

BAB II

TINJAUAN PUSTAKA

2.1 Analisis Regresi

Analisis regresi adalah suatu metode statistika yang umum digunakan untuk melihat pengaruh antara peubah bebas (variabel prediktor) dengan peubah tak bebas (variabel respon), (Houstis 1998). Hal ini dapat dilakukan melalui dua pendekatan untuk mengestimasi kurva regresi yaitu regresi parametrik dan regresi nonparametrik. 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 prediktor, secara umum hubungan antara Y dan X dapat ditulis sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \dots + \beta_n X_n + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.1)$$

Keterangan :

Y_i = Variabel dependen

β_i = Parameter

X_i = Variabel Independen

ε_i = Error

Apabila tidak terdapat informasi apapun tentang bentuk fungsi, maka digunakan pendekatan nonparametrik.

2.2 Regresi Parametrik

Analisis regresi merupakan suatu studi yang digunakan untuk melihat ketergantungan atau hubungan antara suatu variabel respons (variabel terikat) pada satu atau lebih variabel prediktor (variabel bebas). Analisis regresi terdiri dari dua jenis variabel yaitu variabel tak bebas atau variabel respons yang disebut juga sebagai variabel dependen. Hubungan antara variabel respons (Y_i) dengan variabel prediktor (X_i) dapat dinyatakan dalam persamaan berikut.

$$Y_i = f(X_i) + \varepsilon_i ; i = 1, 2, 3, \dots, n \quad (2.2)$$

dalam hal ini Y_i adalah nilai variabel respons dalam amatan ke- i , X_i adalah peubah bebas atau variabel prediktor dalam amatan ke- i , $f(X_i)$ adalah regresi yang telah diketahui bentuknya ε_i adalah suku galat atau sisaan yang diasumsikan independen dan bersifat acak dengan nilai tengah nol dan variansi σ^2 dalam amatan ke- i , dan n adalah banyaknya amatan Model antara dua atau lebih variabel independen atau variabel prediktor ($X_1, X_2, X_3, \dots, X_h$) dengan variabel dependen (Y_i) secara umum dapat ditulis sebagai berikut.

$(Y_i) = \sum_{g=0}^h y_g X_{ig} + \varepsilon_i$ dengan $i=1, 2, 3, \dots, n$; $y_0, y_1, y_2, \dots, y_h$ adalah parameter koefisien regresi, $X_{i1}, X_{i2}, \dots, X_{ih}$ adalah peubah bebas yang diketahui nilainya, dan ε_i adalah galat atau sisaan yang saling bebas.

Adapun koefisien determinasi dalam regresi parametrik adalah sebagai berikut.

$$R^2 = \frac{JKR}{JKT} = 1 - \frac{JKG}{JKT} \quad (2.3)$$

2.3 Regresi Nonparametrik

Menurut Schmidt (2013) Pendekatan parametrik merupakan pendekatan yang digunakan mengatasi nonlinieritas menunjukkan bahwa fungsi regresi secara eksplisit didefinisikan sebagai perantara yang menentukan fleksibilitas kurva. Untuk mengatasi keterbatasan regresi parametrik, pendekatan nonparametrik dikembangkan. Tidak adanya apriori asumsi yang dibuat tentang bentuk kurva. Salah satu dasar pendekatan nonparametrik adalah regresi linier bergherak halus. Menurut Lestari (2008) 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) dan (Hu 2007). Adapun model regresi nonparametrik secara umum disajikan persamaan .

$$Y_i = f(t_i) + \varepsilon_i, i = 1, 2, \dots, n \quad (2.4)$$

Keterangan :

Y_i = variabel dependen

t_i = variabel independen

$f(t_i)$ = fungsi regresi

$\varepsilon_i = \text{error}$ yang berdistribusi independen dengan mean nol dan varians σ^2

$f(t_i)$ merupakan kurva regresi yang tidak diketahui bentuknya. Kurva $f(t_i)$ diasumsikan smooth pada ruang fungsi tertentu (Tilling 2014).

