BAB III

METODE PENELITIAN

3.1. Variabel Penelitian dan Definisi Operasional Variabel

3.1.1. Variabel Dependen

Dalam penelitian ini variabel dependennya adalah jumlah penduduk miskin Kabupaten/Kota di Provinsi DIY. Jumlah penduduk miskin Kabupaten/Kota di Provinsi DIY adalah total penduduk yang berpendapatan kurang dari rata-rata pendapatan rata-rata masyarakat di daerah di mana seseorang tinggal, di mana menurut BPS penduduk miskin di DIY dengan pendapatan Rp.335.886,- per kapita per bulan, sedangkan pada tahun 2014 sebesar Rp 321.056,- per kapita per bulan (BPS, 2015). Jumlah penduduk miskin di Provinsi DIY dalam kurun waktu dari Tahun 2009 sampai dengan Tahun 2013. Satuan jumlah penduduk miskin Kabupaten/Kota di Provinsi DIY adalah jiwa.

3.1.2. Variabel Independen

Variabel independen adalah variabel yang mempengaruhi. Dalam penelitian ini variabel independen yang digunakan adalah:

1. Pertumbuhan Ekonomi

Pertumbuhan ekonomi adalah proses kenaikkan output per kapita dalam jangka panjang. Pertumbuhan ekonomi dalam penelitian ini pada kurun waktu dari Tahun 2009 sampai dengan Tahun 2013. Satuan pertumbuhan ekonomi adalah persen.

2. Rata-rata Lama Sekolah

Rata-rata lama sekolah adalah jumlah penduduk kabupaten/kota di propinsi DIY yang tamat atau lulus SD, SLTP, SLTA, dan PT. Rata-rata lama sekolah dalam penelitian ini pada kurun waktu dari Tahun 2009 sampai dengan Tahun 2013. Satuan rata-rata lama sekolah adalah tahun.

3. Jumlah Pengangguran

Jumlah pengangguran adalah banyaknya atau persentase jumlah penganggur terhadap jumlah angkatan kerja. Penduduk yang sedang mencari pekerjaan, tetapi tidak sedang mempunyai pekerjaan disebut penganggur. Jumlah pengangguran Kabupaten/Kota di Provinsi DIY dalam penelitian ini pada kurun waktu dari Tahun 2009 sampai dengan Tahun 2013. Satuan jumlah pengangguran adalah jiwa.

3.2. Jenis dan Sumber Data

Dalam penelitian ini penulis menggunakan data sekunder *pooling time series* (data panel) yang terdiri dari variabel dependen yaitu jumlah penduduk miskin Kabupaten/Kota di Provinsi DIY dan variabel independen yaitu pertumbuhan ekonomi, rata-rata lama sekolah, dan jumlah pengangguran. Data ini diperoleh dari Badan Biro Pusat Statistik (BPS) serta pihak lain yang berkompeten dengan publikasi data yang relevan dengan dengan penelitian ini dengan kurun waktu antara Tahun 2009 sampai dengan Tahun 2013 dalam bentuk tahunan.

3.3. Metode Pengumpulan Data

Metode pengumpulan data dalam penelitian ini menggunakan suryei data di Badan Biro Pusat Statistik (BPS) serta pihak lain yang berkompeten dengan publikasi data yang relevan dengan penelitian ini.

3.4. Metode Analisis Data

3.4.1. Analisis Regresi Linier

Model yang digunakan dalam analisis ini yaitu model persamaan linier berganda untuk mengetahui pengaruh antara pertumbuhan ekonomi, rata-rata lama sekolah, dan jumlah pengangguran terhadap jumlah penduduk miskin Kabupaten/Kota di Provinsi DIY dengan persamaan atau model linier sebagai berikut:

$Y = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + e_i$..(1)

Model persamaan *log linier* dengan tujuan untuk menyamakan atau memperkecil variasi data dan untuk menghindari terjadinya penyakit asumsi klasik, sehingga terjadinya perubahan pada variabel independen akan menyebabkan perusahaan pada variabel dependen secara absolut untuk melihat elastisitas. Berikut ini model persamaan *log linier*:

$LnY = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 LnX_{3it} + e_i$ ....(2)

Keterangan :

Y = Jumlah Penduduk Miskin Kabupaten/Kota di Provinsi DIY (jiwa)

X_1 = Pertumbuhan Ekonomi (persen)

X_2 = Rata-rata Lama Sekolah (tahun)

X_3 = Jumlah Pengangguran (jiwa)

Ln = Logartima Natural (persen)

b_0 = Konstanta

b_{1-3} = Koefisien Regresi

$_{it}$ = Kabupaten/Kota

ε = *Error term*

Parameter β_1, $\beta_{o2,}$ β_{o3} merupakan koefisien dari masing-masing variabel independen dari X_1 hingga X_3. dalam hal ini variabel jumlah penduduk miskin di daerah Kabupaten/Kota di Propinsi DIY (Y) menyatakan variabel dependen.

