

BAB II

TINJAUAN PUSTAKA

2.1. Analisis Regresi

Analisis regresi merupakan salah satu teknik analisis data dalam statistika yang paling banyak digunakan untuk mengetahui pola hubungan antara variabel respon terhadap satu atau beberapa variabel prediktor. Regresi pertama kali diperkenalkan pada tahun 1886 oleh seorang ahli yang bernama Francis Galton. Menurut Galton, analisis regresi berkenaan dengan studi ketergantungan dari suatu variabel yang disebut variabel tak bebas (*dependet variable*), pada satu atau variabel yang menerangkan dengan tujuan untuk memperkirakan ataupun meramalkan nilai-nilai dari variabel tak bebas apabila nilai variabel yang menerangkan sudah diketahui. Variabel yang menerangkan sering disebut variabel bebas (*independent variable*). Hubungan antar kedua variabel tersebut dapat digambarkan oleh kurva regresi dengan bentuk fungsi tertentu. Pendekatan yang dapat digunakan untuk mengestimasi kurva regresi ada dua jenis, yaitu :

- a. Pendekatan Regresi Parametrik
- b. Pendekatan Regresi Nonparametrik

Regresi parametrik digunakan untuk mengestimasi bentuk hubungan antara variabel respon dengan satu atau beberapa variabel prediktor dimana bentuk kurva regresi diketahui. Pada regresi parametrik terdapat asumsi-asumsi yang harus terpenuhi. Regresi nonparametrik merupakan metode statistika yang digunakan

apabila bentuk hubungan antara variabel respon dengan variabel prediktornya tidak diketahui.

2.2. Regresi Nonparametrik

Regresi nonparametrik dapat digunakan apabila hubungan antara variabel respon dan variabel prediktor memiliki bentuk pola yang tidak diketahui kurva regresinya. Regresi nonparametrik juga dapat digunakan pada data yang memiliki distribusi normal ataupun tidak. Istilah nonparametrik pertama kali diperkenalkan oleh Wolfowitz pada tahun 1942. Pendekatan nonparametrik merupakan pendekatan regresi yang sesuai untuk pola data yang tidak diketahui bentuknya, atau tidak terdapat informasi masa lalu tentang pola data (Budiantara, 2010). Regresi nonparametrik berbeda dengan regresi parametrik, dalam regresi parametrik bentuk kurva regresi diasumsikan mengandung parameter sehingga perlu dilakukannya estimasi parameter untuk memperoleh estimator kurva regresi. Sedangkan regresi nonparametrik tidak membutuhkan asumsi mengenai bentuk kurva regresi. Oleh karena itu, regresi nonparametrik bersifat lebih fleksibel terhadap perubahan pola data (Eubank, 1988).

Secara umum model regresi nonparametrik dapat di tuliskan tersebut sebagai berikut :

$$y = M(X_i) + \varepsilon_i \quad (1)$$

Dimana :

y = Variabel terikat.

$M(X_i)$ = Kurva regresi

ε_i = Error yang diasumsikan independen dengan mean nol dan variansi σ_i .

Budiantara (2010) mengungkapkan bahwa terdapat beberapa teknik untuk mengestimasi kurva regresi dalam regresi nonparametrik, yaitu histogram, *spline*, Deret *Fourier*, *Wavelets*, *orthogonal* dan kernel.

2.3. Estimator Kernel *Nadaraya-Watson*

Model pendekatan nonparametrik yang banyak digunakan adalah estimator kernel. Estimator kernel diperkenalkan oleh Rosenblatt dan Parzen sehingga disebut estimator densitas kernel Rosenblatt-Parzen (Hardle,1994). Menurut Eubank (1999) estimator kernel merupakan estimator linier yang sama dengan estimator lainnya, perbedaannya hanya karena metode kernel lebih khusus dalam penggunaan metode *bandwidth*. Beberapa kelebihan estimator kernel adalah fleksibel, bentuk matematisnya mudah, dan dapat mencapai tingkat kekonvergenan yang relatif cepat (Budiantara dan Mulianah, 2007).