2.4 Regresi Semiparametrik

Regresi semiparametrik merupakan salah satu metode Statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dan prediktor dimana sebagian pola data diketahui bentuknya dan sebagian lagi tidak diketahui bentuknya, (Asmin 2010). Misalkan terdapat data berpasangan x_i, y_i, t_i dan hubungan antara x_i, y_i dan t_i diasumsikan mengikuti model regresi semiparametrik sebagai berikut.

$$y_i = x_{ik}\beta + f(t_{iq}) + \varepsilon_i, \quad t_{iq} \in [a, b], \quad \begin{matrix} i = 1, 2, \dots, n \\ k = 1, 2, \dots, n \\ q = 1, 2, \dots, n \end{matrix} \quad (2.5)$$

Atau model regresi semiparametrik yang dinyatakan sebagai berikut.

$$Y_i = f(t_i) + X_i^1 y + \varepsilon_i$$

Keterangan : $i = 1, 2, 3, \dots, n$,

Y_i : nilai variabel respon dalam amatan ke- i ,

$X_i = (X_{i1}, X_{i2}, X_{i3}, \dots, X_{ih})$: adalah peubah bebas atau variabel prediktor yang

berhubungan secara parametrik dengan variabel respon

$Y_i, y = (y_0, y_1, y_2, \dots, y_h)$ adalah parameter koefisien regresi, sementara itu t_i adalah

variabel prediktor yang tidak diketahui bentuk polanya terhadap variabel respon (hubungan secara nonparametrik dengan variabel respon Y_i) (Mazzia, 2006).

Maka persamaan yang terbentuk sebagai berikut (Laome 2009).

$$Y_i = \sum_{j=0}^{p-1} \beta_p t_i^j + \sum_{l=1}^r \beta_{(p-1+l)} (t_i - k_l)_+^{p-1} + \sum_{g=0}^h y_g X_{ig} + \varepsilon_i \quad (2.6)$$

2.5 Regresi Spline

Menurut Markov (2002). Regresi Spline adalah teknik estimasi nonparametrik lain yang populer. splines umumnya didefinisikan sebagai subjek *piecewise polinomial* kendala kontinuitas tertentu. Model regresi spline sangat mirip dengan dummy (biner) model variabel. Menggunakan sebuah variabel dummy untuk mengubah *intercept* atau kemiringan model umumnya menghasilkan patahan-patahan digaris regresi (Gao 2008). Splines memiliki kemampuan untuk mengumpulkan segmen polinomial yang dihasilkan oleh variabel dummy untuk menghilangkan patahan buatan di garis regresi.

Menurut Eubank (1988) regresi spline merupakan salah satu teknik estimasi dalam regresi nonparametrik dengan model polinomial yang memiliki sifat tersegmen atau terpotong-potong yang mulus. Apabila regresi $f(t_i)$ pada persamaan $Y_i = f(t_i) + \varepsilon_i$; $i = 1, 2, \dots, n$ dihipotesiskan fungsi spline, maka untuk mengatasi $f(t_i)$ dapat digunakan dengan model regresi spline. Secara umum model regresi spline pada suatu fungsi dengan orde p dapat dirumuskan dalam bentuk persamaan berikut.

$$f(t_i) = \sum_{j=0}^{p-1} \beta_j t_i^j + \sum_{l=1}^r \beta_{(p-1+l)} (t_i - k_l)_+^{p-1} \quad (2.7)$$

Keterangan : $\beta_0, \beta_1, \beta_2, \dots, \beta_{p-1}, \beta_{(p-1+l)}$ adalah parameter koefisien regresi, $t_i^1, t_i^2, \dots, t_i^{p-1}$ adalah peubah bebas yang diketahui nilainya, k adalah banyaknya titik knot $k_1, k_2, k_3, \dots, k_r$, dan $(t_i - k_l)_+^{p-1}$ adalah fungsi potongan (*truncated*) yang dapat diuraikan sebagai berikut. (Fiatussholikhah 2014) (Zheng 2003).

fungsi *truncated*

$$(t_i - k_l)_+^{p-1} = \begin{cases} (t_i - k_l)^{p-1}; & t_i \geq k_l \\ 0 & t_i < k_l \end{cases} \quad (2.8)$$

Apabila persamaan $y_i = f(t_i) + \varepsilon_i, i = 1, 2, \dots, n$ disubstitusikan kedalam persamaan $f(t_i) = \sum_{j=0}^{p-1} \beta_j t_i^j + \sum_{l=1}^r \beta_{(p-1+l)} (t_i - k_l)_+^{p-1}$ maka akan diperoleh persamaan regresi nonparametrik *spline runcated* sebagai berikut (Kuzelka 2014).

$$Y_i = \sum_{j=0}^{p-1} \beta_j t_i^j + \sum_{l=1}^r \beta_{(p-1+l)} (t_i - k_l)_+^{p-1} + \varepsilon_i \text{ dengan } i = 1, 2, 3, \dots, n. \quad (2.9)$$

