BAB II
KAJIAN TEORI

2.1 Analisis Regresi

Analisis regresi adalah suatu metode statistika yang umum digunakan untuk melihat pengaruh antara variabel independen dengan variabel dependen. Hal ini dapat dilakukan melalui tiga pendekatan untuk mengestimasi kurva regresi yaitu regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Jika kurva regresi merupakan model parametrik maka disebut sebagai regresi parametrik dan apabila model yang diasumsikan ini benar, maka pendugaan parametrik sangat efisien, tetapi jika tidak, menyebabkan interpretasi data yang menyesatkan (Hardle, 1994). Misalnya Y adalah variabel respon dan X adalah variabel prediktor, secara umum hubungan antara Y dan X dapat ditulis sebagai berikut:

$$Y_i = \beta_0 + \beta_i x_i + \cdots + \beta_n x_n + \varepsilon_i, \ i = 1,2, \dots, n \tag{2.1}$$

Dimana,

Y_i = Variabel respon

β_i = Parameter

x_i = Variabel prediktor

ε_i = Error

Apabila tidak terdapat informasi apapun tentang bentuk fungsi, maka digunakan pendekatan nonparametrik (Hardle,1994).

2.2 Regresi Nonparametrik

Regresi nonparametrik adalah regresi yang diasumsikan tidak diketahui bentuk kurva regresi atau tidak terdapat informasi masa lalu yang lengkap tentang bentuk pola data (Eubank, l999). Beberapa model regresi nonparametrik yang banyak digunakan diantaranya: Spline, MARS, Kernel, Deret Fourier, Deret Orthogonal, Neural Network (NN), Polinomial Lokal, Histogram, Wavelets, k-NN, dan yang lainnya (Budiantara, 2009). Adapun model regresi nonparametrik secara umum disajikan pada persamaan berikut :

$$Y_i = f(x_i) + \varepsilon_i , i = 1,2, \ldots , n \tag{2.2}$$

Dengan,

Y_i = variabel respon

x_i = variabel prediktor

$f(x_i)$ = fungsi regresi

ε_i = *error* yang berdistribusi independen dengan mean nol dan varians σ^2

$f(x_i)$ merupakan kurva regresi yang tidak diketahui bentuknya. Kurva $f(x_i)$ diasumsikan *smooth* pada ruang fungsi tertentu.

2.3 Deret Fourier

Deret Fourier adalah fungsi polinomial trigonometri yang mempunyai tingkat fleksibilitas. Hal ini dikarenakan bahwa deret Fourier merupakan kurva yang menunjukkan fungsi sinus cosinus (Prahutama, 2013). Deret Fourier digunakan apabila data yang diselidiki tidak diketahui polanya dan cenderung berulang (Bilodeau, 1992). Estimator deret Fourier ini, umumnya digunakan apabila pola datanya tidak diketahui dan ada kecenderungan pola musiman (Tripena dan Budiantara, 2006). Fungsi Deret Fourier adalah sebagai berikut :

$$f(t) = \frac{1}{2}\hat{a}_0 + \hat{\gamma}t + \sum_{k=1}^{K} \hat{a}_k \cos\left(\frac{2\pi kt}{2L}\right) + \hat{b}_k \sin\left(\frac{2\pi kt}{2L}\right) \quad (2.3)$$

Jika fungsi $f(t)$ pada interval $(-L, L)$ dan diluar selang ini oleh $f(t \pm 2L) = f(t)$, sehingga $f(t)$ merupakan fungsi perodik dengan periode $2L$. $f(t)$ dapat direpresentasikan dengan deret perluasan fourier sebagai berikut:

$$f(t) = \frac{1}{2}a_0 + \gamma t + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi kt}{2L}\right) + b_k \quad (2.4)$$

$$f(t) = \frac{1}{2}a_0 + \gamma t + \sum_{k=1}^{\infty} a_k \cos\left(\frac{\pi kt}{L}\right) + b_k \quad (2.5)$$

$$a_k = \frac{1}{L}\int_{-L}^{L} f(x) \cos\left(\frac{\pi kt}{L}\right) dt, k = 1,2,3 \ldots \quad (2.6)$$

$$b_k = \frac{1}{L}\int_{-L}^{L} f(x) \sin\left(\frac{\pi kt}{L}\right) dt, k = 1,2,3 \ldots \quad (2.7)$$

Nilai $\frac{2\pi}{T}$ (dengan T adalah periode $f(t)$ merupakan pengali agar t dalam satuan radian. Diberikan n data pengamatan $\{(t_i, y_i)\}_{i=1}^{n}$ yang memenuhi persamaan diatas. Jika $T_i \in [-L, L]$ dan $Y_i \in R$ dan diasumsikan periode $f(t)$ dapat didekati oleh deret Fourier yang didefinisikan sebagai berikut :

$$f(t) = \frac{1}{2}a_0 + \gamma t + \sum_{k=1}^{K} a_k \cos\left(\frac{2\pi kt}{2L}\right) \quad (2.8)$$

Dengan a_0, a_k , dan b_k adalah koefisien Fourier (Asrini, 2012).

