BAB II

TINJAUAN PUSTAKA

2.1 Fertilitas

Fertilitas sebagai istilah demografi diartikan sebagai hasil reproduksi yang nyata dari seorang wanita atau kelompok wanita. Dengan kata lain fertilitas ini menyangkut banyaknya bayi yang lahir hidup. Fertilitas mencakup peranan kelahiran pada perubahan penduduk. Istilah fertilitas adalah sama dengan kelahiran hidup (*live birth*), yaitu terlepasnya bayi dari rahim seorang perempuan dengan ada tanda-tanda kehidupan, misalnya berteriak, bernafas, jantung berdenyut dan sebagainya (Mantra, 2003).

Seorang perempuan yang secara biologis subur (*fecund*) tidak selalu melahirkan anak-anak yang banyak, misalnya dia mengatur fertilitas dengan abstinensi atau menggunakan alat-alat kontrasepsi. Kemampuan biologis seseorang perempuan untuk melahirkan sangat sulit untuk diukur. Ahli demografi hanya menggunakan pengukuran terhadap kelahiran hidup (*live birth*).

2.2 Keluarga Berencana (KB)

Menurut BKKBN (1998) keluarga berencana artinya mengatur jumlah anak sesuai kehendak anda dan mengatur sendiri kapan anda ingin hamil atau salah satu usaha masalah kependudukan sekaligus merupakan bagian terpadu dalam program pembangunan nasional yang bertujuan untuk turut serta menciptakan kesejahteraan ekonomi, spiritual, budaya penduduk Indonesia agar dapat dicapai keseimbangan yang baik dengan kemampuan produksi nasional.

Keluarga berencana (KB) adalah meningkatkan derajat kesehatan dan kesejahteraan ibu dan anak, keluarga serta bangsa pada umumnya, meningkatkan martabat kehidupan rakyat dengan cara menurunkan angka kelahiran sehingga pertambahan penduduk tidak melebihi kemampuan untuk meningkatkan produksi.

2.3 Pekerjaan

Menurut (Manginsihi, 2003) Pekerjaan adalah kegiatan yang dilakukan oleh orang tua untuk mencari nafkah. Pekerjaan yang ditekuni oleh setiap orang pasti berbeda-beda, dari perbedaan itulah akan menyebabkan perbedaan tingkat penghasilan yang rendah sampai tingkat penghasilan yang tinggi, tergantung pada jenis pekerjaan yang di tekuni. Jenis perkejaan akan menentukan status sosial ekonomi karena dari bekerja semua akan terpenuhi.

2.4 Pendidikan

Menurut Bouge (Lucas, 1990) mengemukakan bahwa pendidikan menunjukkan pengaruh yang lebih kuat terhadap fertilitas dari pada variabel lain. Seseorang dengan tingkat pendidikan yang relatif tinggi tentu saja dapat mempertimbangkan berapa keuntungan financial yang di peroleh seorang anak dibandingkan dengan biaya yang harus di keluarkan.

Tingkat pendidikan seseorang dalam masyarakat berbeda-beda. Berikut tingkat pendidikan yang ada:

a. Pendidikan dasar: TK (Taman Kanak-Kanak) dan SD (Sekolah dasar).

b. Pendidikan lanjutan: SMP (Sekolah Menengah Pertama), SMA (Sekolah Menengah Atas) dan SMK (Sekolah Menengah Kejuruan).

c. Pendidikan tinggi: Diploma, Sarjana, Pasca Sarjana, Doktor.

2.5 Usia Pernikahan Pertama

Badan Pusat Statistik (2016) mendefinisikan usia pernikahan pertama sebagai umur pertama menikah yang berarti juga saat dimulainya masa reproduksinya pembuahan. Semakin muda usia pernikahan pertama maka akan semakin panjang masa reproduksinya atau semakin banyak anak dilahirkan, usia pernikahan pertama sebesar 21 berarti rata-rata penduduk menikah saat berusia 21 tahun. Usia pernikahan pertama berhubungan terbalik dengan jumlah kelahiran, usia pernikahan pertama yang semakin rendah mengindikasikan tingkat fertilitas yang tinggi. Usia pernikahan pertama berfungsi sebagai salah satu dasar pengambilan kebijakan untuk menekan laju pertumbuhan penduduk.

