BAB II
LANDASAN TEORI

2.1 Analisis Regresi

Analisis regresi merupakan suatu metode statistika yang digunakan untuk menentukan hubungan antara suatu variabel respon pada satu atau lebih variabel prediktor. Tujuan utama analisis regresi adalah bagaimana mencari bentuk estimasi untuk kurva regresi. Analisis regresi dapat dilakukan dengan tiga pendekatan yaitu parametrik, semiparametrik dan nonparametrik (Mariati, 2015).

Variabel-variabel regresi yang berhubungan secara linear disebut regresi linier. Regresi yang menghubungkan satu variabel respon dengan satu variabel prediktor disebut regresi liner sederhana, sedangkan regresi linear yang menghubungkan satu variabel respon dengan dua atau lebih variabel prediktor disebut regresi linear berganda. Jika bentuk kurva regresi diketahui, maka dapat digunakan pendekatan parametrik, sedangkan jika bentuk kurva regresi tidak diketahui dan tidak terdapat informasi yang lengkap sebelumnya maka dapat menggunakan pendekatan nonparametrik (Budiantara, 2014).

2.2 Regresi Nonparametrik

Regresi nonparametrik merupakan metode regresi yang digunakan ketika bentuk pola hubungan antara variabel respon dengan variabel prediktor tidak diketahui bentuk kurva regresinya. Bentuk model regresi tergantung pada kurva $f(x)$ (Lestarai & Budiantara, 2007). Estimasi kurva sangat tergantung dengan perilaku data sampel. Jika model regresi nonparametrik mempunyai lebih dari satu variabel prediktor maka dapat ditulis sebagai berikut:

$$y_i = f\left(x_{ji}\right) + \varepsilon_i, i = 1,2,\cdots,n, j = 1,2,\dots,t \tag{2.1}$$

dengan y_i adalah variabel respon ke-i, x_{ji} adalah i yaitu banyaknya pengamatan dan j yaitu banyaknya variabel prediktor, $f\left(x_{ji}\right)$ adalah fungsi regresi nonparametrik yang tidak diketahui bentuknya, dan ε_i adalah *error* yang diasumsikan identik, independen, dan berdistribusi normal.

2.3 Deret Fourier

Deret fourier merupakan suatu deret tak hingga dengan suku-suku memuat komponen trigonometri, sinus-cosinus, yang konvergen ke suatu fungsi periodik. Fungsi periodik adalah fungsi yang mengulang nilainya secara berkala. Berikut fungsi deret fourier:

$$f(x) = \frac{\beta_0}{2} + \sum_{n=1}^{\infty}(a_n coskx + b_n sinkx) \quad (2.2)$$

Dalam regresi deret fourier selain menggunakan kombinasi aditif fungsi sinus dan cosinus, dapat juga digunakan kombinasi aditif fungsi linier dan fungsi sinus atau cosinus. Dengan mengambil kombinasi fungsi linier dan salah satu fungsi sinus atau cosinus akan memberikan efisiensi estimasi (Suparti, 2018). Pendekatan kombinasi aditif fungsi linier dan cosinus sebagai berikut:

$$f(x) = \beta x + \frac{1}{2}\beta_0 + \sum_{k=1}^{K} \alpha_k coskx \quad (2.3)$$

2.4 Pemilihan Titik Osilasi

Pada pemodelan regresi nonparametrik dengan menggunakan Deret Fourier, hal yang perlu diperhatikan adalah menentukan nilai k atau parameter osilasi. Pemilihan nilai parameter osilasi harus dilakukan seoptimal mungkin. Penentuan parameter osilasi bisa menggunakan metode GCV (*Generalized Cross Validation*). Fungsi GCV untuk pemilihan panjang titik osilasi dapat ditunjukkan pada persamaan berikut (Wisisono, 2018):

$$GCV(k_1, k_2, \dots, k_r) = \frac{MSE(k_1, k_2, \dots, k_K)}{(n^{-1}tr(\boldsymbol{I} - A(k_1, k_2, \dots, k_K)))^2} \quad (2.4)$$

dimana :

$$MSE(k_1, k_2, \dots, k_r) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_i) \quad (2.5)$$

2.5 Pengujian Parameter Model

Terdapat dua tahap pengujian parameter yaitu pengujian secara serentak dan pengujian secara individu (parsial).

2.5.1 Uji Serentak

Uji serentak yaitu pengujian seluruh parameter yang terdapat dalam model secara bersama-sama. Uji serentak sering disebut uji F. uji F dilakukan untuk mengetahui apakah variabel respon secara serentak berpengaruh signifikan terhadap variabel prediktor atau tidak (Octavanny, 2017). Pengujian uji F dapat dilakukan dengan hipotesis yaitu:

$H_0: \beta_1 = \beta_2 = \cdots = \beta_i = 0$

H_1 : Minimal Ada Satu $\beta_i \neq 0$, $\quad i = 1,2,\cdots,n$

dengan statistik uji

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \tag{2.6}$$

MS regresi dan MS error didapatkan dari analisis ragam. Hasil perhitungan nilai F kemudian dilakukan pembandingan dengan nilai F tabel serta α yang digunakan misalnya 0,05. Jika F hitung lebih besar dari F tabel, maka hipotesis dapat dipergunakan secara simultan (Suharjo, 2013). Berikut adalah tampilan penyajian uji ANOVA:

Tabel 2.1 Uji Analisis Ragam **(**ANOVA)