Parameter variabel independen akan diestimasi dengan menggunakan analisis data panel dengan menggabungkan data *cross section* dan *time series*. Adapun alasan mengapa yang dipilih adalah analisis panel data adalah sebagai berikut: (Gujarati, 2009: 71).

1. Dengan *ordinary least square* (OLS) biasa dilakukan terpisah diasumsikan bahwa parameter regresi tidak berubah antar waktu (*temporal stability*) dan tidak berubah antar unit-unit individualnya (*cross sectional unit*).
2. Dengan OLS biasa akan terjadi asumsi yang sangat sempit tentang asumsi klasik yaitu *homoscedastisiti* dan *non autocorrelation* (homoskedastisitas dan tidak berkolerasi pada variabel kesalahan) pembentukan model dengan menggabungkan data *time series* dan *cross section*.

Penggunaan data panel dalam penelitian ekonomi mempunyai beberapa keunggulan dibandingkan dengan *data time series* atau data *cross section* biasa (Hsiao, 1995).

1. Memberikan kepada peneliti sejumlah data yang banyak meningkatkan derajat kebebasan (*degress of freedom*) dan mengurangi kolinieritas (hubungan) diantara variabel penjelas (*explanatory variables*), sehingga akan menghasilkan estimasi ekonometrik yang efisien.
2. Data longitudinal membolehkan peneliti untuk menganalisis dengan menggunakan data *cross sectional* atau *time series*.

β_0, λ_i adalah koefisien-koefisien regresi yang akan ditaksir. Secara umun model regresi mempunyai bentuk sebagai berikut:

$$Y_{it} = \alpha_1 + \beta x_{it} + \mu_{it} \quad \text{...} \quad (3)$$

Dimana :

i	: 1 , 2 , ... n menunjukkan pada *cross section*
t	: 1, 2 , ... t menunjukkan pada suatu waktu tertentu
Y_{it}	: nilai dari dependen variable dari daerah i pada waktu t ada sejumlah K pada x_{it} , tidak termasuk *constant term*
α_1	: *individual effect* yang konstan antar waktu t dan spesifik untuk masing-masing unit *cross section* i
β	: koefisien regresi
μ_{it}	: variabel pengganggu

Model seperti di atas adalah model regresi klasik.

Jika kita menganggap α_1 adalah sama untuk semua unit individu, maka *ordinary least square* (OLS) memberikan estimasi yang efisien untuk parameter α_1 dan β. Dalam menggunakan panel data terdapat dua pendekatan mendasar yaitu:

1. Pendekatan *fixed effect* yang menetapkan bahwa α_1 adalah sebagai kelompok yang spesifik atau berbeda dalam *constant term* dalam model regresinya.
2. Pendekatan *random effect* yang meletakkan α_1 adalah gangguan spesifik kelompok, sama dengan μ_{it} kecuali untuk masing-masing kelompok.

3.4.2. Pemilihan Model dengan Hausman's *Test*

Sebelum model diestimasi dengan model yang tepat, maka akan dilakukan terlebih dulu uji spesifikasi model apa yang akan dipakai, apakah *common model, fixed effect* atau *random effect* atau ketigaya memberikan hasil yang sama. Pilihan antara *fixed effect* dan *random effect* ditentukan dengan menggunakan Hausman's *Test* atau masing-masing test melakukan uji spesifikasi.

