

Hubungan Faktor Kolesterol Terhadap Gula Darah Diabetes dengan Spline Kubik Terbobot

Zhazha Alifkhamulki Ramdhani^{1*}, Anna Islamiyati², Raupong³
^{1,2,3}Departemen Statistika, Fakultas MIPA,
Universitas Hasanuddin, Makassar, 90245, Indonesia

* Corresponding author, email: zhazhaalifkha@gmail.com

Abstract

Diabetes Mellitus (DM) is often recognized through an increase in a person's blood sugar level. Factors that can affect the increase in blood sugar levels of DM patients one of which is cholesterol. It usually contains the bookkeeping of several types of cholesterol, including LDL and total cholesterol. DM data are assumed to experience heterokedasticity so that in this study analyzed using regression of weighted cubic spline nonparametric. The estimation method used is weighted least square (WLS). This study aims to obtain a weighted cubic spline model on cholesterol based DM data. The selection of the best model can be seen based on the criteria for the value of generalized cross validation (GCV) minimum. Based on the analysis obtained weighted cubic spline models for cholesterol factors for blood sugar as follows:

$$\hat{y} = 1,59 + 1,51(x_1) - 6,51(x_1)^2 - 6,32(x_1)^3 + 3,78(x_1 - 59)^3_+ - 3,23(x_1 - 81)^3_+ - 5,00(x_2) + 5,09(x_2)^2 - 6,51(x_2)^3 + 1,72(x_2 - 99)^3_+$$

Keywords: Diabetes Mellitus, Heteroscedasticity, LDL, Cubic Spline, Weighted.

Abstrak

Diabetes Mellitus (DM) sering dikenali melalui peningkatan kadar gula darah seseorang. Faktor yang dapat mempengaruhi naiknya kadar gula darah pasien DM salah satunya kolesterol. Itu biasanya memuat pengukuran beberapa jenis kolesterol, diantaranya kolesterol LDL dan total. Pada data DM diasumsikan mengalami heterokedastisitas sehingga dalam penelitian ini dianalisis menggunakan regresi nonparametrik spline kubik terboboti. Metode estimasi yang digunakan yaitu *weighted least square* (WLS). Penelitian ini bertujuan untuk memperoleh model spline kubik terboboti pada data DM berdasarkan kolesterol. Pemilihan model terbaik dapat dilihat berdasarkan kriteria nilai *generalized cross validation* (GCV) minimum. Berdasarkan analisis diperoleh model spline kubik terbobot untuk faktor kolesterol terhadap gula darah sebagai berikut:

$$\hat{y} = 1,59 + 1,51(x_1) - 6,51(x_1)^2 - 6,32(x_1)^3 + 3,78(x_1 - 59)^3_+ - 3,23(x_1 - 81)^3_+ - 5,00(x_2) + 5,09(x_2)^2 - 6,51(x_2)^3 + 1,72(x_2 - 99)^3_+$$

Kata Kunci: Diabetes Melitus, Heterkedastisitas, LDL, Spline Kubik, Terbobot.

1. Pendahuluan

Analisis regresi merupakan salah satu teknik analisis statistika yang digunakan untuk melihat pola hubungan pasangan data antara variabel prediktor dengan variabel respon. Terdapat dua pendekatan estimasi model dalam analisis regresi, yaitu regresi parametrik dan regresi nonparametrik. Pendekatan parametrik digunakan ketika bentuk fungsi data yang diperoleh menunjukkan suatu pola hubungan yang mudah digambarkan dengan fungsi tertentu seperti linier, kuadratik, dan sebagainya, namun jika fungsi data yang diperoleh tidak menunjukkan pola hubungan yang mudah digambarkan dengan

fungsi tertentu maka estimasi fungsi regresi dilakukan dengan menggunakan pendekatan nonparametrik. Regresi nonparametrik memiliki fleksibilitas yang tinggi karena data dapat mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti [1]. Salah satu estimator yang sering digunakan yaitu estimator spline yang terdiri dari beberapa bentuk, diantaranya spline truncated [2], spline smoothing [3] dan spline penalized [4]. Spline merupakan jenis potongan polinomial yang tersegmen sehingga memberikan fleksibilitas yang tinggi dibandingkan dengan model polinomial pada umumnya [5].