Tingkat kemulusan estimator deret Fourier ditentukan oleh pemilihan parameter pemulus K. Semakin kecil parameter pemulus K, semakin mulus estimasinya dan semakin besar parameter pemulus K, semakin kurang mulus estimasi dari f. Oleh karena itu, perlu dipilih K yang optimal.

2.4 *Generalized Cross Validation* (GCV)

Pada pemodelan regresi nonparametrik untuk menentukan parameter pemulus adalah dengan menggunakan deret Fourier, hal yang perlu diperhatikan adalah menentukan nilai K. Salah satu metode yang dapat digunakan adalah metode *Generalized Cross Validation* (GCV). Penentuan nilai K optimal akan menghasilkan nilai koefisien (R^2) yang tinggi. *Generalized Cross Validation* (GCV) didefinisikan sebagai berikut :

$$GCV(K) = \frac{MSE(K)}{\left[1-\left(\frac{trace(A_K)}{n}\right)\right]^2} \tag{2.9}$$

(Wu dan Zhang ,1988)

Dimana :

$$MSE(K) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2.10}$$

$$(\boldsymbol{A_K}) = \boldsymbol{A}(\boldsymbol{A^T A})^{-1}\boldsymbol{A^T} \tag{2.11}$$

$$\hat{y}_i = (\boldsymbol{A_K})y \tag{2.12}$$

n : banyak pengamatan

Nilai GCV terkecil akan menghasilkan K yang optimal (Asrini, 2012 dalam Prahutama, 2013).

2.5 *Cross Validation* (CV)

Metode CV membagi data menjadi n sub-sampel berukuran $n-1$. Di asumsikan $f_{K_{(i)}}$ merepresentasikan fungsi $f_{K_{(i)}}$ yang dihitung tanpa observasi ke-i. Fungsi $f_{K_{(i)}}$ dibangun dari sub-sampel berukuran $n-1$ yang diambil dari data asli. Jika observasi ke-i diperlukan sebagai observasi tambahan (*holdout sample*), maka fungsi CV adalah :

$$CV(K) = \frac{1}{n}\sum_{i=1}^{n}(y_i - f_{K_{(i)}})^2 \tag{2.13}$$

Fungsi $f_{K_{(i)}}$ bila diuraikan sedemikian sehingga

$$f_{K_{(i)}} = f_{K_i} - \frac{s_i(y_i - f_{K_i})}{(1-s_i)}, \tag{2.14}$$

Dimana s_i adalah elemen diagonal ke-I dari **A(K)**. sehingga,

$$\begin{aligned}
CV(k) &= \frac{1}{n}\sum_{i=1}^{n}(y_i - f_{K_{(i)}})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \left(f_{K_i} - \frac{(y_i - f_{K_i})s_i}{(1-s_i)}\right)\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(y_i - f_{K_i} + \frac{(y_i - f_{K_i})s_i}{(1-s_i)}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{(y_i - f_{K_i})(1-s_i)}{(1-s_i)} + \frac{(y_i - f_{K_i})s_i}{(1-s_i)}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\left(y_i - f_{K_i}\right)(1-s_i+s_i)}{(1-s_i)}\right)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - f_{K_i}}{(1-s_i)}\right)^2
\end{aligned}$$

Sehingga didapatkan kriteria CV yaitu :

$$CV(K) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - f_{K_i}}{1 - s_i}\right)^2 \qquad (2.15)$$

Pemilihan K optimal dengan metode CV didapatkan dengan meminimalkan fungsi CV (Devi, 2015).

2.6 Estimasi Parameter Regresi Nonparametrik Deret Fourier

Diberikan data berpasangan sebagai berikut (y_j, t_{ij}) dengan $j = 1,2,3,\ldots,n$ menyatakan banyaknya variabel pengamatan dan $i = 1,2,3,\ldots,p$ menyatakan banyaknya variabel independen. Model regresi nonparametriknya adalah sebagai berikut :

$$Y_i = T_i\beta + \varepsilon \text{ atau} \qquad (2.16)$$

$$Y_i = \beta_0 + \beta_1 T_{1i} + \beta_2 T_{2i} + \beta_3 T_{3i} + \cdots + \beta_p T_{pi} + \varepsilon_i \qquad (2.17)$$

Jika persamaan (2.16) dibuat dalam bentuk matrik $Y_i = T_i\beta + \varepsilon$ maka sebagai berikut :

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{nx1} \quad x = \begin{bmatrix} 1 & t_{11} & t_{12} & \ldots & t_{1p} \\ 1 & t_{21} & t_{22} & \ldots & t_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_{n1} & t_{n2} & \ldots & t_{np} \end{bmatrix}_{nx(px1)} \qquad (2.18)$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}_{(p+1)x1} \quad \varepsilon = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}_{nx1}$$

Jika $y = f(t) + \varepsilon$ maka :

$$f(t) = [f(t_1)\, f(t_2)\, f(t_3) \ldots f(t_n)] \qquad (2.19)$$

$f(t)$ merupakan kurva yang tidak diketahui bentuknya maka $f(t)$ didekati dengan menggunakan Deret Fourier yaitu sebagai berikut :

$$f(t) = \frac{1}{2}\alpha_0 + \gamma_t + \sum_{k=1}^{K} \alpha_k \cos\, kt \qquad (2.20)$$

