Prosiding Seminar Nasional Matematika, Universitas Jember, 2014.
- [4] Chausse, P. *Computing Generalized Method of Moments and Generalized Empirical Likelihood with R*. *Journal of Statistical Software*, 2010.
- [5] Kadir S. Perbandingan Estimasi Model Respon Kualitatif Menggunakan Metode OLS, GMM dan Maximum Likelihood. Palembang: Universitas Sriwijaya, 2018.
- [6] Kuan, C. M. *Generalized Method of Moment*. Taiwan: Department of Finance & CRETA National Taiwan University, 2010.
- [7] Muliati, A. Pendugaan Parameter Model Regresi Logistik dengan Metode Ridge. Makassar: Statistika Universitas Hasanuddin, 2018.