Analisis regresi terbobot digunakan ketika terjadi kondisi heterokedastisitas. Heterokedastisitas terjadi akibat adanya ketidaksamaan nilai variansi *error* [6]. Beberapa penelitian telah dilakukan dalam menangani kasus heterokedastisitas baik parametrik maupun nonparametrik. Maziyya (2015) menggunakan regresi linear berganda terboboti dengan pembobot $1/x$ [7]. Hendayanti (2015) mengatasi kurva regresi nonparametrik pada kasus heterokedastisitas dengan pembobot $1/y$ [8]. Islamiyati, Fatmawati dan Chamidah (2018, 2019) menggunakan pembobot matriks variansi kovariansi [9,10].

Selanjutnya, penelitian ini mengkaji hubungan faktor kolesterol terhadap gula darah menggunakan spline kubik terbobot dengan pembobot variansi dan melibatkan dua prediktor. Islamiyati (2019) telah mengembangkan penggunaan spline biprediktor tanpa pembobot pada kasus linier dan diaplikasikan pada data DBD [11]. Adapun model ini diterapkan pada data diabetes mellitus. Diabetes mellitus merupakan penyakit kronis dan telah menjadi penyebab kematian terbesar ke empat di dunia. Tanda awal seseorang terkena diabetes dapat dilihat dari peningkatan kadar gula darah [12]. Menurut WHO, Diabetes Melitus (DM) didefinisikan sebagai suatu penyakit atau gangguan metabolisme kronis dengan multi etiologi yang ditandai dengan tingginya kadar gula darah disertai dengan gangguan metabolisme kaborhidrat, lipid dan kolesterol. Islamiyati dkk (2019) mempolakan kadar gula darah setelah 2 jam makan selama perawatan [13]. Adapun Islamiyati, Raupong dan Anisa (2019) meneliti hubungan berat badan terhadap kenaikan gula darah [14]. Hasil penelitian ini menunjukkan bahwa terdapat 3 pola perubahan gula darah pasien DM berdasarkan kolesterol LDL dan 2 pola perubahan berdasarkan kolesterol total. Perubahan pola gula darah dari hasil estimasi kubik menunjukkan adanya pengaruh yang berbeda pada interval kolesterol tertentu terhadap naik turunnya gula darah pasien diabetes.

2. Material dan Metode

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data penderita diabetes melitus tipe II pada tahun 2014-2018 di Rumah sakit pendidikan Universitas Hasanuddin. Variabel dalam penelitian ini menggunakan satu variabel respon (y) dan dua variabel prediktor (x_1 dan x_2). Gula darah (mg/DL) merupakan variabel respon (y), x_1 adalah kolesterol LDL (mg/DL), dan x_2 adalah kolesterol total (mg/DL).

Regresi nonparametrik adalah suatu metode yang digunakan untuk mengetahui pola hubungan perilaku data ketika ketersediaan informasi mengenai bentuk kurva hanya sedikit. Berbeda dengan regresi parametrik, dalam regresi nonparametrik data akan mencari sendiri bentuk estimasi dari kurva regresi tanpa harus dipengaruhi oleh faktor subjektifitas peneliti. Oleh karena itu, dalam regresi nonparametrik digunakan pendekatan yang sangat adaptif dengan persamaan regresi yang ditentukan dari data [3]. Model regresi secara umum dapat ditulis sebagaimana persamaan :

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{ki}) + \varepsilon_i ; i = 1, 2, \dots, n. \quad (1)$$

Model regresi secara umum pada persamaan (1) didekati dengan fungsi regresi nonparametrik estimator spline kubik, seperti pada persamaan dibawah ini:

$$f(x_{ji}) = \sum_{l=0}^3 \beta_{lj} x_{ji}^l + \sum_{h=1}^r \beta_{j(p+h)} (x_{ji} - K_{hj})_+^3 \quad (2)$$

dengan:

$$(x_{ji} - k_{hj})_+^3 = \begin{cases} (x_{ji} - k_{hj})^3, & x_{ji} \geq k_{hj} \\ 0 & , x_{ji} < k_{hj} \end{cases}$$

