

BAB II

TINJAUAN PUSTAKA

Pada Bab II akan dibahas mengenai konsep landasan teori yang menunjang penelitian ini yaitu kadar trombosit pada pasien DBD menggunakan *Generalized Semiparametric Regression Spline* pada data longitudinal. Landasan teori yang akan dibahas adalah sebagai berikut.

2.1 Analisis Regresi

Analisis regresi adalah metode statistika yang memberikan penjelasan mengenai pola hubungan antara dua variabel atau lebih (Draper & Smith, 1992). Dalam membentuk sebuah model dengan satu atau lebih variabel harus dilakukan pengecekan apakah variabel-variabel secara rasional memiliki korelasi atau tidak. Apabila memiliki korelasi, maka dapat dilakukan pemodelan menggunakan analisis regresi.

2.2 Regresi Parametrik

Hardle (dalam Toruan, 2018) mendefinisikan regresi parametrik merupakan pendekatan regresi yang sering digunakan, dimana asumsi yang mendasari yaitu bentuk kurva regresi digambarkan oleh sekumpulan parameter tertentu atau dapat digambarkan dalam pola tertentu. Pemilihan bentuk kurva yang dipilih akan berpengaruh pada ketepatan analisis parametrik.

2.3 Regresi Nonparametrik

Menurut Eubank (dalam Toruan, 2018) regresi nonparametrik merupakan salah satu pendekatan yang biasanya digunakan untuk mengetahui pola hubungan antara variabel penjelas dan variabel respon dimana tidak diketahui bentuk kurva regresinya atau tidak terdapat informasi masa lampau tentang bentuk pola dari data yang lengkap. Secara umum regresi nonparametrik menurut Wahba (dalam Toruan, 2018) memiliki bentuk sebagai berikut.

$$y = f(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Dengan y_i merupakan variabel respon sedangkan fungsi $f(t_i)$ adalah kurva regresi dengan t_i sebagai variabel prediktor dan ε_i adalah error random yang diasumsikan berdistribusi normal independen dimana mean nol dan variansi σ^2 .

2.4 Regresi Semiparametrik

Menurut Sugiantari *et al.* (2013) regresi semiparametrik digunakan untuk mengetahui pola hubungan antara variabel respon dan variabel prediktor dimana sebagian pola diketahui bentuknya dan sebagian pola lagi tidak diketahui bentuknya.

Bentuk umum model regresi semiparametrik sebagai berikut.

$$y_i = \sum_{k=1}^q \beta_k x_{ik} + \sum_{r=1}^R f_r(t_{ri}) + \varepsilon_i$$

dengan $i = 1, 2, \dots, n$; $k = 1, 2, \dots, q$; $r = 1, 2, \dots, R$

2.5 Spline

Spline merupakan salah satu pendekatan dalam nonparametrik yang dideskripsikan sebagai potongan polinomial tersegmen yang memiliki sifat fleksibilitas, dimana hal tersebut yang menjadi pembeda spline dengan polinomial. Titik perpaduan bersama dari potongan-potongan atau interval-interval yang berbeda dan sering disebut sebagai knot.

Eubank (dalam Suparti *et al.*, 2017)) Secara umum, fungsi spline berorde m adalah sembarang fungsi yang ditulis sebagai berikut:

$$f(t) = \sum_{s=1}^m \alpha_s t^s + \sum_{u=1}^U \theta_u (t - K_u)_+^m \quad (1)$$

dimana $\sum_{s=1}^m \alpha_s t^s$ adalah komponen polinomial dan $\sum_{u=1}^U \theta_u (t - K_u)_+^m$ adalah komponen *truncated*, dengan

$$(t - K_u)_+^m = \begin{cases} (t - K_u)^m & ; t \geq K \\ 0 & ; t < K \end{cases}$$

Diketahui α_s dan θ_u adalah parameter serta K_1, K_2, \dots, K_U adalah titik knot, $f(t)$ adalah fungsi regresi, serta t adalah variabel prediktor.

