

## **BAB III**

### **METODOLOGI PENELITIAN**

#### **1. Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk mendapatkan pengetahuan yang tepat, valid, dan dapat dipercaya (dapat diandalkan atau reliable), tentang:

1. Pengaruh luas panen terhadap produksi beras di Indonesia pada tahun 2007-2011
2. Pengaruh tenaga kerja terhadap produksi beras di Indonesia pada tahun 2007-2011
3. Pengaruh luas panen dan tenaga kerja terhadap produksi beras di Indonesia pada tahun 2007-2011

#### **2. Jenis dan Sumber Data**

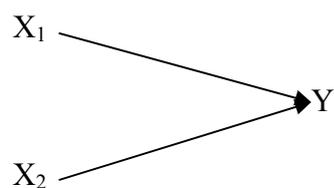
Penelitian ini menggunakan data panel yang merupakan penggabungan antara data kerat lintang sekunder (*cross section*) 27 propinsi yang ada di Indonesia dengan data deret waktu (*time series*) yang berbentuk periode tahunan dari tahun 2007 hingga tahun 2011. Hal ini dilakukan untuk melihat produksi beras pada waktu tertentu (2007-2011) di setiap provinsi. Data yang dibutuhkan adalah data produksi beras, luas panen padi, dan tenaga kerja di sub sektor tanaman padi. Data yang digunakan diperoleh dari Badan Pusat Statistik.

### 3. Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode *ekspos facto* dengan pendekatan korelasional. Metode ini dipilih karena merupakan metode yang sistematis dan empirik. Metode *Ex Post Facto* adalah suatu penelitian yang dilakukan untuk meneliti peristiwa yang telah terjadi dan kemudian meruntut ke belakang untuk mengetahui faktor-faktor yang dapat menimbulkan kejadian tersebut.<sup>33</sup> Sehingga dengan pendekatan korelasional ini, akan dapat dilihat pengaruh antara tiga variabel yaitu, variabel bebas (luas panen dan tenaga kerja), yang mempengaruhi dan variabel terikat (produksi beras), yang dipengaruhi.

### 4. Konstelasi Hubungan Antar Variabel/Desain Penelitian

Konstelasi pengaruh antar variabel dalam penelitian ini bertujuan untuk memberikan arah atau gambaran dalam penelitian ini, yang dapat digambarkan sebagai berikut:



Keterangan:

- X1 : luas panen
- X2 : tenaga kerja
- Y : produksi beras
- : arah pengaruh

---

<sup>33</sup> Sugiyono, *Metode Penelitian Bisnis*, (Jakarta: Alfabeta, 2004), hal. 7

## **5. Operasionalisasi Variabel Penelitian**

### **1. Produksi Beras**

#### **a. Definisi Konseptual**

Produksi beras adalah jumlah produk (beras) yang dihasilkan dari proses pengolahan lahan pertanian yang dipengaruhi oleh berbagai faktor produksi lain.

#### **b. Definisi Operasional**

Data produksi beras dalam penelitian ini diukur dengan menggunakan data produksi padi yang dikalikan dengan angka rendemen penggilingan lapangan yaitu sebesar 63,2%. Angka rendemen merupakan angka yang dirilis oleh BPS dan Direktorat Pengolahan dan Pemasaran Hasil Pertanian, kementerian Pertanian. Produksi beras diperoleh dengan cara:

$$\text{Produksi Beras} = \text{Produksi padi} \times 63,2\%$$

Data produksi beras yang digunakan berdasarkan data masing-masing provinsi di Indonesia pada tahun 2007-2011.

### **2. Luas Panen**

#### **a. Definisi Konseptual**

Luas panen adalah jumlah seluruh lahan yang dapat diambil hasil tanamannya (padi) setelah tanaman tersebut cukup umur.

#### **b. Definisi Operasional**

Data luas panen merupakan luas tanaman padi yang berhasil di panen dari tahun 2007-2011. Data ini dipublikasikan oleh Badan Pusat

Statistik dan berdasarkan data masing-masing provinsi di Indonesia yang dinyatakan dalam satuan Ha.

