

BAB II

LANDASAN TEORI

2.1 Analisis Regresi

Analisis regresi pertama kali dikenalkan oleh Sir Francis Galton pada tahun 1877 (Tediwibawa dkk, 2019). Sir Francis Galton melakukan penelitian tentang hubungan antara tinggi badan anak dan tinggi badan orang tuanya. Saat ini analisis regresi sangat banyak digunakan dalam penelitian.

Analisis regresi adalah metode statistika yang digunakan untuk memprediksi nilai variabel respon berdasarkan pada nilai variabel prediktor (Kusnandar dkk, 2019). Pada analisis regresi terdapat tiga pendekatan yaitu pendekatan regresi parametrik, nonparametrik, dan semiparametrik. Hubungan antara variabel dapat dijelaskan dalam sebuah kurva yang disebut kurva regresi. Apabila pola pada hubungan antara variabel respon dan variabel prediktor tidak diketahui maka pendekatan yang digunakan adalah pendekatan regresi nonparametrik.

2.2 Regresi Nonparametrik

Regresi nonparametrik adalah analisis yang digunakan untuk mengetahui bentuk hubungan antara variabel respon dan variabel prediktor jika tidak diketahui bentuk kurva dari regresinya (Pangestikasari dkk, 2018). Regresi nonparametrik memiliki asumsi bahwa kurva regresi yang dibentuk bersifat mulus (*smooth*) sehingga regresi nonparametrik mempunyai fleksibilitas yang tinggi karena data diharapkan untuk mencari sendiri bentuk estimasi dari kurva regresinya. Secara umum model regresi nonparametrik adalah sebagai berikut (Afa, Suparti, dan Rahmawati, 2018):

$$Y_i = f(X_i) + \varepsilon_i, i = 1, 2, 3, \dots, n \quad (2.1)$$

Keterangan:

- Y_i : Variabel respon pada observasi ke- i
- $f(X_i)$: Fungsi regresi ke- i yang tidak diketahui
- ε_i : Residual untuk observasi ke- i

2.3 Koefisien Korelasi

Koefisien korelasi adalah suatu nilai yang digunakan untuk mengukur besarnya hubungan antara dua variabel atau lebih (Sugiyono, 2018). Nilai koefisien korelasi secara umum dapat dihitung menggunakan rumus korelasi *Pearson* sebagai berikut (Dewanti dkk, 2020):

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{(\sum_{i=1}^n (X_i - \bar{X})^2)(\sum_{i=1}^n (Y_i - \bar{Y})^2)}} \quad (2.2)$$

Keterangan:

- r : Koefisien korelasi
 X_i : Variabel prediktor pada observasi ke- i
 \bar{X} : Rata-rata dari variabel prediktor
 Y_i : Variabel respon pada observasi ke- i
 \bar{Y} : Rata-rata dari variabel respon
 n : Banyaknya observasi

Besar nilai r berada pada rentang -1 sampai $+1$ ($-1 \leq r \leq +1$). Apabila nilai koefisien korelasi semakin mendekati $+1$ artinya terdapat hubungan sempurna positif antara variabel sedangkan jika nilai koefisien korelasi semakin mendekati -1 artinya terdapat hubungan sempurna negatif antara variabel (Femadiyanti, Suparti, dan Warsito, 2020). Nilai koefisien korelasi dapat diinterpretasikan pada tabel kriteria pengukuran berikut (Sugiyono, 2018):

Tabel 2.1 Interpretasi Nilai Koefisien Korelasi

Nilai Koefisien Korelasi	Interpretasi
0,000 – 0,199	Korelasi sangat lemah
0,200 – 0,399	Korelasi lemah
0,400 – 0,599	Korelasi cukup
0,600 – 0,799	Korelasi kuat
0,800 – 1,000	Korelasi sangat kuat

Berdasarkan Tabel 2.1 interpretasi dari nilai koefisien korelasi hanya dapat digunakan untuk memberikan interpretasi dari besarnya nilai koefisien korelasi, tetapi tidak dapat digunakan untuk menentukan apakah terdapat korelasi atau tidak

antara dua variabel. Pengujian koefisien korelasi dapat dilakukan dengan uji hipotesis sebagai berikut (Dewanti dkk, 2020):

a. Hipotesis

$H_0 : \rho = 0$ (tidak terdapat korelasi antara kedua variabel)

$H_1 : \rho \neq 0$ (terdapat korelasi antara kedua variabel)

b. Taraf Signifikansi

Taraf signifikansi yang digunakan adalah $\alpha = 5\% = 0,05$

c. Statistik Uji

$$t_{hitung} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (2.3)$$

Keterangan:

r : Koefisien korelasi

n : Banyaknya observasi

d. Kriteria Pengujian

H_0 ditolak jika nilai $|t_{hitung}| > t_{\frac{\alpha}{2}; n-2}$

H_0 diterima jika nilai $|t_{hitung}| < t_{\frac{\alpha}{2}; n-2}$

e. Pengambilan Keputusan

Apabila H_0 ditolak maka terdapat korelasi antara kedua variabel.

Apabila H_0 diterima maka tidak terdapat korelasi antara kedua variabel.