2.6 Regresi Logistik Biner

Regresi logstik merupakan analisis yang penting dalam melihat hubungan antara suatu variabel respon dengan satu atau lebih variabel penjelas dengan data kategorik.

Apabila variabel respon terdiri atas dua kategori yaitu $Y = 1$ (Sukses) dan $Y = 0$ (Gagal), metode regresi logistik yang dapat diterapkan adalah regresi logistik biner. Variabel yang dikotomik/biner adalah variabel yang hanya mempunyai dua kategori saja, yaitu kategori yang menyatakan kejadian sukses (Y = 1) dan kategori yang menyatakan kejadian gagal (Y = 0). Variabel terikat Y ini, diasumsikan mengikuti distribusi Bernoulli.

2.6.1 Bentuk Umum Model Regresi Logistik

Bentuk umum model peluang regresi logistik adalah sebagai berikut:

$$P(Y=1)=\pi(x)=\frac{\exp(\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_pX_p)}{1+\exp(\beta_0+\beta_1X_1+\beta_2X_2+...+\beta_pX_p)} \qquad (2.1)$$

dimana:

X_i : Variabel penjelas, dengan $i = 1,2,\ldots,p$.

$\pi(x)$: Peluang terjadinya kejadian yang sukses yaitu Y = 1.

β_j : Nilai parameter ke-j ($j = 0,1,2,\ldots,p$).

Persamaan (2.1) merupakan model nonlinear maka perlu ditransformasikan ke dalam bentuk *logit* agar dapat dilihat hubungan variabel penjelas dengan variabel respon (Hosmer & Lemeshow, 2000).

2.6.2 Bentuk Linear dari Regresi Logistik

Dengan melakukan transformasi logit dari $\pi(x)$ diperoleh persamaan yang lebih sederhana yang merupakan fungsi linear yaitu:

$$g(x)=\ln\left[\frac{\pi(x)}{1-\pi(x)}\right]$$

$$= \left[\frac{\dfrac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}}{1 - \dfrac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}} \right]$$

$$g(x) = \ln \left[\frac{\dfrac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}}{\dfrac{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p) - \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}} \right]$$

$$= \ln \left[\frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p) - \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)} \right]$$

$$= \ln \left[\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p) \right]$$

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \tag{2.2}$$

Persamaan (2.2) merupakan fungsi linear dalam parameter-parameternya atau g(x) telah linear dalam parameter-parameternya (Cullagh, 1989).

2.6.3 Penaksiran Parameter

Model regresi logistik menggunakan metode *Maximum Likelihood* untuk menduga parameter-parameternya. Dalam model regresi logistik variabel terikat mengikuti distribusi Bernoulli dengan fungsi kepadatan peluang sebagai berikut:

$$f(\beta, \mathrm{y}_i) = \pi(x)^{yi} \left[1 - \pi(x)\right]^{1-yi} \tag{2.3}$$

Karena nilai variabel terikat (Y_i) diasumsikan saling bebas, maka diperoleh fungsi *likelihood* sebagai berikut:

$$l(\beta) = \prod_{i=1}^{1} f(\beta, y_i)$$

Sedangkan persamaan *likelihood* adalah:

$$L(\beta) = \ln\left[l(\beta)\right] = \sum_{i=1}^{k}\left\{y_i \ln\left[\pi(x)\right] + \left(1-y_i\right)\ln\left[1-\pi(x)\right]\right\} \quad (2.4)$$

Untuk mendapatkan nilai β yang memaksimumkan $L(\beta)$ dilakukan diferensiasi (turunan) terhadap β, dengan syarat: (1) $\frac{\partial L}{\partial \beta} = 0$, dan (2) $\frac{\partial^2 L}{\partial \beta^2} < 0$

Sehingga diperoleh persamaan:

$$\sum_{i=1}^{k} x\left[y_i - \pi(x)\right] = 0 \quad (2.5)$$

Persamaan tidak linear dalam β, sehingga solusi bagi $\hat{\beta} = \hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p$ tidak dapat dituliskan secara eksplisit. Untuk mendapatkan nilai β digunakan metode iterasi *Newton Raphson*. Iterasi merupakan metode yang umum dalam membantu perhitungan estimasi dari β (Hosmer & Lemeshow, 2000).