Apabila $m=1$ dan banyaknya knot adalah satu, maka persamaan (1) dapat ditulis kedalam fungsi *spline linear* menjadi:

$$f(t) = \alpha_1 t + \theta_2 (t - K)_+^1 \quad (2)$$

Fungsi (2) dinamakan *spline linear* satu knot pada $t=K$. Fungsi (2) dapat pula ditulis (Budiantara, 2005), menjadi:

$$f(t) = \begin{cases} \alpha_1 t & ; t < K \\ \alpha_1 t + \theta_2 (t - K) & ; t \geq K \end{cases}$$

Model fungsi *spline* dapat ditulis dalam bentuk:

$$y_i = \sum_{s=1}^m \alpha_s t_i^s + \sum_{u=1}^U \theta_u (t_i - K_u)_+^m + \varepsilon_i \quad (3)$$

Selanjutnya model (3) dapat ditulis dalam bentuk :

$$y_i = \alpha_i t_i + \dots + \alpha_m t_i^m + \theta_i (t_i - K_1)_+^m + \dots + \theta_U (t_i - K_U)_+^m + \varepsilon_i \quad (4)$$

Apabila model regresi diubah dalam bentuk matriks, menjadi:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} t_1 & \dots & t_1^m & (t_1 - K_1)_+^m & \dots & (t_1 - K_U)_+^m \\ t_2 & \dots & t_2^m & (t_2 - K_1)_+^m & \dots & (t_2 - K_U)_+^m \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ t_n & \dots & t_n^m & (t_n - K_1)_+^m & \dots & (t_n - K_U)_+^m \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \\ \theta_1 \\ \vdots \\ \theta_U \end{pmatrix}$$

$$+ \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Selanjutnya, persamaan (4) dapat ditulis menjadi:

$$\tilde{y} = X(K_1, \dots, K_U) \tilde{\beta} + \tilde{\varepsilon}$$

dengan,

$$\tilde{y} = (y_1, y_2, \dots, y_n)',$$

$$X(K_1, \dots, K_U) = \begin{bmatrix} t_1 & \dots & t_1^m & (t_1 - K_1)_+^m & \dots & (t_1 - K_U)_+^m \\ t_2 & \dots & t_2^m & (t_2 - K_1)_+^m & \dots & (t_2 - K_U)_+^m \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ t_n & \dots & t_n^m & (t_n - K_1)_+^m & \dots & (t_n - K_U)_+^m \end{bmatrix},$$

$$\tilde{\beta} = (\alpha_1, \dots, \alpha_m, \theta_1, \dots, \theta_U)' \text{ dan } \tilde{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$$

2.6 Data Longitudinal

Data longitudinal adalah data dari hasil pengukuran yang berulang untuk satu atau beberapa variabel setiap anggota sejumlah subjek atau individu yang sama, dimana diamati pada sejumlah titik di waktu yang berbeda. Misalkan t_{ij} adalah waktu pengamatan untuk setiap subjek ke- i pada waktu ke- j , y_{ij} adalah variabel respon yang diamati pada waktu. Wu dan Zhang (dalam Suparti *et al.*, 2017) mengatakan bahwa pada data longitudinal dimana subjek-subjek diasumsikan saling independen satu sama lain, namun terdapat korelasi antar pengamatan didalam subjek yang saling dependen. Analisis data longitudinal bertujuan untuk mempelajari mengenai bagaimana perubahan subjek yang diamati dari waktu ke waktu. Bentuk struktur data longitudinal adalah sebagai berikut.

Tabel 2.1 Struktur Data Longitudinal

Subjek ke- j	Pengamatan ke- k	Prediktor	Respon
Subjek ke- 1	1	x_{11}	y_{111}
	2	x_{12}	y_{112}
	⋮	⋮	⋮
	n	x_{1n}	y_{11n}
Subjek ke- 2	1	x_{21}	y_{121}
	2	x_{22}	y_{122}
	⋮	⋮	⋮
	n	x_{2n}	y_{12n}
⋮	⋮	⋮	⋮
Subjek ke- m	1	X_{m1}	y_{1m1}
	2	X_{m2}	y_{1m2}
	⋮	⋮	⋮
	n	X_{mn}	y_{1mn}

2.7 Model *Generalized Semiparametric Regression Spline* pada Data

Longitudinal

Metode analisis regresi biasanya akan menggunakan *Least square* n untuk mengestimasi parameter. Pada data *cross section* dan data longitudinal terdapat perbedaan penggunaan *Least Square*, dimana estimasi parameter yang digunakan pada data *cross section* adalah *Ordinary Least Square* (OLS). Sedangkan untuk estimasi parameter pada data longitudinal adalah *Weighted Least Square* (WLS), hal ini dikarenakan pada data longitudinal korelasi antar pengamatan dalam subjek yang sama harus dipertimbangkan. Metode *weighted least square* (WLS) mendefinisikan mengenai estimasi parameter sebagai suatu nilai yang meminimumkan jumlah kuadrat antara pengamatan dan model yang biasanya disebut jumlah kuadrat *error*, dituliskan dengan rumus sebagai berikut:

$$\boldsymbol{\varepsilon}^T \mathbf{V}^{-1} \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