2.4 Matriks

Matriks merupakan susunan dari suatu bilangan berbentuk empat persegi panjang yang diletakkan dalam kurung siku. Susunan dari suatu bilangan yang mendatar disebut baris dan susunan dari suatu bilangan yang menurun disebut kolom (Sessu, 2014). Bilangan yang terdapat dalam susunan tersebut disebut dengan elemen. Bentuk umum dari matriks dapat dituliskan sebagai berikut:

$$C_{m \times n} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2n} \\ c_{31} & c_{32} & c_{33} & \dots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & c_{m3} & \dots & c_{mn} \end{bmatrix}$$

matriks \mathbf{C} disebut matriks $m \times n$ karena terdiri dari m baris dan n kolom. Matriks \mathbf{C} juga dapat ditulis $\mathbf{C} = [c_{ij}]$ dengan $i = 1, 2, 3, \dots, m$ dan $j = 1, 2, 3, \dots, n$. Setiap c_{ij} disebut elemen dari matriks, sedangkan indeks i menyatakan baris dan j menyatakan kolom. Jadi, elemen c_{ij} terdapat pada baris ke- i dan kolom ke- j .

Matriks bujur sangkar adalah matriks yang mempunyai baris dan kolom yang jumlah elemennya sama (Anton dan Rorres, 2004). Apabila \mathbf{C} adalah sebuah matriks bujur sangkar ($n \times n$) maka *trace* dari matriks \mathbf{C} merupakan jumlahan dari elemen-elemen pada diagonal utama matriks \mathbf{C} . Matriks bujur sangkar dapat dituliskan sebagai berikut:

$$\mathbf{C}_{n \times n} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2n} \\ c_{31} & c_{32} & c_{33} & \dots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \dots & c_{nn} \end{bmatrix}$$

maka *trace* dari matriks $\mathbf{C} = c_{11} + c_{22} + c_{33} + \dots + c_{nn}$. *Trace* dari matriks \mathbf{C} tidak dapat didefinisikan jika \mathbf{C} bukan matriks bujur sangkar (Anton dan Rorres, 2004). Matriks identitas adalah matriks bujur sangkar yang memiliki ordo $n \times n$ dengan elemen-elemen pada diagonal utama bernilai satu dan elemen yang lain bernilai nol. Apabila matriks identitas memiliki ordo $n \times n$ maka dapat ditulis dengan \mathbf{I}_n .

Apabila diketahui \mathbf{A} dan \mathbf{B} adalah matriks bujur sangkar $\mathbf{AB} = \mathbf{BA}$, maka \mathbf{B} disebut invers dari matriks \mathbf{A} yang dapat ditulis dengan $\mathbf{B} = \mathbf{A}^{-1}$ dan \mathbf{A} disebut invers dari matriks \mathbf{B} yang dapat ditulis dengan $\mathbf{A} = \mathbf{B}^{-1}$. Matriks yang mempunyai invers disebut matriks nonsingular, sedangkan matriks yang tidak mempunyai invers disebut matriks singular. Matriks *transpose* diperoleh dengan menukarkan baris-baris dan kolom-kolom pada matriks \mathbf{A} . *Transpose* dari matriks \mathbf{A} berordo $m \times n$ didefinisikan sebagai matriks \mathbf{A}^T dengan ordo $m \times n$ yang diperoleh dari menukarkan baris dan kolom pada matriks \mathbf{A} (Anton dan Rorres, 2004).

2.5 Koefisien Determinasi

Koefisien determinasi (R^2) adalah besaran yang dapat digunakan untuk mengukur besarnya persentase keragaman dalam variabel respon yang mampu dijelaskan oleh variabel prediktor. Koefisien determinasi (R^2) digunakan sebagai

kriteria dalam mengetahui kebaikan dari model (Padatuan dkk, 2021). Nilai pada koefisien determinasi (R^2) berada pada $0 \leq R^2 \leq 1$ (Femadiyanti dkk, 2020). Apabila nilai dari R^2 sebesar 1 maka terjadi kecocokan yang sempurna dan apabila nilai dari R^2 sebesar 0 maka tidak terdapat hubungan antara variabel prediktor dengan variabel respon. Jika nilai dari koefisien determinasi (R^2) semakin besar maka model yang diperoleh juga semakin baik. Adapun rumus untuk menghitung koefisien determinasi (R^2) sebagai berikut (Pangestikasari dkk, 2018):

$$R^2 = \frac{\sum_{p=1}^2 \sum_{i=1}^n \left((\hat{Y}_i^{(p)} - \bar{Y}^{(p)}) \right)^2}{\sum_{p=1}^2 \sum_{i=1}^n \left((Y_i^{(p)} - \bar{Y}^{(p)}) \right)^2} \quad (2.4)$$

Keterangan:

- R^2 : Koefisien determinasi
 $Y_i^{(p)}$: Variabel respon ke- p pada observasi ke- i
 $\bar{Y}^{(p)}$: Rata-rata dari variabel respon ke- p
 $\hat{Y}_i^{(p)}$: Nilai estimasi variabel respon ke- p pada observasi ke- i
 n : Banyaknya observasi