Persamaan (2.20) didapatkan elemen diagonalnya yaitu :

$$f(t) = \frac{1}{2}\alpha_0 + \gamma_t + (\alpha_{i1} \cos t_i + \alpha_{12} \cos t_i + \ldots + \alpha_{ik} t_i) \qquad (2.21)$$

Jika $f(t) = B\hat{\theta}$, maka :

$$B = \begin{bmatrix} 1 & t_1 & \cos t_1 & \ldots & \cos kt_1 \\ 1 & t_2 & \cos t_2 & \ldots & \cos kt_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & \cos t_n & \ldots & \cos kt_n \end{bmatrix}_{n(k+2)} \hat{\theta} = \begin{bmatrix} \phi \\ \gamma \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}_{(k+2)x1} \qquad (2.21)$$

Dengan $\phi = \frac{n}{2}\alpha_0$

2.7 Kemiskinan Penduduk

Kemiskian adalah kondisi kehidupan yang serba kekurangan yang dialami seseorang yang pegeluaran per kapitanya selama sebulan tidak cukup untuk memenuhi kebutuhan standar hidup minimum. Kebutuhan standar hidup minimum digambarkan dengan garis kemiskinan (GK), yaitu batas minimum pengeluaran per kapita per bulan untuk memenuhi kebutuhan minimum makanan dan nonmakanan. Batas kecukupan minimum makanan mengacu pada Widya Karya Nasional Pangan dan Gizi pada tahun 1978, yaitu besarnya rupiah yang dikeluarkan untuk makanan yang memenuhi kebutuhan minimum energy 2100 kalori per kapita per hari, sedangkan kebutuhan minimum nonmakanan mencakup pendapatan untuk perumahan, penerangan, bahan bakar, pakaian, pendidikan, kesehatan, transportasi, barang-barang tahan lama serta barang dan jasa esensial lainnya. Jumlah orang miskin dapat dilihat melalui jumlah orang yang berada di bawah atau sama dengan garis kemiskinan (Badan Pusat Statistik, 2002).

Kemiskinan adalah kondisi dimana seseorang atau sekelompok orang tidak mampu memenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Hak-hak dasar antara lain (a) terpenuhinya kebutuhan pangan, (b) kesehatan, pendidikan, pekerjaan, perumahan, air bersih, pertanahan, sumber daya alam dan lingkungan hidup, (c) rasa aman dari perlakuan atau ancaman tindak kekerasan, (d) hak untuk berpartisipasi dalam kehidupan sosial-politik (Badan Pusat Statistik, 2009).

2.8 Pendapatan Per kapita

Pendapatan per kapita atau produk domestik bruto per kapita adalah pendapatan rata-rata penduduk suatu negara (Untoro, 2010:13). Pendapatan per kapita menunjukan tingkat pendapatan masyarakat dalam suatu negara. Rumus untuk mencari pendapatan per kapita adalah :

$$\text{Pendapatan per kapita} = \frac{Produk\ Nasional\ Bruto\ (GNP)}{Jumlah\ Penduduk} \tag{2.16}$$

Menurut Rakiman (2011:80) pendapatan per kapita suatu negara merupakan tolak ukur kemajuan dari negara tersebut, apabila pendapatan per kapita suatu negara rendah dapat dipastikan mekanisme ekonomi masyarakat di negara tersebut mengalami penurunan dan begitu pula sebaliknya apabila pendapatan per kapita suatu negara tinggi maka dapat dipastikan mekanisme ekonomi masyarakat tersebut mengalami peningkatan, tapi pendapatan tersebut bukan hanya didapat atau diperoleh dari mekanisme ekonomi masyarakatnya saja banyak faktor yang mempengaruhi penurunan atau peningkatan pendapatan tersebut seperti keadaan alam yang tidak dapat diperkirakan keadaannya, kondisi alam ini dapat berubah sewaktu-waktu yang dapat menimbulkan bencana alam yang akan membuat pendapatan suatu negara akan mengalami penurunan.

Pendapatan per kapita pada skala daerah dapat digunakan sebagai pengukur pertumbuhan ekonomi yang lebih baik karena lebih baik dan lebih tepat mencerminkan kesejahteraan penduduk suatu negara atau daerah yang bersangkutan. Manfaat perhitungan pendapatan per kapita sebagai indikator ekonomi yang mengukur tingkat kemakmuran penduduk suatu negara, pendapatan per kapita dihitung secara berkala, biasanya 1 tahun. Manfaat dari perhitungan pendapatan per kapita antara lain sebagai berikut (Alam, 2007: 50) :

a. Untuk melihat tingkat perbandingan kesejahteraan masyarakat suatu negara dari tahun ke tahun.
b. Sebagai data perbandingan tingkat kesejahteraan suatu negara dengan negara lain.
c. Sebagai perbandingan tingkat standar hidup suatu negara dengan negara lainnya.
d. Sebagai data untuk mengambil kebijakan di bidang ekonomi.