Selanjutnya, akibat asumsi heterokedastisitas sehingga dibutuhkan pembobot dalam proses estimasi, Proses estimasi yang akan digunakan yaitu *Weighted Least Square* (WLS). Berdasarkan prosedur yang sama pada OLS, metode ini mendefinisikan estimasi parameter sebagai suatu nilai yang meminimumkan jumlah kuadrat *error* yang didefinisikan sebagai berikut [15]:

$$(\boldsymbol{\varepsilon}'\mathbf{W}\boldsymbol{\varepsilon}) = ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})) \quad (3)$$

Misalkan fungsi $\mathbf{Q}(\boldsymbol{\beta}) = ((\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'\mathbf{W}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}))$, kemudian didiferensialkan terhadap $\boldsymbol{\beta}$ dan menyamakan dengan 0, diperoleh hasil estimasi parameter sebagai berikut:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}'\mathbf{W}\mathbf{Y}) \quad (4)$$

Akibatnya, estimasi fungsi regresi spline kubik terbobot diperoleh seperti pada persamaan (5) dan selanjutnya diterapkan pada analisis data diabetes melitus:

$$\hat{f}(x_{ji}) = \sum_{l=0}^3 \beta_{lj} x_{ji}^l + \sum_{h=1}^r \beta_{j(p+h)} (x_{ji} - K_{hj})_+^3 \quad (5)$$

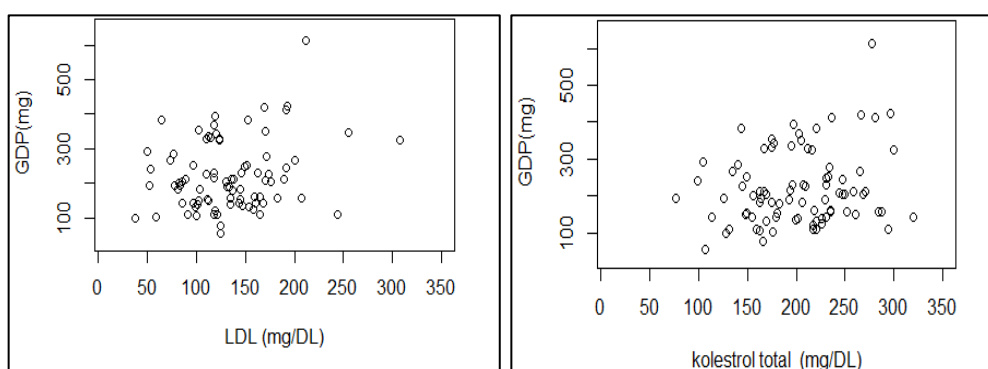
Adapun langkah-langkah analisis yang digunakan pada data diabetes melitus sebagai berikut:

1. Langkah pertama dalam memodelkan gula darah yang mengandung heterokedastisitas menggunakan spline kubik, dengan membuat scatter plot untuk melihat pola perubahan gula darah pasien berdasarkan kolesterol LDL dan kolesterol total .
2. Langkah kedua dilakukan pemodelan gula darah pasien terhadap kolesterol LDL menggunakan spline kubik dengan bantuan titik knot.
3. Langkah ketiga melakukan *Uji heterokedastistas*.

- Langkah keempat menghitung pembobot variansi yang akan digunakan.
- Langkah kelima dilakukan pemodelan gula darah pasien terhadap kolesterol LDL dan kolesterol total menggunakan spline polynomial truncated terboboti dengan bantuan titik knot.
- Langkah keenam mendapatkan model spline polynomial truncated terboboti optimal dengan melihat nilai GCV yang minimum. Langkah terakhir yaitu menginterpretasikan model yang diperoleh

3. Hasil dan Diskusi

Tahapan pemodelan kadar gula darah terhadap kolesterol LDL dan kolesterol total, dimulai dengan membuat *scatter plot* untuk melihat pola hubungan antara kadar gula darah dengan kolesterol LDL dan kolesterol total yang ditunjukkan pada Gambar 1.



Gambar 1. *Scatter plot* variabel respon dan prediktor pada data penderita Diabetes Melitus

Hasil plot pada Gambar 1 terlihat bahwa tidak ada kecenderungan pola parametrik yang terjadi pada hubungan antara gula darah dengan kolesterol LDL begitupun dengan kolesterol total. Oleh karena itu, pendekatan yang cocok digunakan pada data adalah pendekatan nonparametrik.