Dalam penelitian ini penulis menggunakan pendekatan *fixed effect* dan *random effect* dengan pemilihan model yang menggunakan Hausman *Test*. atau masing-masing tes melakukan uji signifikansi. Dalam penelitian ini penulis menggunakan pendekatan *fixed effect* dan *random effect* dengan pemilihan model yang menggunakan Hausman test. Jika nilai Hausman test, jika nilai Hausman test > dari nilai Chi Square$_{\text{-tabel}}$, maka model yang baik untuk diestimasi adalah *Fixed Effect*, sedangkan jika nilai Hausman test < dari nilai Chi Square$_{\text{-tabel}}$, maka model yang baik untuk diestimasi adalah *Random Effect.*

Berikut ini penjelasan masing-masing model:

1. Pendekatan *Common* Model

Bentuk umum regresi *common* model dapat di tulis sebagai berikut (Widarjono, 2005:77) :

$Y_{it} = \beta_o + \beta_1 X_{1it} + \beta_2 X_{2it} + \dots\dots\dots.\beta_k X_{kit} + e_i \dots\dots\dots\dots\dots\dots\dots\dots\dots$ (4)

Di mana Y adalah variabel dependen, X_1 dan X_1 adalah variabel independen dan e_i adalah residual. Subskrip i menunjukkan observasi ke i untuk data *cross section* dan jika digunakan data *time series*, maka digunakan subskrip t yang menunjukkan waktu. Di dalam persamaan (3.4) tersebut β_o disebut intersep, sedangkan $\beta_1.\beta_2$ disebut koefisien regresi parsial. Adapun asumsi dari regresi *common* model adalah :

1. Hubungan antara Y (variabel dependen) dan X (variabel independen) adalah linier dalam parameter.
2. Nilai X nilainya tetap untuk observasi yang berulang-ulang *(non-stocastic)* dengan asumsi tidak ada hubungan linier antara variabel independen atau tidak ada multikolinearitas.
3. Nilai harapan *(expected vakue)* atau rata-rata dari variabel gangguan e_i adalah nol.

   $E(eIX_i) = 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots......$ (5)

4. Varian dari variabel dari variabel gangguan e_i adalah sama (homoskedastisitas).

   $Var(eiIX_i) = E[(ei\text{-}E(eiIX_i)]^2$

   $= E(ei^2IX_i)$

   $= \sigma^2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots...\dots\dots$ (6)

5. Tidak ada serial korelasi antara residual ei atau residual ei tidak saling berhubungan dengan residual ej yang lain.

   $Cov(ei,ejIX_{i,}\ X_j) = E[(ei\text{-}E(eiIX_i)]\ [(ej0E(ej)Ixj)\ ]$

   $= E(eiIX_i)\ (ejIXj)$

$= 0$……………………………………………......….......…. (7)

6). Variabel gangguan e_i berdistribusi normal.

2. Pendekatan *Fixed Effect* Model

Asumsi dari *fixed effect,* perbedaan antar unit dapat dilihat melalui perbedaan dalam *constant term. Fixed effect* model mengasumsikan bahwa tidak terdapat *time specific effect* dan hanya memfokuskan pada *individual specific effect* (Hsiao, 1995: 29-30). Sehingga nilai dari dependent variabel untuk unitke-I pada waktu y, Yit tergantung pada variabel eksogen K, (X_{Yit} ……., X_{Kit}) $=X_{it}$ yang berbeda antar individu dalam *cross section* pada satu waktu yang menunjukkan variasi selama waktu itu, sebagaimana dalam variabel yang specific pada unit ke-I dan konstan antar waktu. Misal Yi dan Xi adalah sejumlah T observasi dalam unit ke-I, dan menjadi satu dalam vektor gangguan Txi. Sehingga dalam persamaan dapat dituliskan menjadi (Green, 2000: 560-561).

$Yi = i\alpha i + Xi\,\beta + \varepsilon i$……………………..............................……......…. (8)

Kumpulan batasan menjadi :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ \cdot \\ Y_n \end{bmatrix} = \begin{bmatrix} i & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & i & \cdot & \cdot & \cdot & 0 \\ & & \cdot & & & \\ & & \cdot & & & \\ & & \cdot & & & \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_n \end{bmatrix} + \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_n \end{bmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix} \quad \text{atau } Y = \begin{bmatrix} d_1 & d_2 & \cdots & d_n & X \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \varepsilon$$

Dimana d_1 adalah variabel dummy yang mengindikasikan unit ke-i. Jika matriks nTXn D = $[d_1 \; d_2 \; \cdots \; d_n]$. Kemudian, semua baris nT di pasangkan, sehingga menjadi:

$$Y = D\alpha + X\beta + \varepsilon \text{ ……………………………………………....… (9)}$$

Model ini disebut *Least Square Dummy Variabel* (LSDV). Model LSDV adalah model regresi klasik, sehingga tidak ada hasil regresi yang baru dibutuhkan untuk menganalisisnya. Jika n hanya dalam jumlah kecil, maka model ini bisa diestimasi dengan *Ordinary Least Square* (OLS). Dengan K regresor dalam X dan n kolom dalam D, sebagaimana regresi berganda dengan n+k parameter. Jika jumlah n dalam ribuan, maka model seperti diatas akan melebihi kapasitas yang tersedia pada komputer.