2.6.4 Uji Signifikansi Parameter

2.6.4.1 Uji Simultan

Uji yang digunakan untuk menguji kelayakan model dengan menggunakan keseluruhan variabel penjelas secara bersama-sama digunakan uji *likelihood ratio* atau uji signifikansi model.

Hipotesis:

H_0 : $\beta_1 = \beta_2 = ... = \beta_j = 0$ (tidak ada pengaruh antara variabel-variabel penjelas terhadap variabel respon).

H_1 : Minimal ada satu $\beta_j \neq 0$ (minimal ada satu variabel penjelas yang berpengaruh terhadap variabel respon).

Statistik uji:

$$G = -2\ln\left[\frac{(L_0)}{(L_1)}\right] \tag{2.6}$$

dimana:

L_0 : fungsi *likelihood* tanpa variabel penjelas.

L_1 : fungsi *likelihood* dengan semua variabel penjelas.

Statistik G mengikuti sebaran khi-kuadrat dengan derajat bebas banyaknya parameter yang terdapat dlam model (k-1) (Agresti, 1990).

Kriteria uji:

$G > \chi^2_{(df;\alpha)}$ atau *p-value* < α, maka H_0 ditolak.

$G \leq \chi^2_{(df;\alpha)}$ atau *p-value* ≥ α, maka H_0 diterima.

2.6.4.2 Uji Parsial

Uji yang digunakan untuk mengetahui variabel penjelas yang mempengaruhi variabel respon, digunakan uji koefisien parameter β secara parsial dengan menggunakan uji wald.

Hipotesis:

H_0 : $\beta_j = 0$ (tidak ada pengaruh antara variabel penjelas ke-*j* dengan variabel respon)

H_1 : $\beta_j \neq 0$ (ada pengaruh antara variabel penjelas ke-*j* dengan variabel respon)

Statistik uji:

$$W = \left(\frac{\hat{\beta}_j}{se\left(\hat{\beta}_j\right)} \right)^2 \qquad (2.7)$$

dimana:

$\hat{\beta}_j$: penduga β_j

$se\left(\hat{\beta}_j\right)$: galat baku dari $\hat{\beta}_j$

Kriteria uji:

Kriteria uji yang digunakan berdasarkan statistik uji W diasumsikan mengikuti sebaran khi-kuadrat dengan derajat bebas = 1 (Agresti, 1990).

Jika $W > \chi^2_{(df;\alpha)}$ atau *p-value* < α, maka H_0 ditolak.

Jika $W \leq \chi^2_{(df;\alpha)}$ atau *p-value* ≥ α, maka H_0 diterima.

2.6.5 Uji Kesesuaian Model

Untuk menguji kesesuaian model, dilakukan uji Hosmer Lemeshow terhadap model regresi logistik.

Hipotesis:

H_0 : Model sesuai/tepat (tidak terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model)

H_1 : Model tidak sesuai/tepat (terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil prediksi model)

Kriteria uji:

Jika *p-value* $< \alpha$, maka H_0 ditolak.

Jika *p-value* $\geq \alpha$, maka H_0 diterima.

2.6.6 Odds Ratio

Odds ratio merupakan perbandingan tingkat resiko relatif dari dua buah nilai variabel penjelas X_j atau resiko kecenderungan misalkan $X_j = 1$ terhadap $X_j = 0$. Dengan kata lain, resiko kecenderungan pengaruh observasi $X_j = 1$ adalah n kali lipat dibandingkan observasi $X_j = 0$. Odds ratio dilambangkan dengan ψ yang merupakan ukuran untuk mengetahui tingkat resiko, yaitu perbandingan antara dua nilai variabel penjelas X_j antara kejadian-kejadian yang masuk kategori sukses dan gagal.