3.1 Memodelkan Kadar Gula Darah Pasien Diabetes Terhadap Kolesterol LDL dan Kolesterol Total Menggunakan Spline Kubik

Pemodelan spline kubik dilakukan dengan titik knot optimal. Titik knot optimal dipilih berdasarkan nilai GCV minimum, pemilihan titik knot menggunakan metode *trial and error* pada nilai koletsrol LDL dan nilai kolesterol total. Berikut nilai GCV minimum untuk titik knot optimal dapat dilihat pada Tabel 1. Dari hasil Tabel 1 diperoleh GCV minimum sebesar 11.085,3 terletak pada nilai kolesterol LDL sebesar 59 mg/DL dan nilai kolesterol total 149 mg/DL. Artinya terjadi perubahan kadar gula darah ketika nilai kolesterol LDL pasien berada pada nilai 59 mg/DL dan nilai kolesterol total pasien berada pada 149 mg/DL. Model spline kubik dengan satu titik knot:

$$y = \beta_0 + \beta_{11}(x_1) + \beta_{21}(x_1)^2 + \beta_{31}(x_1)^3 + \beta_{1(4)}(x_1 - k_{11})^3_+ + \beta_{12}(x_2) + \beta_{22}(x_2)^2 + \beta_{32}(x_2)^3 + \beta_{2(4)}(x_2 - k_{12})^3_+$$

Tabel 1. Nilai GCV minimum untuk titik knot optimal

No	Titik knot		Nilai GCV
	x_1	x_2	
1	59	99	11.159,0
2	59	113	11.237,1
3	59	149	11.085,3
4	77	128	11.228,9
5	77	131	11.178,5

Sumber: Data diolah 2020

Selanjutnya, hasil estimasi parameter dapat dilihat pada Tabel 2.

Tabel 2. Hasil estimasi spline kubik dengan knot optimal

Parameter	Estimasi
β_0	-1,52
β_{11}	6,98
β_{21}	-1,19
β_{31}	6,81
$\beta_{1(4)}$	-6,82
β_{12}	4,28
β_{22}	-2,79
β_{23}	6,15
$\beta_{2(4)}$	-6,07

Sumber: Data diolah 2020

Berdasarkan estimasi parameter pada Tabel 2, maka diperoleh stimasi model spline kubik secara simultan dengan knot optimal sebagai berikut:

$$\hat{y} = -1,52 + 6,98(x_1) - 1,19(x_1)^2 + 6,81(x_1)^3 - 6,82(x_1 - 59)^3_+ + 4,28(x_2) - 2,79(x_2)^2 + 6,15(x_2)^3 - 6,07(x_2 - 149)^3_+$$

Selanjutnya uji asumsi residual identik menggunakan uji gletjser. Hipotesis dari Uji *Glejser* :

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$$

$$H_1: \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2, \text{ dengan } i = 1, 2, \dots, n$$

Pada pengujian residual identik dengan menggunakan uji *Glejser* diperoleh $p\text{-value} = 0,005$ dengan $\alpha = 0,05$, sehingga $p\text{-value} < \alpha$ yang berarti tolak H_0 dan dapat disimpulkan bahwa residual bersifat heterokedastisitas atau variansi *error* berbeda. Karena terjadi heterokedastisitas maka akan dilakukan pembobotan, pembobot yang akan digunakan yaitu $\mathbf{W} = \sigma_\varepsilon^2$. Selanjutnya dilakukan uji gletjser untuk hasil analisis pembobot yang memberikan $p\text{-value} = 0,26$ untuk $\alpha = 0,05$ sehingga $p\text{-value} > \alpha$. Hal ini

berarti bahwa keputusan terima dan dapat disimpulkan bahwa asumsi residual homoskedastisitas terpenuhi.

3.2 Memodelkan Kadar Gula Darah Pasien Diabetes terhadap Kolesterol LDL dan Kolesterol Total Menggunakan Spline Kubik Terboboti

Pemodelan spline kubik terboboti juga dilakukan dengan pemilihan knot optimal terboboti berdasarkan nilai GCV minimum. Berikut nilai GCV minimum berdasarkan knot optimal terboboti pada Tabel 3.

Tabel 3. Nilai GCV minimum untuk titik knot optimal terboboti

No	Titik Knot			GCV
	x_1		x_2	
1	50	77	76	10.298,6
2	59	81	99	10.216,8
3	78	110	113	10.547,0
4	81	110	135	10.567,3
5	81	121	155	10.585,9

Sumber: Data diolah 2020

Berdasarkan Tabel 3 diperoleh GCV minimum sebesar 10.216,8 terletak pada nilai kolesterol LDL sebesar 59 mg/DL dan 81 mg/DL sedangkan untuk nilai kolesterol total terletak pada nilai 99 mg/DL.