### **3. Tenaga Kerja**

#### **a. Definisi Konseptual**

Tenaga kerja merupakan sejumlah penduduk yang mampu melakukan pekerjaan menghasilkan barang/jasa untuk memenuhi kebutuhan sendiri maupun masyarakat.

#### **b. Definisi Operasional**

Data tenaga kerja merupakan jumlah tenaga kerja produktif yang bekerja di sub sektor tanaman padi yang tersedia di masing-masing provinsi di Indonesia, dinyatakan dalam orang/tahun pada tahun 2007-2011 yang dipublikasikan oleh BPS.

### **6. Teknik Analisis Data**

#### **1. Data Panel**

Dalam penelitian ini pendekatan yang digunakan untuk mengestimasi parameter model adalah pendekatan data panel. Data panel atau *pooled data* merupakan kombinasi dari data time series dan cross-section. Data panel secara substansial mampu menurunkan masalah *omitted-variables*, model yang mengabaikan variabel yang relevan (Wibisono, 2005).<sup>34</sup> Untuk mengatasi interkorelasi di antara variabel-variabel bebas yang pada akhirnya dapat mengakibatkan tidak tepatnya penaksiran regresi, metode

---

<sup>34</sup> Shochrul Rohmatul Ajija, Cara Cerdas Menguasai Eviews (Jakarta: Salemba Empat, 2011), hal. 52

data panel lebih tepat untuk digunakan (Griffiths, 2001: 351). Gujarati<sup>35</sup> menyebutkan bahwa data panel mempunyai kelebihan berikut:

1. Estimasi data panel dapat menunjukkan adanya heterogenitas dalam tiap unit.
2. Data panel dapat memberikan data yang lebih informatif, lebih bervariasi, mengurangi kolinearitas antara variabel, derajat kebebasan yang lebih besar dan lebih efisien.
3. Data panel lebih sesuai untuk mempelajari dinamika perubahan.
4. Data panel dapat lebih baik mendeteksi dan mengukur efek yang tidak dapat diamati dalam data *cross section* dan data *time series*.
5. Data panel dapat digunakan untuk mempelajari model perilaku yang kompleks.
6. Data panel dapat meminimalisir bias yang mungkin ditimbulkan oleh regresi.

Karena data panel merupakan gabungan dari data *cross section* dan data *time series*, maka Nachrowi dan Usman<sup>36</sup> menuliskan modelnya sebagai berikut:

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

$$i = 1, 2, \dots, N \text{ dan } t = 1, 2, \dots, T$$

---

<sup>35</sup> Damodar N. Gujarati, *Basic Econometrics 4<sup>th</sup>* (New York: Mc Graw Hill, 2004)

<sup>36</sup> Nachrowi dkk, *Pendekatan Populer dan Praktis Ekonometrika untuk Analisis ekonomi dan Keuangan* (Jakarta: Lembaga Penerbit FE UI, 2006)

Dimana:

$N$  = banyak observasi

$T$  = banyaknya waktu

$N \times T$  = banyaknya data panel

Estimasi fungsi produksi dengan linearisasi fungsi produksi Cobb

Douglas dalam data panel dapat dinyatakan dengan persamaan:

$$\ln \text{produksi} = \ln \alpha + \beta_1 \ln \text{luas panen} + \beta_2 \ln \text{tenaga kerja} + \varepsilon$$

Terdapat tiga pendekatan dalam mengestimasi parameter model dengan data panel, yaitu:

1. Pendekatan kuadrat terkecil (*Ordinary Least Square*)

Teknik ini mengkombinasikan atau mengumpulkan semua data *cross section* dan *time series*, lalu mengestimasi model tersebut dengan metode *Ordinary Least Squared* (OLS). Asumsi yang digunakan dalam metode OLS adalah data gabungan tersebut menunjukkan kondisi yang sebenarnya sehingga hasil analisis regresi dianggap berlaku pada semua objek pada semua waktu atau dengan kata lain pengaruh spesifik dari masing-masing individu diabaikan atau dianggap tidak ada. Dengan demikian, akan dihasilkan sebuah persamaan regresi yang sama untuk setiap unit *cross section*. Sesuatu yang secara realistis tentunya kurang dapat diterima. Oleh karena itu, model ini sangat jarang digunakan dalam analisis data panel. Persamaan regresi untuk model *common effect* dapat dituliskan sebagai berikut:

$$\ln \text{produksi beras}_{it} = \alpha + \beta_1 \ln \text{luas panen}_{it} + \beta_2 \ln \text{tenaga kerja}_{it} + \varepsilon_{it}$$

