

## **BAB III**

### **METODE PENELITIAN**

#### **A. Obyek/Subyek Penelitian**

Variabel penelitian dan definisi operasional atau konsep yang dapat diukur dengan berbagai macam nilai untuk memberikan gambaran yang nyata mengenai peristiwa yang diteliti. Dalam penelitian ini menggunakan dua variabel yaitu variabel dependen dan variabel independen. Variabel dependen yang digunakan dalam penelitian ini yaitu jumlah penduduk miskin di Indonesia pada tahun 2011-2016. Variabel independen yang digunakan dalam penelitian ini yaitu Pengeluaran Pemerintah Sektor Pendidikan, Produk Domestik Regional Bruto, dan Angkatan Kerja di 30 Provinsi Indonesia pada tahun 2011-2016 yang diperoleh dari Badan Pusat Statistik, BAPPENAS dan Departemen Jenderal Perimbangan Keuangan.

#### **B. Jenis dan Sumber Data**

Jenis data yang digunakan dalam penelitian ini adalah data panel dalam bentuk data sekunder dan tergolong dalam jenis data kuantitatif. Data sekunder itu sendiri berarti data yang tidak dikumpulkan sendiri oleh peneliti melainkan data yang diperoleh dari buku, arsip yang telah ada dan tidak dipublikasikan secara umum. Dengan kata lain, peneliti harus berkunjung ke perpustakaan, pusat kajian, atau pusat arsip tersebut. Adapun yang dimaksud data panel adalah “gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*)”. Menurut Widarjono dalam Basuki A.T dan Yuliadi I (2015) penggunaan data panel dalam

sebuah observasi mempunyai beberapa keuntungan yang diperoleh, yaitu data panel dapat menyediakan data yang lebih banyak sehingga akan lebih menghasilkan *degree of freedom* yang lebih besar dan data panel juga dapat mengatasi masalah yang timbul ketika ada masalah penghilangan variabel. Data sekunder ini bersumber dari Badan Pusat Statistik, BAPPENAS, dan Departemen Jenderal Perimbangan Keuangan, data yang peneliti pakai adalah data Jumlah Penduduk Miskin, Pengeluaran Pemerintah Sektor Pendidikan, Produk Domestik Regional Bruto, Angkatan Kerja di Indonesia menurut provinsi 30 Periode 2011-2016.

### **C. Teknik Pengumpulan Data**

Teknik pengumpulan data yang digunakan adalah dengan studi dokumentasi. Studi dokumentasi adalah pengumpulan data yang dilakukan dengan kategori dan klasifikasi bahan-bahan tertulis yang berhubungan dengan masalah penelitian, baik itu datang secara langsung maupun dengan mengunjungi *website*-nya.

### **D. Definisi Operasional Variabel Penelitian**

#### 1. Variabel Dependen

##### a. Kemiskinan

Dalam penulisan ini yang menjadi variabel terikat atau variabel yang di pengaruhi oleh variabel lainnya. Menurut BAPPENAS (2004) kemiskinan sebagai keadaan dimana kekurangan yang terjadi bukan karena kehendak karena oleh orang miskin, tetapi karena keadaan yang tidak bisa dihindari oleh kekuatan yang apa adanya.

#### 2. Variabel Independen

a. Pengeluaran Pemerintah Sektor Pendidikan

Pengeluaran pemerintah merupakan bentuk investasi langsung untuk memberikan kesejahteraan bagi masyarakat. Berdasarkan Perpres No. 13 Tahun 2009 dilanjutkan dengan No.15 Tahun 2010, Pengeluaran Pemerintah Sektor Pendidikan termasuk Kluster I yaitu fokus dalam pemenuhan hak dasar ditujukan untuk memperbaiki kualitas kehidupan masyarakat miskin untuk kehidupan yang lebih baik, seperti pemenuhan hak atas pangan, pelayanan kesehatan, dan pendidikan.

b. Produk Domestik Bruto Regional (PDRB)

PDRB di sini menggunakan PDRB atas harga konstan. PDRB atas harga konstan 2010 menurut Badan Pusat Statistik (2016) adalah jumlah nilai produksi atau pendapatan atau pengeluaran yang dinilai atas harga dasar tetap (harga pada tahun dasar) yang digunakan dalam satu tahun. PDRB atas harga konstan dapat digunakan untuk menunjukkan laju pertumbuhan ekonomi secara keseluruhan atau setiap sektor dari tahun ke tahun.

c. Angkatan Kerja

Menurut (Suparmoko, 2002), angkatan kerja adalah penduduk yang belum bekerja namun siap untuk bekerja atau sedang mencari pekerjaan pada tingkat upah yang berlaku. Angkatan kerja terdiri atas golongan yang bekerja, dan golongan yang menganggur dan mencari pekerjaan (Payaman, 1985).