Sumber	Df	Sum of Square	Mean Square	F_{hitung}
Regresi	$p-1$	$\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	$\frac{SSR}{p-1}$	$\frac{MS_{regresi}}{MS_{error}}$
Error	$n-p$	$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$	$\frac{SSE}{n-p}$	
Total	$(n-1)$	$\sum_{i=1}^{n}(y_i - \bar{y}_i)^2$		

2.5.2 Uji Parsial

Uji parsial sering disebut juga uji T. digunakan untuk menguji suatu parameter mempunyai pengaruh yang signifikan terhadap variabel respon. Hipotesisnya yaitu sebagai berikut (Octavanny, 2017):

$H_0: \beta_i = 0$ (tidak terdapat pengaruh variabel prediktor terhadap variabel respon)
$H_1: \beta_i \neq 0$, (terdapat pengaruh signifikan dari variabel prediktor terhadap variabel respon)

Statistik uji yang digunakan yaitu sebagai berikut:

$$t_{hitung} = \frac{\beta_i}{SE(\beta_i)} \tag{2.7}$$

dengan $SE(\beta_i)$ merupakan *Standard Error*. H_0 ditolak jika $|t_{hit}| \geq t_{(\frac{\alpha}{2}, n-p)}$.

2.6 Pemeriksaan Asumsi Residual

Dalam mendeteksi residual telah memenuhi asumsi maka diperlukan adanya pemeriksaan terhadap residual tersebut. Berikut ini asumsi residual yang harus dipenuhi dalam regresi nonparametrik deret fourier:

2.6.1 Asumsi Residual Identik

Mendeteksi ada tidaknya korelasi dapat dilakukan dengan membuat plot antara $\hat{y}$ dengan residual. Apabila terdapat pola maka dapat diindikasikan terjadi heterokedastisitas dan asumsi identik tidak terpenuhi. Selain itu dapat dilakukan dengan uji *Glejser* yakni dengan meregresikan nilai mutlak residual dengan variabel prediktor (Octavanny, 2017). Hipotesis yang digunakan yaitu:

$H_0: \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_i^2 = \sigma^2$
H_1 : minimal ada satu $\sigma_i \neq \sigma^2, \qquad i = 1,2,\ldots,n$

dimana n adalah banyaknya variabel prediktor. Statistik Uji *Glejser*:

$$F_{hitung} = \frac{\sum_{i=1}^{n}(|\hat{e}_i| - |\bar{e}_i|)^2/(p-1)}{\sum_{i=1}^{n}(|e_i| - |\hat{e}_i|)/(n-p)} = \frac{MS_{regresi}}{MS_{error}} \tag{2.8}$$

dimana:

$\hat{e}_i$: estimasi residual
$\bar{e}$: rata-rata residual
e_i : residual ke-i

Keputusannya adalah terima H_0 jika nilai statistik uji kurang dari nilai $\chi^2_{df;\alpha}$. Atau $p - value$ lebih besar dari α.

2.6.2 Asumsi Residual Independen

Uji independen dilakukan untuk memastikan bahwa tidak terdapat korelasi antar residual atau autokorelasi. Pendeteksian autokorelasi dapat dilakukan dengan melihat plot ACF dari residual. Untuk mendapatkan nilai ACF digunakan rumus sebagai berikut (Octavanny, 2017):

$$\hat{\rho}_w = \frac{\sum_{t=w+1}^{n}(e_t - \bar{e})(e_{t-w} - \bar{e})}{\sum_{t=1}^{n}(e_t - \bar{e})^2} \tag{2.9}$$

Asumsi independent dapat dideteksi dengan menggunakan interval konfidensi dengan rumus sebagai berikut:

$$-t_{n-1,\frac{\alpha}{2}} SE(\rho_w) < \rho_w < t_{n-1,\frac{\alpha}{2}} SE(\rho_w) \tag{2.10}$$

Apabila terdapat ACF yang keluar dari interval konfidensi, maka diindikasikan adanya autokorelasi antar residual.

2.6.3 Uji Asumsi Berdistribusi Normal

Untuk mengetahui residual berdistribusi normal atau tidak, dapat dilakukan uji *Kolmogorov Smirnov* (Suharjo, 2013). Berikut hipotesisinya:

$H_0: F_0(x) = F(x)$ (residual berdistribusi normal)

$H_1: F_0(x) \neq F(x)$ (residual tidak berdistribusi normal)

Statistik uji:

$$D = Maks|F_0(x) - S_n(x)| \tag{2.11}$$

dimana $F_0(x)$ merupakan fungsi distribusi frekuensi kumulatif. $S_n(x) = k/N$ merupakan fungsi peluang kumulatif yang diobservasi dari suatu sampel random dengan N observasi. Keputusan pada uji *Kolmogorov-Smirnov* adalah tolak H_0 jika $|D| > Ks_{df;\alpha}$ dimana nilai Ks berdasarkan tabel *Kolmogorov-Smirnov*.

2.7 Koefisien Determinasi

Koefisien determinasi (R^2) merupakan ukuran kontribusi variabel-variabel prediktor terhadap variabel respon. Adapun rumus koefisien determinasi adalah (Sahidah, 2020):

$$R^2 = 1 - \frac{SSE}{SST} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.12}$$

dimana R^2 adalah koefisien determinasi; SSE merupakan *Sum of Square Error*; SST merupakan *Sum of Square Total*.