## 2. Pendekatan efek tetap (*Fixed Effect*)

Dalam pendekatan efek tetap, satu objek mempunyai konstan yang tetap besarnya untuk berbagai periode waktu. Koefisien regresinya juga tetap besarnya dari waktu ke waktu (*time variant*). Untuk membedakan satu objek dengan objek lainnya digunakan variabel boneka (*dummy*). Secara matematis, model efek tetap dapat dituliskan sebagai berikut:

$$\ln \text{produksi beras}_{it} = \alpha + \beta_1 \ln \text{luas panen}_{it} + \beta_2 \ln \text{tenaga kerja}_{it} + \gamma_2$$

$$W_{2i} + \gamma_3 W_{3i} + \dots + \gamma_N W_{Ni} + \sigma_2 Z_{i2} + \sigma_3 Z_{i3} + \dots + \sigma_i Z_{iT} + \varepsilon_{it}$$

$$W_{Ni} + \sigma_2 Z_{i2} + \sigma_3 Z_{i3} + \dots + \sigma_i Z_{iT} + \varepsilon_{it}$$

Dimana:

$W_{it}$  dan  $Z_{it}$  variabel dummy yang didefinisikan sebagai berikut:

$$W_{it} = 1; \text{ untuk provinsi } i; i = 1, 2, \dots, N$$

$$= 0; \text{ lainnya}$$

$$Z_{it} = 1; \text{ untuk tahun } t; t = 1, 2, \dots, T$$

$$= 0; \text{ lainnya}$$

Dari model di atas, terlihat bahwa sesungguhnya model efek tetap adalah sama dengan regresi yang menggunakan *dummy variable* sebagai variabel bebas, sehingga dapat diestimasi dengan OLS. Dengan diestimasi model tersebut dengan OLS, maka akan diperoleh estimator yang tidak bias dan konsisten.

### 3. Pendekatan efek random (*Random Effect*)

Dalam pendekatan efek random, perbedaan karakteristik objek dan waktu diakomodasikan pada residual dari model. Mengingat terdapat dua komponen yang berkontribusi pada pembentukan residual, yaitu objek dan waktu, maka residual pada teknik ini juga diurai untuk komponen objek, komponen waktu dan gabungan. Winarno menyebutkan bahwa teknik ini digunakan untuk mengakomodir kelemahan pendekatan efek tetap yang menggunakan variabel *dummy* sehingga model mengalami ketidakpastian. Tanpa menggunakan variabel semu, metode efek random menggunakan residual, yang diduga memiliki hubungan antarwaktu dan antarobjek.<sup>37</sup> Persamaan yang digunakan dalam model efek random, yaitu:

$$\ln \text{produksi beras}_{it} = \alpha + \beta_1 \ln \text{luas panen}_{it} + \beta_2 \ln \text{tenaga kerja}_{it} + \epsilon_i$$

Tidak seperti pada model efek tetap ( $\alpha$  dianggap tetap), pada model ini  $\alpha$  diasumsikan bersifat random, sehingga dapat dituliskan dalam persamaan:

$$\alpha = \alpha + u_i, i = 1, \dots, n$$

Untuk memilih salah satu model estimasi yang dianggap paling tepat dari tiga jenis model data panel, maka perlu dilakukan serangkaian uji (tes), yaitu uji F (*Chow Test*), uji LM (*Lagrange Multiplier*), dan uji Hausman.