## **E. Uji Kualitas Instrumen dan Data**

1. Uji asumsi klasik.

Menurut Basuki dan Yuliadi (2015) dalam regresi data panel, uji asumsi klasik yang digunakan hanya uji *multikolinearitas* dan uji *heterokedastisitas* saja.

a. Uji *multikolinearitas*.

Menurut Basuki dan Yuliadi (2015) uji *multikolinearitas* adalah uji yang dilakukan untuk memastikan apakah terdapat interkorelasi atau kolinearitas antar variabel bebas di dalam suatu analisis regresi. Menurut Ghazali (2006) uji *Multikolonieritas* bertujuan untuk menguji apakah model regresi ditemukan adanya korelasi antar variabel bebas (*Independent*). Model korelasi yang baik seharusnya tidak terjadi korelasi di antara variabel *independent*. Jika variabel *independent* saling berkorelasi, maka variabel ini tidak ortogonal. Variabel ortogonal adalah variabel *independent* yang nilai korelasi antar sesama variabel *independent* sama dengan nol. Basuki dan Yuliadi (2015) uji *multikolinearitas* dengan model regresi sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + E$$

Dimana:

Y = variabel dependen

$X_1, X_2, X_3$  = variabel Pengeluaran Pemerintah Sektor Pendidikan, PDRB, dan Angkatan Kerja

b. Uji *heterokedastisitas*.

Menurut Basuki dan Yuliadi (2015) uji *Heteroskedastisitas* adalah situasi tidak konstannya varians, uji yang menilai apakah ada ketidaksamaan varian dari residual untuk semua pengamatan pada model regresi *linear*. Konsekuensi *Heteroskedastisitas* adalah biasanya varians sehingga uji signifikansi menjadi invalid. Uji ini merupakan salah satu dari uji asumsi klasik yang harus dilakukan pada regresi *linear*. Apabila asumsi *heteroskedastisitas* tidak terpenuhi, maka model regresi dinyatakan tidak valid sebagai alat peramalan. Menurut Ghazali (2006) uji *heteroskedastisitas* bertujuan untuk menguji apakah dalam model regresi terjadi ketidaksamaan *variance* dari residual satu pengamatan ke pengamatan lainnya. Penelitian ini menggunakan Uji Glejser untuk meregres nilai *absolut residual* terhadap variabel *independent* (Gujarati, 2006) dengan menggunakan dasar pengambilan keputusan, jika nilai Sig. variabel independen  $< 0,05$  terjadi *Heterokedastisitas* sedangkan jika nilai Sig. variabel independen  $> 0,05$  tidak terjadi *Heterokedastisitas*.

## **F. Uji Hipotesis dan Analisis Data**

### **1. Statistik Deskriptif**

Analisis deskriptif statistik yaitu analisis yang dilakukan untuk menunjukkan gambaran statistik variabel dependen dan independen yang digunakan dalam penelitian seperti, nilai maksimum, nilai minimum, nilai rata-rata dan standar deviasi. Pengukuran analisis deskriptif dilakukan dengan menggunakan alat analisis *Ms.Excel*.

#### **a. Mean (Rata-rata)**

Dalam output Excel pada menu *descriptive statistics*, nilai rata-rata yang ditampilkan adalah rata-rata hitung (*arithmetic mean*). Rata-rata hitung ini adalah pengukuran nilai sentral yang paling umum digunakan. Dalam keseharian kita biasanya mengenal hanya dengan istilah rata-rata. Rumus untuk menentukan nilai rata-rata hitung:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

Dimana:

$\bar{X}$  = rata-rata hitung

$X_1, X_2, X_3$  = jumlah nilai data dari data pertama sampai data n

$n$  = jumlah data

#### b. *Median*

Median merupakan ukuran yang kuat (*robust*) dari nilai sentral. Hal ini dikarenakan nilai median tidak dipengaruhi oleh nilai ekstrim. Median adalah nilai yang berada di tengah-tengah dari sekumpulan data, jika data tersebut diurutkan baik dari nilai terkecil ke nilai terbesar maupun dari nilai terbesar ke nilai terkecil. Secara rumus, median terletak pada urutan ke:

$$\frac{n + 1}{2}$$

dimana n adalah banyaknya data

c. *Maximum* adalah nilai tertinggi dari data

d. *Minimum* adalah nilai terendah dari data

e. Standar Deviasi

Standar deviasi merupakan akar dari varians (karena pada varians kita mengkuadratkan selisih data dari rata-ratanya, maka dengan mengakarkannya, kita mendapatkan kembali nilai asalnya).