Persamaan dari odds ratio (Hosmer & Lemeshow, 2000) dituliskan sebagai berikut:

$$\psi = \frac{\pi(1)/1-\pi(1)}{\pi(0)/1-\pi(0)} \tag{2.8}$$

dimana:

Untuk $X_j = 1$, digunakan:

$$\frac{\pi(1)}{1-\pi(1)} = \frac{\dfrac{\exp\{\beta_0+\beta_j\}}{1+\exp\{\beta_0+\beta_j\}}}{1-\dfrac{\exp\{\beta_0+\beta_j\}}{1+\exp\{\beta_0+\beta_j\}}} = \frac{\exp\{\beta_0+\beta_j\}}{1} = \exp\{\beta_0+\beta_j\}$$

Untuk $X_j = 0$, digunakan:

$$\frac{\pi(0)}{1-\pi(0)} = \frac{\dfrac{\exp\{\beta_0\}}{1+\exp\{\beta_0\}}}{1-\dfrac{\exp\{\beta_0\}}{1+\exp\{\beta_0\}}} = \frac{\dfrac{\exp\{\beta_0\}}{1+\exp\{\beta_0\}}}{\dfrac{1+\exp\{\beta_0\}-\exp\{\beta_0\}}{1+\exp\{\beta_0\}}} = \exp\{\beta_0\}$$

Sehingga diperoleh persamaan sebagai berikut:

$$\psi = \frac{\pi(1)/1-\pi(1)}{\pi(0)/1-\pi(0)}$$

$$= \frac{\exp\{\beta_0+\beta_j\}}{\exp\{\beta_0\}}$$

$$= \exp\{\beta_0+\beta_j-\beta_0\}$$

$$= \exp\{\beta_j\} \tag{2.9}$$

2.7 Multivariate Adaptive Regression Spline (MARS)

MARS merupakan salah satu pendekatan regresi nonparametrik multivariat yang berguna untuk mengatasi permasalahan data yang berdimensi tinggi, yaitu data yang memiliki jumlah variabel prediktor sebanyak $3 \leq n \leq 20$ (Friedman, 1991). Selain itu model MARS mampu menghasilkan prediksi variabel respon yang akurat. Metode MARS menjadi popular karena tidak menentukan tipe khusus seperti hubungan (linear, kuadratik, dan kubik) diantara variabel prediktor dan respon pada proses pembentukan model MARS tidak memerlukan asumsi (Otok, Guritno, Subanar & Haryatmi, 2006).

Beberapa istilah yang perlu diperhatikan dalam pemodelan MARS adalah sebagai berikut.

1. Knot

Knot merupakan suatu nilai/titik tempat perubahan pola apabila suatu garis regresi tidak bisa menjelaskan keseluruhan data yang ada dari variabel prediktor. Knot merupakan akhir dari sebuah garis regresi dan juga awal dari garis regresi yang lain (Nash & Bradford, 2001).

2. Basis Fungsi

Basis fungsi merupakan suatu fungsi yang digunakan untuk menjelaskan hubungan antara variabel respon dan variabel prediktor. Basis fungsi bisa memiliki lebih dari satu variabel yang merupakan fungsi dari tiap garis regresi yang dihasilkan. Maksimum basis fungsi yang diperbolehkan adalah 2 sampai 4 kali jumlah variabel prediktornya.

3. Interaksi

Interaksi merupakan hubungan korelasi antar variabel dengan maksimum interaksi (MI) adalah 1, 2, dan 3. Jika MI lebih dari tiga maka akan menghasilkan model yang lebih kompleks.

4. Minimum Observasi

Minimum Observasi (MO) merupakan jumlah pengamatan paling minimal antar knot sebesar 0, 1, 2, dan 3.