Jika diberikan data observasi y_{ij}, x_{kij}, t_{rij} maka model *generalized semiparametric regression spline* untuk data longitudinal adalah sebagai berikut yang memuat sebanyak n subjek dengan subjek ke- i sebanyak p pengamatan:

$$y_{ij} = \sum_{k=1}^q \boldsymbol{\beta}_{kj} x_{kij} + \sum_{r=1}^R f_r(t_{ij}) + \varepsilon_{ij}$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, p$$

Selanjutnya mendekati $f_r(t_{ij})$ dengan fungsi *spline* berorde m dan knot

$K_{11}, K_{12}, \dots, K_{1U}, K_{21}, K_{22}, K_{2U}, \dots, K_{R1}, K_{R2}, K_{RU}$ yaitu :

$$f_r(t_{rij}) = \sum_{s=1}^m \alpha_{rs} t_{rij}^s + \sum_{u=1}^U \theta_u (t_{rij} - K_{ru})_+^m$$

Dimana $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1m}, \alpha_{21}, \alpha_{22}, \dots, \alpha_{Rm}, \theta_1, \theta_2, \dots, \theta_U$ adalah parameter-parameter yang belum diketahui serta *spline function* $(t_{rij} - K_{ru})_+^m$

didefinisikan sebagai:

$$(t_{rij} - K_{ru})_+^m = \begin{cases} (t_{rij} - K_{ru})^m, & t_{rij} \geq K_{ru} \\ 0, & t_{rij} < K_{ru} \end{cases}$$

Bentuk model *generalized semiparametric regression spline* untuk data longitudinal adalah sebagai berikut.

$$y_{ij} = \sum_{k=1}^q \beta_{kj} x_{kij} + \sum_{r=1}^R \left(\sum_{s=1}^m \alpha_{rs} t_{rij}^s + \sum_{u=1}^U \theta_u (t_{rij} - K_{ru})_+^m \right) + \varepsilon_{ij}$$

Berdasarkan penjabaran diatas, maka persamaan-persamaan tersebut dapat disajikan dalam bentuk matriks berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1[K] & 0 & \dots & 0 \\ 0 & \mathbf{X}_2[K] & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \mathbf{X}_p[K] \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

Model regresi dapat dituliskan sebagai berikut:

$$Y = X[K]B + e$$

Pembobot digunakan untuk mengestimasi β yang diperoleh dengan menyelesaikan optimasi *weighted least square* dibawah ini:

$$\min_{B,f}\{(Y - X[K]\beta)^T V^{-1}(Y - X[K]\beta)\} \quad (6)$$

Metode *generalized regression spline* merupakan salah satu metode regresi *spline* untuk data longitudinal. Metode ini bertujuan untuk mengestimasi fungsi rata-rata populasi dari kumpulan data longitudinal. *Generalized regression spline* menyatakan korelasi data longitudinal dalam pembobot, dimana $V = \text{cov}(y) = \text{diag}(V_1, V_2, \dots, V_n)$ dan merupakan matriks yang berdiagonal 1 per pengamatan untuk setiap subjek yang diamati (Wu dan Zhang, 2006:127).

2.8 Pemilihan Titik Knot Optimal pada Data Longitudinal

Dalam regresi nonparametrik *spline*, pemilihan titik knot yang optimal adalah hal yang sangat penting. Metode yang biasanya digunakan dalam memilih titik knot yang optimal adalah *Generalized Cross Validation* (GCV). Nilai GCV terkecil akan menghasilkan titik knot optimal dalam menentukan model *spline* terbaik (Budiantara, 2005).