## 2. Uji Hipotesis

Dalam penelitian ini, metode *Panel Least Square* digunakan untuk melakukan uji hipotesis. Menurut Basuki dan Yuliadi (2015) terdapat 3 pendekatan dalam metode analisis ini, yaitu:

### a. *Common effect model/pooled model*.

*Common effect model* adalah pendekatan regresi data panel sederhana, di mana pada pendekatan ini model hanya mengkombinasikan data *cross section* dan data *time series*. Pendekatan ini mengasumsikan perilaku data perusahaan sama dalam berbagai kurun waktu, selain itu model ini juga tidak memperhatikan dimensi waktu maupun individu. Analisis regresi dengan pendekatan ini dapat dilakukan dengan model OLS dalam mengestimasi data panel. Menurut Basuki dan Yuliadi (2015) adapun persamaan regresi dalam model *common effects* dapat ditulis sebagai berikut:

$$Y_{it} = \alpha + X_{it}\beta + \varepsilon_{it}$$

Dimana:

i = Aceh, Sumatra Utara, ..., Papua

t = 2011, 2012, 2013, 2014, 2015, 2016

dimana  $i$  menunjukkan *cross section* (individu) dan  $t$  menunjukkan periode waktunya. Dengan asumsi komponen *error* dalam pengolahan kuadrat terkecil biasa, proses estimasi secara terpisah untuk setiap unit *cross section* dapat dilakukan.

b. *Fixed effect model*.

Pendekatan *fixed model* mengasumsikan perbedaan intersep dapat mengakomodasi perbedaan antar individu. Estimasi dalam model *fixed effect* ini dilakukan dengan teknik variabel *dummy*, teknik variabel *dummy* digunakan untuk menangkap perbedaan intersep antar perusahaan. Menurut Basuki dan Yuliadi (2015) dalam model *fixed effects*, setiap merupakan parameter yang tidak diketahui dan akan diestimasi dengan menggunakan teknik variabel *dummy* yang dapat ditulis sebagai berikut:

$$Y_{it} = \alpha + i\alpha_{it} + X'_{it}\beta + \varepsilon_{it}$$

$$\begin{bmatrix} y_1 \\ y_1 \\ y_n \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \\ \alpha \end{bmatrix} + \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} x_{11} & x_{21} & x_{p1} \\ x_{12} & x_{22} & x_{p2} \\ x_{1n} & x_{2n} & x_{pn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_n \end{bmatrix}$$

Model estimasi ini disebut dengan teknik *Least Square Dummy Variable* (LSDV). Selain diterapkan untuk efek tiap individu, LSDV ini juga dapat mengakomodasi efek waktu yang bersifat sistemik.

c. *Random effect model*.

Pendekatan *random effect* melakukan estimasi data panel ketika variabel gangguan mungkin akan saling berhubungan antar waktu antar



individu. Berbeda dengan *fixed effect model*, efek spesifik dari masing-masing individu dianggap sebagai bagian dari komponen error yang bersifat acak dan tidak berkorelasi dengan variabel penjelas yang teramati. Pendekatan dengan model *random effect* dinamakan ECM (*Error Component Model*) atau teknik GLS (*Generalized Least square*). Menurut Basuki dan Yuliadi (2015) dengan demikian, persamaan model *random effects* dapat dituliskan sebagai berikut:

$$Y_{it} = \alpha + X_{it}^i \beta + w_{it}$$

$i$  = Aceh, Sumatra Utara, ..., Papua

$t$  = 2011, 2012, 2013, 2014, 2015, 2016

Dimana:

$$w_{it} = \varepsilon_{it} + u_i; E(w_{it}) = 0; (w_{it}^2) = \alpha^2 + \alpha_u^2;$$

$$w_{it}'w_{jt-1}) = 0; i \neq j; E(u_i' \varepsilon_{it}) = 0;$$

$$E(\varepsilon_i' \varepsilon_{is}) = E(\varepsilon_{it}' \varepsilon_{js}) = E(\varepsilon_{it}' \varepsilon_{js}) = 0$$

Walaupun komponen error  $w_t$  bersifat homoskedastik, nyatanya terdapat korelasi antara  $w_t$  dan  $w_{t-1}$  (*equicorrelation*), yakni:

$$\text{Corr}(w_{it}'w_{i(t-1)}) = \alpha_u^2 / (\alpha^2 + \alpha_u^2)$$

Karena itu, metode OLS tidak bisa digunakan untuk mendapatkan estimator yang efisien bagi model *random effects*. Metode yang tepat untuk

mengestimasi model *random effects* adalah GLS (*Generalized Least Squares*) dengan asumsi homokedastik dan tidak ada *cross sectional correlation*.