Model spline kubik terboboti dengan knot optimal:

$$y = \beta_0 + \beta_{11}(x_1) + \beta_{21}(x_1)^2 + \beta_{31}(x_1)^3 + \beta_{1(4)}(x_1 - k_{11})^3_+ + \beta_{1(5)}(x_1 - k_{21})^3_+ + \beta_{12}(x_2) + \beta_{22}(x_2)^2 + \beta_{32}(x_2)^3 + \beta_{2(4)}(x_2 - k_{12})^3_+$$

Hasil estimasi parameter dapat dilihat pada Tabel 4.

Tabel 4. Hasil estimasi spline kubik terboboti dengan knot optimal

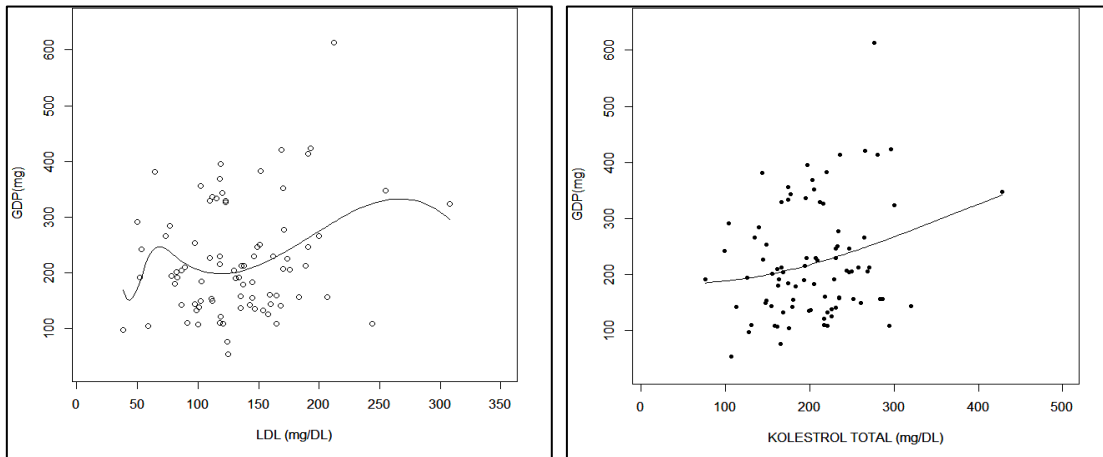
Parameter	Estimasi
β_0	1,59
β_{11}	1,51
β_{21}	-6,51
β_{31}	-6,32
$\beta_{1(4)}$	3,78
$\beta_{1(5)}$	-3,23
β_{12}	-5,00
β_{22}	5,09
β_{23}	-6,51
$\beta_{2(4)}$	1,72

Sumber: Data Diolah 2020

Berdasarkan estimasi parameter pada Tabel 2, diperoleh estimasi model spline kubik terboboti dengan knot optimal adalah:

$$\hat{y} = 1,59 + 1,51(x_1) - 6,51(x_1)^2 - 6,32(x_1)^3 + 3,78(x_1 - 59)_+^3 - 3,23(x_1 - 81)_+^3 - 5,00(x_2) + 5,09(x_2)^2 - 6,51(x_2)^3 + 1,72(x_2 - 99)_+^3$$

Adapun grafik estimasi fungsi spline kubik terboboti pada setiap prediktor dapat dilihat pada Gambar 2:



Gambar 2. Grafik estimasi spline kubik terboboti dengan knot optimal

4. Kesimpulan

Berdasarkan hasil yang diperoleh, model spline kubik terboboti pada data diabetes mellitus memiliki knot optimal untuk kolesterol LDL terletak pada 59 dan 81 mg/DL, sedangkan untuk kolesterol total terletak pada 99 mg/DL, sehingga model yang terbentuk seperti berikut:

$$\hat{y} = 1,59 + 1,51(x_1) - 6,51(x_1)^2 - 6,32(x_1)^3 + 3,78(x_1 - 59)_+^3 - 3,23(x_1 - 81)_+^3 - 5,00(x_2) + 5,09(x_2)^2 - 6,51(x_2)^3 + 1,72(x_2 - 99)_+^3$$

Pola hubungan faktor kolesterol dengan peningkatan kadar gula darah pasien diabetes mellitus mengikuti pola kubik. Pada penelitian selanjutnya pemodelan regresi spline kubik terbobot dapat menggunakan pembobot lain.