Fungsi GCV untuk data longitudinal adalah sebagai berikut.

$$GCV(K) = \frac{N^{-1}Y^T(I - A(K))^T V^{-1}(I - A(K))Y}{(N^{-1}\text{trace}(I - A(K)))^2}$$

2.9 Mean Square Error

Menurut Suparti (2013), bentuk estimator dari model *spline* dipengaruhi oleh banyaknya titik knot. Parameter K adalah pengontrol dalam melihat kelulusan kurva dan keseimbangan antara kesesuaian

kurva terhadap data sehingga pemilihan parameter K yang optimal menjadi berkaitan dengan pemilihan optimasi yang optimal. Pemilihan K yang optimal harus memenuhi beberapa kriteria tertentu misalnya nilai *Mean Square Error* (MSE) yang minimum. Perhitungan nilai MSE adalah sebagai berikut:

$$MSE = N^{-1}Y^T(I - A(K))^T V^{-1}(I - A(K))Y$$

2.10 Koefisien Determinasi

Menurut Gujarati (dalam Ariana, 2018), koefisien determinasi atau biasanya dilambangkan dengan R^2 menyatakan sebagai koefisien determinasi majemuk yang dikenal untuk mengetahui proporsi variasi dalam y yang dijelaskan beberapa variabel x secara bersama-sama (gabungan). Nilai koefisien terletak antara nol dan satu. Jika nilai koefisien determinasi sama dengan satu, maka model semakin baik dalam menjelaskan variasi dalam y . Namun, sebaliknya jika nilai koefisien determinasi sama dengan nol, maka model tidak sedikitpun menjelaskan variasi dalam y . Kecocokan model lebih baik dalam menjelaskan variabel y bila nilainya semakin dekat dengan satu.

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

Dimana $\sum(y_i - \bar{y})^2 = \sum(\hat{y}_i - \bar{y})^2 + \sum(y_i - \hat{y}_i)^2$

2.11 Demam Berdarah *Dengue* (DBD)

Penyakit DBD merupakan penyakit menular yang ditularkan oleh nyamuk *Aedes aegypti* yang membawa virus dengue. Departemen

Kesehatan Republik Indonesia (dalam Wati, 2009) menyebutkan bahwa penyakit demam berdarah *dengue* ditandai dengan adanya demam mendadak selama 2 sampai 7 hari tanpa penyebab jelas, badan terasa lemah atau lesu, gelisah, terasa nyeri pada ulu hati dan disertai tanda perdarahan dikulit berupa bintik merah perarahan, lebam atau ruam. Kadang-kadang juga terjadi mimisan, berak darah, muntah darah, dan kesadaran menurun atau shock.

Virus dengue yang menyebabkan Penyakit Demam Berdarah *Dengue* (DBD) dan Penyakit Demam *Dengue* (DD), dimana virus tersebut termasuk dalam kelompok B *Arthropod Borne Virus* (*Arboviroses*) yang pada saat ini dikenal sebagai genus *Flavivirus*, famili *Flaviviricae*, dan memiliki 4 jenis serotipe yaitu : DEN-1, DEN-2, DEN-3, DEN-4. Antibodi terhadap salah satu serotipe ditimbulkan oleh infeksi yang bersangkutan, sedangkan antibodi yang terbentuk terhadap serotipe lain sangat kurang, sehingga perlindungan yang diberikan tidak dapat memadai terhadap serotipe lain. Serotipe DEN-3 adalah serotipe yang dominan dan diasumsikan banyak yang menunjukkan manifestasi klinik yang berat (Hadinegoro *et al.*, 2001).

Anggraeni (dalam Tobing, 2011) mengatakan bahwa nyamuk *Aedes aegypti* mempunyai kemampuan terbang yang bisa mencapai radius 100-200 meter. Oleh karena itu, masyarakat harus waspada jika dalam radius tersebut ada kasus DBD. Tobing (dalam Kurniadi, 2018) mengatakan nyamuk *Aedes aegypti* menyukai tempat yang lembab, gelap

dan berbau seperti kawasan yang padat dimana keadaan sanitasi yang kurang memadai, digenangan air dalam rumah, seperti pada vas bunga, bak mandi, pot ataupun tempat penyimpanan air lain seperti pada ember plastik, tempayan dan drum. Masa inkubasi virus *Dengue* dalam manusia (inkubasi intrinsik) yang berkisar antara 3 sampai dengan 14 hari sebelum gejala muncul, untuk gejala klinis yang rata-ratanya muncul pada hari keempat sampai dengan hari ketujuh, sedangkan masa inkubasi didalam tubuh nyamuk (ekstrinsik) dapat berlangsung sekitar 8 sampai 10 hari (Taryono et al., 2018).