Menurut Halim dan Bisono (2006) terdapat beberapa macam estimator kernel, salah satunya adalah estimator *Nadaraya-Watson* yang didefinisikan :

$$m(x) = \frac{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) y_i}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)} \quad (2)$$

Keterangan:

Y_i = Variabel respon

x_i = Variabel prediktor

K = Fungsi kernel

h = *Bandwidth* atau *smoothing* parameter

$$s_0 = 0 \quad (3)$$

$$s_i = \frac{(x_i + x_{i+1})}{2} \quad (4)$$

$$i=1, \dots, n-1 \text{ dan } s_n=1$$

Secara umum estimator kernel *Nadaraya-Watson* merupakan estimator yang paling sering digunakan. Seperti persamaan (1), yang terdapat fungsi bobot :

$$m(x) = w_{hi}(x) y_i \quad (5)$$

$$w_{hi}(x) = \frac{\frac{1}{h} K\left(\frac{x-x_i}{h}\right)}{f_h(x)} = \frac{K\left(\frac{x-x_i}{h}\right)}{f_h(x)} \quad (6)$$

Fungsi ini tergantung pada kernel K . Dimana $f_h(x)$ adalah estimator densitas kernel

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n K_h\left(\frac{x-x_i}{h}\right) = \frac{1}{h} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (7)$$

Sehingga

$$w_{hi}(x) = \frac{\frac{1}{h} K\left(\frac{x-x_i}{h}\right)}{\frac{1}{h} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)} = \frac{K\left(\frac{x-x_i}{h}\right)}{\sum_{i=1}^n K\left(\frac{x-x_i}{h}\right)} \quad (8)$$

Akan menjadi

$$m(x) = \frac{\sum_{i=1}^n K_h\left(\frac{x-x_i}{h}\right) y_i}{\sum_{i=1}^n K_h\left(\frac{x-x_i}{h}\right)} \quad (9)$$

Kemudian dari persamaan (1) akan menjadi

$$y = \frac{\sum_{i=1}^n K_h \left(\frac{x - x_i}{h} \right) y_i}{\sum_{i=1}^n K_h \left(\frac{x - x_i}{h} \right)} + \varepsilon \quad (10)$$

(Kurniasih, 2013)

Estimasi estimator kernel *Nadaraya-Watson* akan diestimasi menggunakan *Weighted Least Square*. Metode estimasi *Weighted Least Square* (WLS) merupakan salah satu metode estimasi parameter yang digunakan untuk menentukan nilai taksiran parameter yang meminimumkan jumlah kuadrat residu dengan dilakukan pembobotan pada suatu faktor yang tepat. Berikut merupakan perumusan dari metode *Weighted Least Square* (WLS) :

$$WLS = \sum_{i=1}^n (y_i - \sum_{j=1}^2 m_j(t_j))^2 K_h(x_i - x) \quad (11)$$

2.4. Fungsi Kernel

Selain estimator kernel terdapat juga fungsi kernel. Menurut Hardle (1990), secara umum fungsi kernel didefinisikan sebagai berikut :

$$K_h(x) = \frac{1}{h} K \left(\frac{x}{h} \right), \text{ untuk } -\infty < x < \infty, h > 0 \quad (12)$$

Dimana :

K = Fungsi Kernel

h = *Bandwidth* atau parameter pemulus

Suatu fungsi kernel harus memenuhi beberapa syarat, yaitu :

1. $K(x) \geq 0$ untuk semua x .
2. $\int_{-\infty}^{\infty} K(x)dx = 1$
3. $\int_{-\infty}^{\infty} x^2 K(x)dx = \sigma^2 > 0$
4. $\int_{-\infty}^{\infty} xK(x)dx = 0$

(Hansen, 2009)

Kriteria pemilihan fungsi kernel yang baik berdasarkan pada resiko kernel minimum yang dapat diperoleh dari kernel optimal atau variansi minimum. Menurut Hardle (1994) dan Sukarsa (2012) terdapat macam-macam fungsi kernel, yaitu :