2.6 Generalized Cross Validation (GCV)

Pada pemodelan regresi nonparametrik untuk menentukan parameter pemulus adalah dengan menggunakan spline. Titik knot merupakan titik perpaduan bersama yang memperlihatkan terjadinya perubahan perilaku dari fungsi spline pada interval-interval yang berbeda sehingga kurva yang terbentuk tersegmen pada titik tersebut. Estimator spline terbaik diperoleh dengan

menggunakan titik knot optimal. Titik knot merupakan titik perpaduan bersama dimana terdapat perubahan pola perilaku fungsi atau kurva. Titik knot optimal dapat diperoleh dengan menggunakan metode *Generalized Cross Validation* (GCV). Pada penentuan model regresi spline dapat dilakukan dengan melihat nilai Generalized Cross Validation (GCV) yang minimum. Adapun rumus untuk menghitung GCV adalah sebagai berikut (Eubank, 1988) dan (Kang, 2014) .

$$GCV(k) = \frac{MSE(k)}{\left[n^{-1} \text{trace}(I - A(k)) \right]^2} \quad (2.10)$$

Keterangan : $MSE(k) = n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2$

$$A(k) = X(k) \left(X(k)' X(k) \right)^{-1} X(k)'$$

2.7 Pengangguran Terbuka

Pengangguran terbuka didefinisikan sebagai orang yang sedang mencari pekerjaan atau yang sedang mempersiapkan usaha atau juga yang tidak mencari pekerjaan karena merasa tidak mungkin lagi mendapatkan pekerjaan, termasuk juga mereka yang baru mendapat kerja tetapi belum mulai bekerja. Pengangguran terbuka tidak termasuk orang yang masih sekolah atau mengurus rumah tangga, sehingga hanya orang yang termasuk angkatan kerja saja yang merupakan pengangguran terbuka.

Definisi pengangguran menurut BPS pengangguran terbuka (*open unemployment*) didasarkan pada konsep seluruh angkatan kerja yang mencari pekerjaan, baik yang mencari pekerjaan pertama kali maupun yang pernah

bekerja sebelumnya. Sedang pekerja yang digolongkan setengah penganggur (*underemployment*) adalah pekerja yang masih mencari pekerjaan penuh atau sambilan dan mereka yang bekerja dengan jam kerja rendah. Setengah penganggur sukarela adalah setengah penganggur tetapi tidak mencari pekerjaan atau tidak bersedia menerima pekerjaan lain. Setengah penganggur terpaksa adalah setengah penganggur yang masih mencari pekerjaan atau bersedia menerima pekerjaan. Pekerja digolongkan setengah penganggur parah (*severe underemployment*) apabila ia masuk setengah menganggur dengan jam kerja kurang dari 25 jam seminggu.

a. Pengangguran dalam Sektor Informal

Pengangguran terbuka biasanya terjadi pada generasi muda yang baru menyelesaikan pendidikan menengah dan tinggi. Ada kecenderungan mereka yang baru menyelesaikan pendidikan berusaha mencari kerja sesuai dengan aspirasi mereka. Aspirasi mereka biasanya adalah bekerja di sektor modern atau di kantor. Untuk mendapatkan pekerjaan itu mereka bersedia menunggu untuk beberapa lama. Tidak tertutup kemungkinan mereka berusaha mencari pekerjaan itu di kota atau di provinsi atau daerah yang kegiatan industri telah berkembang. Hal ini menyebabkan angka pengangguran tinggi di perkotaan atau di daerah kegiatan industri atau sektor modern berkembang. Sebaliknya pengangguran terbuka rendah di daerah atau provinsi yang tumpu pada sektor pertanian. Hal tersebut penyediaan pekerjaan di sektor informal oleh sebab rendahnya pendidikan dan kurang menjamin kelangsungan hidup.

b. Pengukuran Tingkat Pengangguran

Badan statistik negara mengelompokkan orang dewasa pada setiap rumah tangga yang disurvei ke dalam satu kategori berikut.

1. Bekerja
2. Pengangguran
3. Tidak termasuk angkatan kerja

Setelah mengelompokkan seluruh individu yang disurvei ke dalam tiga kategori tersebut, badan statistik negara menghitung berbagai statistik untuk merangkum kondisi angkatan kerja. Angkatan kerja (*labor force*) adalah jumlah orang yang berkerja dan tidak berkerja.

Angkatan kerja = Jumlah orang yang bekerja + Jumlah yang tidak bekerja.

Tingkat pengangguran (*unemployment rate*) adalah persentase angkatan kerja yang tidak bekerja:

$$\text{Tingkat pengangguran} = \frac{\text{jumlah orang yang tidak bekerja}}{\text{angkatan kerja}} \times 100 \quad (2.11)$$

Setelah itu, tingkat pengangguran untuk seluruh populasi penduduk dewasa dan untuk kelompok yang lebih sempit, seperti laki-laki dan perempuan dapat dihitung.