### 3. Pemilihan model estimasi analisis data.

Menurut Basuki dan Yuliadi (2015) untuk memilih model yang paling tepat dalam mengolah data panel, maka perlu dilakukan beberapa pengujian melalui:

#### a. *Chow test*.

*Chow Test* adalah uji yang dilakukan untuk menentukan ketepatan model estimasi antara *pooled* atau *fixed* di mana:

$H_0$ : *Pooled Model*

$H_1$ : *Fixed Effect*

Jika nilai probabilitas *Chow Test* lebih kecil dari 0,05 (5%) maka  $H_0$  ditolak yang berarti model yang paling tepat untuk digunakan adalah *Fixed Effect Model*. Sebaliknya jika nilai probabilitas *Chow Test* lebih besar dari 0,05 (5%) maka model yang cocok dalam mengestimasi data panel adalah *Pooled Model*.

#### b. Koefisien Determinasi ( $R^2$ )

Koefisien determinasi ( $R^2$ ) menunjukkan besar distribusi dari variabel independen terhadap variabel dependen. Koefisien determinasi yaitu proporsi variasi dalam Y yang dijelaskan oleh pengaruh linier X. Untuk memperoleh  $R^2$  dapat diformulasikan sebagai berikut :

$$R^2 = \frac{ESS}{TSS}$$

Dimana: ESS = Jumlah kuadrat error

TSS = Jumlah kuadrat total

$R^2$  berada diantara 0 dan 1. Bila nilai koefisien determinasi ( $R^2$ ) mendekati angka 1, berarti variabel independen semakin mendekati hubungan dengan variabel dependen, sehingga pengaruh model tersebut dapat dikatakan benar. Jadi, semakin besar nilai dari  $R^2$ , maka semakin bagus garis regresi yang terbentuk dan semakin kecil nilai dari  $R^2$ , maka semakin tidak tepat garis regresi tersebut mewakili data hasil observasi. Hal ini diperlukan karena dalam suatu penelitian atau observasi, perlu dilihat seberapa jauh model yang terbentuk dapat menerangkan dependen variabelnya.

c. *Hausman test*.

*Hausman test* ialah uji yang dilakukan untuk menentukan ketepatan model estimasi antara *Fixed* atau *Random*, di mana:

$H_0$ : *Fixed Effect Model*

$H_1$ : *Random Effect Model*

Jika nilai probabilitas *hausman test* lebih besar dari 0,05 maka  $H_0$  diterima, yang artinya model yang cocok dalam menganalisis data panel adalah *fixed effect model*. Namun jika hasil probabilitas *hausman test* lebih kecil dari 0,05 maka model yang tepat untuk mengestimasi data panel adalah *random effect model*.

d. *Uji langrange multiplier*.

Uji ini dilakukan agar dapat mengetahui ketepatan model estimasi antara *pooled* dan *random*.

#### 4. Model Analisis Regresi Data Panel.

Adapun model regresi data panel yang digunakan dalam penelitian ini adalah dengan model regresi data panel. Analisis regresi merupakan studi mengenai ketergantungan satu variabel lainnya atau variabel independen. Data Panel adalah gabungan antara data runtut waktu (*time series*) dan data silang (*cross section*) dimana *cross section* yang sama diukur pada waktu yang berbeda. Dengan kata lain, data panel merupakan data dari beberapa individu sama yang diamati dalam kurun waktu tertentu.

Analisis regresi data panel adalah analisis regresi yang didasarkan pada data panel untuk mengamati hubungan antara satu variabel terikat dengan satu atau lebih variabel bebas. Analisis regresi dengan menggunakan data panel memungkinkan peneliti mengetahui karakteristik antar waktu dan antar individu dalam variabel yang bisa saja berbeda-beda. Adapun model rumus regresi data panel adalah sebagai berikut (Basuki dan Yuliadi, 2015)

$$Y_{it} = \alpha + b_1X_{1it} + b_2X_{2it} + b_3X_{3it} + e$$

$$K_{it} = \alpha + b_1PDRB_{it} + b_2PPSP_{it} + b_3AK_{it} + e$$

Keterangan:

Y = Variabel Dependen (Kemiskinan)

$\alpha$  = Konstanta

$X_1$ PDRB = Variabel Independen 1 (PDRB)

$X_2$ PPSP = Variabel Independen 2 (Pengeluaran Pemerintah Sektor Pendidikan)

$X_3AK$  = Variabel Independen 3 (Angkatan Kerja)

$b(1,2,3)$  = Koefisien regresi masing-masing variabel independen

$i$  = Aceh,.....,Papua

$t$  = 2011,....,2016

$e$  = Error term