Secara umum estimator model MARS dapat ditulis pada persamaan berikut.

$$\hat{f}(x) = \hat{\alpha}_0 + \sum_{m=1}^{M} \hat{\alpha}_m \prod_{k=1}^{K_m} [S_{km}(X_{v(km)} - t_{km})] \quad (2.10)$$

dengan :

$\hat{\alpha}_0$ = koefisien konstanta basis fungsi B_0

α_m = koefisien dari basis fungsi ke-*m*

M = maksimum basis fungsi

m = banyaknya basis fungsi

k_m = banyaknya interaksi pada basis fungsi *m*

k = banyaknya interaksi

S_{km} = nilainya 1 atau -1 jika data berada di sebelah kanan atau kiri titik knot

$X_{v(km)}$ = variabel prediktor

t_{km} = nilai knot dari variabel prediktor $x_{v(k,m)}$

Berdasarkan persamaan (2.10) model MARS dapat dituliskan sebagai berikut.

$$y_i = \alpha_0 + \sum_{m=1}^{M} \alpha_m B_{im}(x) + \varepsilon_i \quad (2.11)$$

dengan $B_{im}(x) = \prod_{k=1}^{K_m} [S_{km}(X_{v(k,m)i} - t_{km})]$. Sehingga jika ditulis dalam bentuk matriks dapat menjadi

$$y = B\,\alpha + \varepsilon \quad (2.12)$$

dengan,

$$y = (y_1 y_2 y_3, \dots, y_n)^T, \alpha = (\alpha_1 \alpha_2 \alpha_3, \dots, \alpha_M)^T, \varepsilon = (\varepsilon_1 \varepsilon_2 \varepsilon_3, \dots, \varepsilon_n)^T$$

$$\boldsymbol{B}=\begin{pmatrix} 1 & \prod_{k=1}^{K_1}\left[S_{k1}\left(x_{v(k,1)1}-t_{k1}\right)\right] & \cdots & \prod_{k=1}^{K_m}\left[S_{kM}\left(x_{v(k,M)1}-t_{kM}\right)\right] \\ 1 & \prod_{k=1}^{K_1}\left[S_{k1}\left(x_{v(k,1)2}-t_{k1}\right)\right] & \cdots & \prod_{k=1}^{K_m}\left[S_{kM}\left(x_{v(k,M)2}-t_{kM}\right)\right] \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{k=1}^{K_1}\left[S_{k1}\left(x_{v(k,1)n}-t_{k1}\right)\right] & \cdots & \prod_{k=1}^{K_m}\left[S_{kM}\left(x_{v(k,M)n}-t_{kM}\right)\right] \end{pmatrix}$$

Persamaan (2.10) dapat dijabarkan sebagai berikut.

$$\hat{f}_{(X)} = \hat{\alpha}_0 + \sum_{m=1}^{M} \alpha_m \left[S_{1m}\left(X_{v(1m)} - t_{1m}\right)\right] + \sum_{m=1}^{M} \alpha_m \left[S_{1m}\left(X_{1m} - t_{1m}\right)\right] + \left[S_{2m}\left(X_{v(2m)} - t_{2m}\right)\right] + \sum_{m=1}^{M} \alpha_m \left[S_{1m}\left(X_{v(1m)} - t_{1m}\right)\right] + \left[S_{2m}\left(X_{Vv2m)} - t_{2m}\right)\right] + \left[S_{3m}\left(X_{v(3m)} - t_{3m}\right)\right] + \cdots \tag{2.13}$$

Secara umum persamaan (2.13) dapat dituliskan sebagai berikut.

$$\hat{f}(x) = \hat{\alpha}_0 + \sum_{i=1}^{v} f_i(x_i) + \sum_{\substack{i,j=1 \\ i\neq j}}^{v} f_{ij}(x_i, x_j) + \sum_{\substack{i,j=1 \\ i\neq j}}^{v} f_{ijk}(x_i, x_j, x_k) + \ldots \tag{2.14}$$

Persamaan (2.14) menunjukkan bahwa penjumlahan suku pertama meliputi semua basis fungsi untuk satu variabel. Penjumlahan suku kedua meliputi semua basis fungsi untuk interaksi antara dua variabel. Penjumlahan suku ketiga meliputi semua basis fungsi untuk interaksi antara tiga variabel, dan seterusnya.