3. Pendekatan *Random Effect* Model

Dalam pendekatan *random effect*, bentuk umum regresi data panel :

$$Y_{it} = \beta_{it} + \beta_2 X_{2it} + \beta_3 X_{3it} + ... + \beta_n X_{nit} + \upsilon_{it} \text{ …………………………..…. (10)}$$

Dimana β_{it} tidak lagi dianggap *fixed effect* tetapi dianggap sebagai variabel *random effect* dengan nilai rata-rata β_1. Dengan demikian nilai intersep untuk masing-masing individu 1 dapat dinyatakan sebagai berikut :

$$\beta_{it} = \beta_1 + \varepsilon_{it} \text{ ………………………………………………………………. (11)}$$

Dimana e_{it} adalah galat random effect dengan nilai rata-rata nol dan varians σ_{ε}^2

$$Y_{it} = \beta_{1it} + \beta_2 X_{2it} + ... + \beta_n X_{nit} + \omega_{it} \text{ ………………………………….…… (12)}$$

Dimana :

1. $\omega_{it} = \varepsilon_{it} + \upsilon_{it}$ galat ω_{it} terdiri atas dua komponen, yaitu ε_{it} yang merupakan komponen galat dari individu spesifik atau *cross section*, dan υ_{it} yang merupakan komponen galat w_{it} gabungan dari *time series* dan *cross section.*
2. $E(w_{it}) = 0$
3. var(w_{it}) = $\sigma_{\varepsilon}^2 + \sigma_{it}^2$
4. Dari nomor tiga di atas, galat w_{it} adalah homokedastis. Namun untuk t $\neq$ s, w_{it} dan berkorelasi yaitu galat untuk unit *cross section* pada dua titik waktu yang berbeda korelasi. Sehingga koefisien korelasinya dapat dinyatakan sebagai :

$$\text{Corr} (w_{it,} w_{is}) = \frac{\sigma_{\varepsilon t}^2}{\sigma_{\varepsilon}^2 + \sigma_{it}^2} \text{ ……………………………………. (13)}$$

Ada dua hal mengenai koefisien korelasi, pertama unit *cross section* tertentu, nilai korelasi antara galat pada dua waktu yang berbeda adalah tetap. Kedua, struktur korelasi tetap sama untuk unit *cross section* artinya identik untuk semua individu. Kondisi yang tepat untuk *random effect* adalah jika diyakini bahwa unit *cross section* yang diambil sebagai sampel diambil dari populasi yang besar.

3.4.3. Regresi Linier Metode *Generalized Least Square (GLS)*

Dalam penelitian ini penulis menggunakan metode GLS untuk mengolah data panel yang tersedia. Metode ini dipilih karena adanya nilai lebih yang dimiliki GLS dibanding OLS dalam hal mengestimasi parameter regresi. Gujarati (2009:41) menyebutkan bahwa OLS yang umum tidak mengasumsikan bahwa varian variabel adalah heterogen. Pada kenyataanya variasi data panel cenderung heterogen. Metode GLS lebih memperhitungkan heterogenitas yang terdapat pada variabel independent secara axplisit. Sehingga metode ini mampu menghasilkan estimator yang memenuhi kriteria *BLUE* (*Best Linier Unbiased Estimator*).

3.4.4. Pengujian Statistik

1. Uji t (*t-test*)

Uji t digunakan untuk membuktikan pengaruh variabel independen terhadap variabel dependen secara individual dengan asumsi bahwa variabel yang lain tetap atau konstan. Adapun langkah-langkah dalam uji t untuk pengaruh yang positif dan negatif adalah :

a). Merumuskan hipotesis untuk pengaruh positif

Ho : $\beta i \leq 0$ (Variabel independen tidak berpengaruh secara positif dan signifikan terhadap variabel dependen)

Ha : $\beta i > 0$ (Variabel independen berpengaruh secara positif dan signifikan terhadap variabel dependen)

b). Merumuskan hipotesis untuk pengaruh negatif

Ho : $\beta i \geq 0$ (Variabel independen tidak berpengaruh secara negatif dan signifikan terhadap variabel dependen)

Ha : $\beta i < 0$ (Variabel independen berpengaruh secara negatif dan signifikan terhadap variabel dependen)

c). Menentukan kriteria pengujian pengaruh positif

Penelitian ini menggunakan uji satu sisi kanan, maka daerah penolakannya berada di sisi kanan kurva yang luasnya α dan derajat kebebasan *(degrre of* *freedom)* yaitu : df = n-k, dimana n adalah jumlah sampel dan k adalah konstanta.