1. Kernel Uniform $K(x) = \frac{1}{2} I(|x| \leq 1)$
2. Kernel Triangular $K(x) = (1 - |x|) I(|x| \leq 1)$
3. Kernel Epanechnikov $K(x) = \frac{3}{4} (1 - x^2) I(|x| \leq 1)$
4. Kernel Kuadrat (Quartik) $K(x) = \frac{15}{16} (1 - x^2)^2 I(|x| \leq 1)$
5. Kernel Triweight $K(x) = \frac{35}{32} (1 - x^2)^3 I(|x| \leq 1)$
6. Kernel Gaussian $K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) -\infty < x < \infty$
7. Kernel Cosinus $K(x) = \frac{\pi}{4} \cos\left(\frac{\pi}{2}x\right) I(|x| \leq 1)$

2.5. Pemilihan *Bandwidth* Optimal

Bandwidth (h) merupakan parameter pemulus (*smoothing*) yang berfungsi untuk mengontrol kemulusan dari kurva yang diestimasi. Pemilihan *bandwidth* jauh lebih penting daripada pemilihan fungsi kernel. Jika *bandwidth* yang terlalu kecil akan menghasilkan kurva *under-smoothing* yang sangat kasar dan sangat fluktuatif dan sebaliknya *bandwidth* terlalu lebar akan menghasilkan kurva *over-smoothing* yaitu sangat mulus, tetapi tidak sesuai dengan pola data (Hardle, 1994). Oleh karena itu harus dipilih nilai *bandwidth* yang optimal agar menghasilkan estimasi terbaik (Mustika, dkk., 2005). Salah satu metode yang dapat digunakan dalam pemilihan *bandwidth* yang optimal adalah dengan menggunakan kriteria *Generalized Cross Validation* (GCV), karena metode GCV lebih optimal secara asimtotik dan tidak memuat parameter populasi yang didefinisikan sebagai berikut :

$$GCV = \frac{MSE}{\left(\frac{1}{n} \text{tr}(I - H(h))\right)^2} \quad (13)$$

Dimana :

n = Banyaknya data

I = Matriks identitas

h = *Bandwidth*

X = Matrik data

$H(h) = X(X'X + nhI)^{-1}X'$ (Michael H., dkk., 1979)

Salah satu dari beberapa kriteria untuk menentukan estimator terbaik dalam model regresi nonparametrik adalah *Mean Square Error* (MSE). Penggunaan *bandwith* yang optimal akan menghasilkan estimasi dengan MSE yang terkecil. Secara umum *Mean Square Error* (MSE) didefinisikan dengan :

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (14)$$

2.6. Kemiskinan

2.6.1 Definisi Kemiskinan

Kemiskinan merupakan salah satu persoalan yang menjadi prioritas utama bagi setiap negara. Menurut Piven dan Cloward dan Swanson dalam Suharto (2009) kemiskinan menggambarkan adanya kelangkaan materi atau barang-barang yang diperlukan dalam kehidupan sehari-hari, seperti makanan, pakaian dan perumahan. Kemiskinan berasal dari kata “miskin“ yang dapat diartikan sebagai serba kekurangan, sementara “kemiskinan” dipandang sebagai kondisi ketidakmampuan secara ekonomi dalam memenuhi standar kebutuhan sehari-hari baik berupa sandang, pangan dan papan. Menurut Anggraeny (2016) kemiskinan adalah suatu kondisi ketidakmampuan secara ekonomi untuk memenuhi standar hidup rata-rata masyarakat di suatu daerah. Kondisi ketidakmampuan ini ditandai dengan kemampuan pendapatan untuk memenuhi kebutuhan pokok baik berupa pangan, sandang, maupun papan. Kemiskinan didefinisikan sebagai ketidakmampuan individu dalam memenuhi kebutuhan dasar minimal untuk hidup layak (BPS dan Depsos, 2002). Bappenas (2005) melihat kemiskinan sebagai