Tingginya angka pengangguran dipengaruhi oleh presentase penduduk usia kerja berumur 15 tahun keatas berdasarkan faktor-faktor yang mempengaruhi antara lain :

1. Hubungan Nilai PDRb dengan Pengangguran

Nilai PDRB menunjukkan kemajuan perekonomian di wilayah tersebut, semakin tinggi nilai PDRB semakin sejahtera kehidupan di wilayah tersebut, sehingga TPT menurun. Berdasarkan BPS Provinsi Jawa Timur, PDRB merupakan penjumlahan nilai output bersih perekonomian yang ditimbulkan oleh seluruh kegiatan ekonomi di suatu wilayah tertentu (Provinsi dan Kabupaten/Kota), dan dalam satu kurun waktu tertentu (satu tahun kelender). Seorang peneliti melakukan penelitian tentang Analisis Faktor-Faktor yang Mempengaruhi Kesempatan Kerja pada Kabupaten/Kota di Propinsi Sumatera Utara yang menjadi rujukan dan persamaan dalam penelitian ini menunjukkan bahwa terdapat adanya pengaruh PDRB dan jumlah pengangguran yang bersifat positif dalam Teori Pertumbuhan Ekonomi. Dikatakan berpengaruh positif sebab pertumbuhan ekonomi tidak dibarengi oleh peningkatan kapasitas produksi, sehingga jumlah pengangguran tetap meningkat seiring pertumbuhan ekonomi yang berlangsung. Hal ini disebabkan pertumbuhan ekonomi yang meningkat tersebut berorientasi pada padat modal, dimana kegiatan produksi untuk memacu output dan menghasilkan pendapatan yang meningkat lebih diutamakan ketimbang pertumbuhan ekonomi yang berorientasi pada padat karya.

Populasi Penduduk Dalam arti luas, penduduk atau populasi berarti sejumlah makhluk sejenis yang mendiami atau menduduki tempat tertentu. Bahkan populasi dapat pula dikenakan pada benda-benda sejenis yang terdapat pada suatu tempat. Dalam kaitannya dengan manusia, maka pengertian penduduk adalah manusia yang mendiami dunia atau bagian bagiannya. Menurut Marxist

tekanan penduduk di suatu negara bukanlah tekanan penduduk terhadap bahan makanan, tetapi tekanan terhadap kesempatan kerja (misalnya di negara kapitalis). Marxist juga berpendapat bahwa semakin banyak jumlah manusia semakin tinggi produk yang dihasilkan, jadi dengan demikian tidak perlu diadakan pembatasan penduduk.

2. Hubungan Prosentase kesempatan kerja dengan Pengangguran

Jumlah angkatan kerja yang bekerja biasanya dipandang sebagai jumlah kesempatan yang tersedia di satu wilayah. Dalam pengertian kesempatan kerja tidak sama dengan lapangan kerja yang masih terbuka, yang dimaksud dengan bekerja dalam sensus penduduk 1980 ialah selama seminggu sebelum pencacahan melakukan pekerjaan dengan maksud memperoleh atau membantu memperoleh penghasilan atau keuntungan dan lamanya bekerja paling sedikit satu jam dalam sehari. Berbeda dengan definisi sebelumnya, Sensus penduduk 1971 ialah batas waktu yang dipergunakan minimal dua hari dalam seminggu dengan bekerja paling sedikit satu jam dalam sehari.

Pertumbuhan ekonomi yang pesat mendorong penyediaan berbagai sarana dan prasarana perekonomian yang dibutuhkan untuk mempercepat pembangunan ekonomi. Indikator pembangunan ekonomi suatu negara terletak pada pertumbuhan ekonomi dan kesempatan kerja. Menurut kaum klasik Adam Smith mengatakan bahwa alokasi sumber daya manusia yang efektif adalah pemula pertumbuhan ekonomi. Setelah ekonomi tumbuh, akumulasi modal (fisik) baru mulai dibutuhkan untuk menjaga agar ekonomi tetap tumbuh. Dengan kata lain,

alokasi sumber daya manusia yang efektif merupakan syarat perlu (*necessary condition*) bagi pertumbuhan ekonomi untuk memperluas kesempatan kerja.

Kesempatan kerja adalah banyaknya orang atau tenaga kerja yang dapat terserap untuk bekerja pada suatu perusahaan atau suatu instansi, kesempatan kerja ini akan menyerap semua tenaga kerja yang tersedia apabila lapangan pekerjaan yang tersedia mencukupi atau seimbang dengan banyaknya tenaga kerja yang tersedia. Menurut Tambunan (2001) dengan adanya ketersediaan lapangan kerja yang memadai tentu akan mempercepat proses pembangunan ekonomi apalagi disertai dengan produktivitas yang tinggi. Suatu peluang untuk mendapatkan pekerjaan diberbagai sektor ekonomi disebut dengan kesempatan kerja.