Secara klinis, tahapan penyakit DBD dibagi menjadi 3 fase, yaitu fase febrile (demam), fase kritis, dan fase penyembuhan. Menurut Arif dan Pudijadi (dalam Yuniarti et al., 2017) tahapan penyakit demam berdarah *dengue* mempengaruhi dinamika pada perubahan tanda dan gejala klinis terhadap pasien DBD. Lestari (dalam Candra, 2010) menyebutkan mengenai manifestasi klinis yang dimulai dari infeksi tanpa gejala demam, demam *dengue* (DD) dan DBD, yang ditandai dengan adanya demam tinggi secara terus menerus selama 2-7 hari; perdarahan diatesis seperti uji tourniquet yang positif, trombositopenia dengan jumlah trombosit $\leq 100 \times 10^9/L$ dan juga adanya kebocoran plasma sebagai akibat dari peningkatan permeabilitas pembuluh. Pemeriksaan hematologi adalah salah satu pemeriksaan penunjang yang mana terdiri atas pemeriksaan jumlah trombosit, jumlah leukosit, nilai hematokrit dan kadar hemoglobin.

WHO pada tahun 2011 mengajukan trombositopenia dan hemokonsentrasi sebagai kriteria laboratoris pada diagnosis klinis penyakit DBD.

World Health Organization (dalam Rasyada et al, 2013) menyatakan bahwa ada dua perubahan patologik yang utama pada penyakit DBD. Pertama, terjadinya peningkatan permeabilitas kapiler, dimana dapat menyebabkan volume plasma pada pembuluh darah berkurang sehingga akan terjadi hemokonsentrasi (*World Health Organization*, 1997 dan Nimmamnitya, 1999 dalam (Rasyada et al., 2013)). Harrison (dalam (Hukom *et al.*, 2013) mengatakan bahwa gangguan hemostasis menyebabkan terjadinya manifestasi klinis perdarahan. Teori *Secondary Heterologous Infection* menyatakan peningkatan agregasi trombosit dapat mengakibatkan penghancuran trombosit oleh *Reticuloendothelial System* (RES) sehingga akan terjadi trombositopenia.

Derajat atau *grade* penyakit DBD dapat dikategorikan berdasarkan kriteria dari *World Health Organization* (dalam Widiyantari *et al*, 2012) . Derajat 1 yaitu demam disertai gejala yang tidak khas dan satu-satunya manifestasi perdarahan ialah uji *torniquet*. Derajat II tidak berbeda dengan derajat I, namun disertai juga dengan perdarahan spontan di kulit atau perdarahan lain. Derajat III yaitu terdapat kegagalan sirkulasi yaitu nadi cepat dan lambat, tekanan nadi menurun (20mmHG atau kurang) atau hipotensi, sianosis, kulit dingin dan lembab serta penderita gelisah. Derajat

IV yaitu pasien mengalami syok berat, nadi tidak dapat diraba dan tekanan darah tidak terukur.

2.12 Trombosit dan Hemoglobin pada Pasien DBD

Trombosit adalah sel darah merah yang berfungsi dalam hemostasis. Sloane (dalam Rasyada et al., 2013) mengatakan mengenai ciri dari trombosit adalah tidak adanya nukleus dan dihasilkan dalam sumsum tulang oleh megakariosit. Menurut Santoso (dalam Hukom *et al.*, 2013), bahwa nilai jumlah trombosit dapat dipergunakan sebagai faktor untuk memprediksi kejadian syok. Fungsi trombosit dipengaruhi oleh sebuah unsur yang ada pada sel darah merah, yaitu *adenosine diphosphate* (ADP) yang mana dikenal sebagai salah satu agonis agregasi trombosit. Goldsmith (dalam Hukom *et al.*, 2013) mengatakan mengenai peningkatan agregasi trombosit yang terjadi sebagai akibat dari pelepasan ADP oleh sel darah merah.

Davis (dalam Wahyuni, 2017) menyebutkan peningkatan kadar hemoglobin yang meningkat menunjukkan adanya kebocoran plasma dan banyaknya sel darah merah di dalam pembuluh darah, hal ini mengindikasikan akan adanya infeksi *dengue* dengan tanda bahaya yang meningkatkan resiko terjadinya *dengue shock syndrome*. Penderita demam berdarah *dengue* biasanya mengalami masalah mengenai sistem peredaran darah yang diserang oleh virus *dengue* dari nyamuk. Reaksi imunologis antara virus dan sistem pertahanan (imun) tubuh menyebabkan dinding pembuluh

darah menjadi mudah ditembus cairan atau plasma darah. Seiring terjadinya perembesan tersebut, kekentalan darah akan meningkat karena kekurangan plasma. Apabila perembesan terjadi secara terus menerus, maka akan menyebabkan jumlah trombosit dalam darah mengalami penurunan.