suatu kondisi dimana seseorang atau sekelompok orang, laki-laki dan perempuan, tidak mampu memenuhi hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat. Dari beberapa definisi tersebut dapat disimpulkan bahwa kemiskinan merupakan suatu keadaan kurangnya materi seseorang ataupun golongan sehingga tidak dapat memenuhi kebutuhan dalam kehidupan sehari-hari. Pada dasarnya definisi kemiskinan bermacam-macam akan tetapi secara garis besar kemiskinan selalu berkaitan dengan masalah kesejahteraan. Kemiskinan dapat dilihat dari dua sisi, yaitu kemiskinan absolut dan kemiskinan relatif. Mengukur kemiskinan dengan mengacu pada garis kemiskinan disebut kemiskinan absolut, sedangkan konsep kemiskinan yang pengukurannya tidak didasarkan pada garis kemiskinan disebut kemiskinan relatif (Tambunan, 2001).

a. Kemiskinan Absolut

Kemiskinan absolut merupakan ketidakmampuan individu untuk memenuhi kebutuhan dasar minimum dalam kehidupan sehari-hari dengan pendapatan yang diperoleh berada dibawah garis kemiskinan. Badan Pusat Statistik (BPS) menentukan kemiskinan absolut Indonesia merupakan ketidakmampuan seseorang untuk mencukupi kebutuhan pokok minimum energi kalori (2.100 kilo kalori per kapita per hari) yang dipergunakan tubuh dan kebutuhan dasar minimum untuk sandang, perumahan, kesehatan, pendidikan, transportasi dan kebutuhan dasar lain.

b. Kemiskinan Relatif

Menurut Todaro dan Smith (2006) kemiskinan relatif merupakan ukuran kesenjangan dalam distribusi pendapatan, biasanya terkait dengan ukuran dibawah tingkat rata-rata distribusi pendapatan nasional. Kategori kemiskinan relatif ditentukan berdasarkan perbandingan relatif tingkat kesejahteraan antar penduduk (Hendra, 2011). Seseorang dapat dikatakan masuk sebagai katagorimiskin relatif apabila telah mampu memenuhi kebutuhan dasar hidupnya akan tetapi masih lebih rendah dibandingkan dengan keadaan masyarakat sekitarnya.

2.6.2. Ciri-ciri Kemiskinan

Badan Pusat Statistik telah menetapkan 14 kriteria rumah tangga yang masuk kategori miskin, yaitu sebagai berikut :

- a. Luas lantai bangunan tempat tinggal kurang dari 8 m² per orang.
- b. Jenis lantai bangunan tempat tinggal terbuat dari tanah/bambu/kayu murahan.
- c. Jenis dinding tempat tinggal terbuat dari bambu/rumbia/kayu berkualitas rendah/tembok tanpa diplester.
- d. Tidak memiliki fasilitas buang air besar/bersama-sama dengan rumah tangga lain.
- e. Sumber penerangan rumah tangga tidak menggunakan listrik.
- f. Sumber air minum bersumber dari sumur/mata air tidak terlindungi/sungai/air hujan.

- g. Bahan bakar untuk memasak sehari-hari adalah kayu bakar/arang/minyak tanah.
- h. Hanya mengkonsumsi daging/susu/ayam satu kali dalam seminggu.
- i. Hanya membeli satu stel pakaian baru dalam setahun .
- j. Hanya sanggup makan sebanyak satu/dua kali dalam sehari.
- k. Tidak sanggup membayar biaya pengobatan di puskesmas/poliklinik.
- l. Sumber penghasilan kepala rumah tangga adalah : petani dengan luas lahan 0,5 ha. Buruh tani, nelayan, buruh bangunan, buruh perkebunan atau pekerjaan lainnya dengan pendapatan dibawah Rp 600.000 per bulan.
- m. Pendidikan tertinggi kepala rumah tangga : tidak bersekolah/tidak tamat SD/ hanya SD.
- n. Tidak memiliki tabungan/barang yang mudah dijual dengan nilai Rp 500.000, seperti sepeda motor (kredit / non kredit), emas, ternak, kapal motor atau barang modal lainnya.