Daftar Pustaka

- [1] Islamiyati, A. Taksiran Kurva Regresi Spline pada Data Longitudinal dengan Kuadrat Terkecil. *Jurnal Matematika, Statistika & Komputasi*, 11 (1) : 77-83, 2014.
- [2] Islamiyati, A. Spline Polynomial Truncated dalam Regresi Nonparametrik. *Jurnal Matematika, Statistika & Komputasi*, 14 (1) : 54-60, 2017.
- [3] Lestari, B., Budiantara, I.N. & Chamidah, N. Estimation of Regression Function in Multi Response Nonparametric Regression Model Using Smoothing Spline and

- Kernel Estimators. *Journal of Physics: Conference Series*, 1097 (1), 012091, 2018.
- [4] Islamiyati, A., Fatmawati & Chamidah, N. Fungsi Goodness of Fit dalam Kriteria Penalized Spline pada Estimasi Regresi Nonparametrik Birespon untuk Data Longitudinal. *Proseding Seminar Nasional Matematika dan Aplikasinya*. UNAIR Surabaya. 2017.
- [5] Islamiyati, A., Fatmawati & Chamidah, I.N. Penalized Spline Estimator with Multi Smoothing Parameters in Biresponse Multipredictor Regression Model for Longitudinal Data. *Songklanakar Journal of Science and Technology*, In Press SJST-2018-0423.R2, 2019.
- [6] Chamidah, N. & Lestari, B. Spline Estimator in Homoscedastic Multirespon Nonparametric Regression Model in Case of Unbalance Number of Observations. *Far East Journal of Mathematical Sciences*, 100 (9) : 1433, 2016.
- [7] Maziyya. P.A. *Mengatasi Heterokedastisitas pada Regresi dengan Menggunakan Weighted Least Square*. Skripsi, Universitas Udayana, Bali, 2015.
- [8] Hendayanti, N.P.N. *Estimasi Kurva Regresi Nonparametrik Heterokedastisitas Spline*, Tesis, Institut Teknologi Sepuluh November, Surabaya, 2015.
- [9] Islamiyati, A., Fatmawati & Chamidah, N. Estimation of Covariance Matrix on Bi-Response Longitudinal Data Analysis with Penalized Spline Regression. *Journal of Physics: Conf. Series*. 979 pp 012093, 2018.
- [10] Islamiyati, A., Fatmawati & Chamidah, N. Ability of Covariance Matrix in Bi-Response Multi-Predictor Penalized Spline Model through Longitudinal Data Simulation. *Int. J. Acad. Appl. Res.* 3 Issue 3 : 8-11, 2019.
- [11] Islamiyati, A. Regresi Spline Polynomial Truncated Biprediktor untuk Identifikasi Perubahan Jumlah Trombosit Pasien Demam Berdarah Dengue. *Al khwarizmi*, 7 (2): 97-110, 2019.
- [12] Porcellati, F., Lucidi, P., Bolii, G.B. & Fanelli, C.G. Thirty Years of Research on The Dawn Phenomenon: Lessons to Optimize Blood Glucose Control in Diabetes. *Diabetes Care*, 36 : 3860-3862, 2013.
- [13] Islamiyati, A., Fatmawati & Chamidah, I.N. Changes in Blood Glucosa 2 Hours After Meals in Type 2 Diabetes Patients based on Length of Treatment at Hasanuddin University Hospital, Indonesia. *Rawal Medical Journal*, 45 (1) : 31-34, 2020.
- [14] Islamiyati, A., Raupong & Anisa. Use of Penalized Spline Linear to Identify Change in Pattern of Blood Sugar based on The Weight of Diabetes Patients. *Int. J. Acad. Appl. Res.*, 3 Issue 12 : 75-78, 2019.
- [15] Searle, S. R., Casella, G. & McCulloch, C. E. *Variance Components*. New Jersey : John Wiley and Sons, Inc., Hoboken, 2006.