Pemodelan MARS ditentukan berdasarkan *trialanderror* untuk kombinasi BF, MI, dan MO untuk mendapatkan nilai GCV yang minimum (Nisa & Budiantara, 2012). Pemilihan model pada MARS dapat menggunakan metode *stepwise* (*forward* dan *backward)*. Pemilihan model dengan menggunakan *forward stepwise* dilakukan untuk mendapatkan jumlah basis fungsi maksimum, sedangkan pada *backward stepwise* dilakukan pemilihan basis fungsi yang dihasilkan dari *forwardstepwise* dengan meminimumkan nilai *Generalized Cross*

Validation (GCV). Model terbaik dipilih berdasarkan nilai GCV yang paling minimum. Fungsi GCV minimum didefenisikan sebagai berikut.

$$GCV\ (M) = \frac{\frac{1}{N}\sum_{i=1}^{N}[y_i - \hat{f}_m(x_i)^2]}{\left[1-\frac{C(M)}{N}\right]^2} \tag{2.15}$$

dengan,

N : banyaknya pengamatan

y_i : variabel respon

x_i : variabel prediktor

$\hat{f}_m(x_i)$: nilai taksiran variabel respon pada pengamatan ke-i

$c(M)$: jumlah parameter dalam model = $Trace\ [B(B^T B)^{-1} B^T] + 1$

B : matriks basis fungsi

2.7.1 Klasifikasi MARS Respon Biner

Klasifikasi pada model MARS didasarkan pada analisis regresi. Jika variabel respon terdiri dari dua nilai, maka dikatakan sebagai regresi dengan respon biner (Cox & Snell, 1989), sehingga dapat digunakan model probabilitas dengan persamaan sebagai berikut.

$$P(Z = 1\ |X = x) = \pi(x) = \frac{e^{f(x)}}{1 + e^{f(x)}}$$

dan

$$P(Z = 0\ |X = x) = 1 - \pi(x) = \frac{1}{1 + e^{f(x)}}$$

Dengan $f\ (x) = z = logit\ \pi\ (x)$. Model MARS untuk klasifikasi dapat dinyatakan sebagai berikut.

$$z = logit\ \pi\ (x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right)$$

$$= \alpha_0 + \sum_{m=1}^{M} \alpha_0 \prod_{k=1}^{Km} [S_{km}(x_{v(km)} - t_{km})] \tag{2.16}$$

2.8 Ketepatan Klasifikasi

Ketepatan klasifikasi diperlukan untuk mengetahui pengelompokkan data yang digolongkan dengan tepat pada kelompoknya. *Apparent Error Rate* (APER) didefenisikan sebagai proporsi sampel yang tidak tepat diklasifikasikan (Johnson & Wichern, 2007). Untuk mengetaui proporsi sampel yang tepat diklasifikasikan dapat dihitung dari nilai TAR *(Total Accuracy Rate)*. Berikut ini merupakan tabel pengklasifikasian untuk respon biner.

Tabel 2.1 Klasifikasi Respon Biner

Observasi	Taksiran Observasi	
	y_0	y_1
y_0	n_{00}	n_{01}
y_1	n_{10}	n_{11}

Nilai APER dan TAR didapatkan dengan perhitungan sebagai berikut.

$$APER(\%) = \frac{n_{10} + n_{01}}{n} X\ 100\% \tag{2.17}$$

dan

$$TAR\ (\%) = 1 - APER = 1 - \left(\frac{n_{10} + n_{01}}{n} X\ 100\%\right) \tag{2.18}$$

dengan,

n = jumlah observasi

n_{00} = jumlah observasi dari y_0 yang tepat diklasifikasikan sebagai y_0

n_{11} = jumlah observasi dari y_1 yang tepat diklasifikasikan sebagai y_1

n_{01} = jumlah observasi dari y_0 yang salah diklasifikasikan sebagai y_1

n_{10} = jumlah observasi dari y_1 yang salah diklasifikasikan sebagai y_0