2.6.3. Faktor Penyebab Kemiskinan

Berbagai penelitian telah banyak dilakukan untuk mengetahui penyebab dan faktor-faktor yang terkait dengan kemiskinan diantaranya :

a. Angka Partisipasi Sekolah

Menurut BPS (2015) Angka Partisipasi Sekolah proporsi anak sekolah pada usia jenjang pendidikan tertentu dalam kelompok usia yang sesuai dengan jejang pendidikan tersebut. Partisipasi sekolah berkaitan dengan aktivitas pendidikan formal dan non formal seseorang. APS

merupakan indikator penting dalam pendidikan yang menunjukkan persentase penduduk usia 7-12 tahun yang masih terlibat dalam sistem persekolahan. Dalam penelitian ini, angka partisipasi sekolah yang digunakan adalah persentase penduduk usia 7-12 tahun yang masih terlibat dalam sistem persekolahan tahun 2015.

b. Rata-rata Lama Sekolah

Orang yang memiliki tingkat pendidikan lebih tinggi, diukur dengan lamanya waktu untuk sekolah akan memiliki pekerjaan dan upah yang lebih baik dibanding dengan orang yang pendidikannya rendah (Atmanti, 2005). Rata-rata lama sekolah adalah rata-rata jumlah tahun yang dihabiskan oleh penduduk berusia 15 tahun ke atas untuk menempuh semua jenis pendidikan formal. Rata-rata lama sekolah menggambarkan jumlah tahun yang digunakan oleh penduduk usia 15 tahun ke atas dalam menjalani pendidikan formal. Rata-rata lama sekolah diukur dari rata-rata jumlah tahun yang dihabiskan oleh penduduk usia 15 tahun keatas di masing-masing Kabupaten/Kota di suatu wilayah. Rata-rata lama sekolah yang digunakan dalam penelitian ini adalah jumlah tahun yang digunakan oleh penduduk usia 15 tahun ke atas dalam menjalani pendidikan formal.

c. Tingkat Pengangguran Terbuka

Pengangguran terbuka merupakan penduduk yang sedang mencari pekerjaan atau mempersiapkan suatu usaha atau merasa tidak mungkin mendapatkan pekerjaan atau sudah punya pekerjaan tetapi belum mulai

bekerja (BPS, 2015). Pengangguran terbuka timbul karena pertumbuhan kesempatan kerja lebih rendah dari pada pertumbuhan tenaga kerja yang mengakibatkan banyaknya tenaga kerja yang tidak memperoleh pekerjaan. Tingkat pengangguran terbuka adalah adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja (BPS, 2015). Islam (2003) melakukan penelitian di 23 negara berkembang dan menyimpulkan bahwa kemiskinan dapat berkurang seiring dengan peningkatan pendidikan (menurunnya persentase buta huruf) dan peningkatan persentase tenaga kerja di sektor industri. Pada penelitian ini digunakan data tingkat pengangguran terbuka tahun 2015 dalam bentuk persentase.

d. Laju Pertumbuhan Ekonomi

Pertumbuhan ekonomi merupakan indikator untuk melihat keberhasilan pembangunan. Laju pertumbuhan ekonomi suatu daerah dapat dilihat dari nilai produk domestik regional bruto (PDRB) daerah tersebut. Menurut BPS, PDRB didefinisikan sebagai jumlah nilai tambah yang dihasilkan oleh seluruh unit dalam suatu wilayah atau merupakan jumlah seluruh nilai barang dan jasa akhir yang dihasilkan oleh unit ekonomi suatu wilayah. Terdapat 2 jenis PDRB, yaitu PDRB atas dasar harga berlaku dan PDRB atas dasar harga konstan. Menurunnya PDRB suatu daerah berdampak pada kualitas dan pada konsumsi rumah tangga. Dalam penelitian ini, untuk mengetahui laju pertumbuhan ekonomi Jawa Tengah menggunakan PDRB atas dasar

harga konstan tahun 2010 menurut Kabupaten dan Kota di Provinsi Jawa Tengah yang dinyatakan dalam persen tahun 2015.