- Bila $t_{\text{-statistik}} \leq t_{\text{-tabel}}$, maka Ho diterima, artinya tidak ada pengaruh secara positif dan signifikan antara variabel independen terhadap variabel dependen.

- Bila $t_{\text{-statistk}} > t_{\text{-tabel}}$, maka Ho ditolak, artinya ada pengaruh secara positif dan signifikan antara variabel independen terhadap variabel dependen.

d). Menentukan kriteria pengujian pengaruh negatif

Penelitian ini menggunakan uji satu sisi kiri, maka daerah penolakannya berada di sisi kanan kurva yang luasnya α dan derajat kebebasan *(degrre of freedom)* yaitu : df = n-k, dimana n adalah jumlah sampel dan k adalah konstanta.

- Bila $-t_{\text{-statistik}} \geq -t_{\text{-tabel}}$, maka Ho diterima, artinya tidak ada pengaruh secara negatif dan signifikan antara variabel independen terhadap variabel dependen.

- Bila $t_{\text{-statistk}} > t_{\text{-tabel}}$, maka Ho ditolak, artinya ada pengaruh secara negatif dan signifikan antara variabel independen terhadap variabel dependen.

e). Mencari nilai $t_{\text{-statistik}}$ (Gujarati, 2009: 74)

$$t_{\text{-hitung}} = \frac{\beta i}{Se\ \beta i}$$

Keterangan :

t = Nilai $t_{\text{-statistik}}$

βi = Koefisien regresi

Se βi = Standard error βi

2. Uji F (*F-test*)

Uji F adalah uji serempak yang digunakan untuk mengetahui pengaruh variabel independen secara serempak terhadap variabel dependen.

Langkah-langkah :

a). Merumuskan hipotesis :

Ho : $\beta_1 = \beta_2 = \beta_3 = 0$ (Tidak ada pengaruh yang signifikan antara variabel independen terhadap variabel dependen secara simultan).

Ha : $\beta_1 \neq \beta_2 \neq \beta_3 \neq 0$ (Ada pengaruh yang signifikan antara variabel independen terhadap variabel dependen secara simultan).

b). Menentukan kriteria pengujian dengan *level of significant* (α) 5 %, dan df pembilang k-1 dan penyebut n-k.

-Bila $F_{-statistik} > F_{-tabel}$, maka Ho ditolak, artinya secara simultan variabel independen berpengaruh terhadap variabel dependen.

-Bila $F_{-statistik} \leq F_{-tabel}$, maka Ho diterima, artinya secara simultan variabel independen tidak berpengaruh terhadap variabel dependen.

c). Mencari $F_{-statistik}$ (Gujarati, 2009: 141) :

$$F_{-hitung} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

Keterangan :

R^2 = Koefisien Determinasi

K = Jumlah Variabel Independen

n = Jumlah Observasi

3. R^2 (Koefisien Determinasi

R^2 (Koefisien Determinasi) untuk mengetahui seberapa besar kemampuan variabel independen dalam menjelaskan variabel dependen. Nilai R^2 (Koefisien Determinasi) mempunyai *range* antara 0-1. Semakin besar R^2 mengindikasikan semakin besar kemampuan variabel independen dalam menjelaskan variabel independen. Perumusan yang digunakan untuk mencari nilai R^2 adalah (Gujarati, 2009: 139) :

$$R^2 = \frac{\left\{N\sum XiYi - \left(\sum Xi\right)\left(\sum Yi\right)\right\}^2}{\left\{N\sum Xi^2 - \left(\sum Xi\right)^2\right\}\left\{N\sum Yi^2 - \left(\sum Yi\right)^2\right\}}$$

Keterangan :

R^2 = Koefisien determinasi

X_i = Variabel independen

Y_i = Variabel dependen

N = Observasi