---

<sup>37</sup> Wing Wahyu Winarno, *Analisis Ekonometrika dan Statistika dengan Eviews*, ed. 3 (Yogyakarta: STIM YKPN Yogyakarta, 2011), hal. 9.17

### 1. Pengujian Signifikansi *Common Effects* atau *Fixed Effects*

Untuk mengetahui model *Pooled Least Square/Common Effect* atau *Fixed Effect Model* yang akan dipilih untuk estimasi data yang dapat dilakukan dengan uji F-test atau uji *Chow Test*. PLS adalah *restricted model* dimana ia menerapkan intercept yang sama untuk seluruh individu. Dalam pengujian ini dilakukan dengan hipotesa sebagai berikut:

$H_0$ : Model PLS (*Restricted*)

$H_1$ : Model *Fixed Effect (Unrestricted)*

Dasar penolakan terhadap hipotesis nol tersebut diputuskan melalui hasil uji dengan menggunakan F Statistik seperti yang dirumuskan oleh Chow sebagai berikut:

$$F = \frac{(RRSS - URSS) / (N - 1)}{URSS / (NT - N - K)}$$

Dimana:

RRSS = *Restricted Residual Sum Square* (merupakan *Sum of Square Residual* yang diperoleh dari estimasi data panel dengan metode *pooled least square/common intercept*)

URSS = *Unrestricted Residual Sum Square* (merupakan *Sum of Square Residual* yang diperoleh dari estimasi data panel dengan metode *fixed effect*)

N = jumlah data *cross section*

T = jumlah data *time series*

K = jumlah variabel penjelas

Pengujian ini mengikuti distribusi F Statistic, yaitu  $F_{n-1, nT-n-k}$ . Jika nilai CHOW Statistic (F Stat) hasil pengujian lebih besar dari F Tabel, maka terdapat cukup bukti untuk melakukan penolakan terhadap hipotesis nol, sehingga model yang digunakan adalah model *fixed effect*, begitu juga sebaliknya.

## 2. Pengujian Signifikansi *Common Effects* atau *Random Effects*

Untuk mengetahui apakah menggunakan model *random effect* atau *common effect* dapat dilakukan dengan uji LM. Breush dan Pagan (1980) telah mengembangkan pengujian *Lagrange Multiplier* untuk mengetahui signifikansi dari *random effects* berdasarkan residual dari OLS (*common effects*).

Secara matematis, statistik uji untuk *LM test (Lagrange Multiplier)* dapat dituliskan sebagai berikut:

$$LM = \frac{nT}{2(T-1)} \left[ \frac{\sum_{i=1}^n (\sum_{t=1}^T e_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2 = \frac{nT}{2(T-1)} \left[ \frac{\sum_{i=1}^n |T \bar{e}_i|^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right]^2$$

Di bawah hipotesis nul, LM mengikuti sebaran *chi-square* dengan derajat bebas satu. Jika hasil LM statistik lebih besar dari nilai kritis statistik *chi-square*, maka hipotesis nul akan ditolak, yang berarti estimasi yang tepat untuk regresi data panel adalah metode *random effect* dibandingkan metode *common effects*.

### 3. Pengujian Signifikansi *Fixed Effects* atau *Random Effects*

Untuk mengetahui apakah menggunakan model *fixed effect* atau *random effect* dapat dilakukan dengan uji Hausman. Uji ini dilakukan dengan membandingkan untuk *subset* dari koefisien variabel-variabel yang bervariasi antar unit waktu (*time-varying variables*). Secara matematis dengan menggunakan notasi matriks, statistik uji Hausman ( $H$ ) dapat dituliskan sebagai berikut:

$$H = \left( \hat{\beta}_{FE} - \hat{\beta}_{RE} \right)' \left[ \text{var} \left( \hat{\beta}_{FE} \right) - \text{var} \left( \hat{\beta}_{RE} \right) \right]^{-1} \left( \hat{\beta}_{FE} - \hat{\beta}_{RE} \right)$$

di bawah hipotesis nul, statistik uji ini mengikuti sebaran *chi-square* dengan derajat bebas  $M$ , di mana  $M$  adalah jumlah variabel penjelas yang nilainya bervariasi antar unit waktu di dalam model.

Hipotesis null pada uji Hausman adalah efek spesifik individu tidak berkorelasi dengan peregresi atau dengan kata lain model *random effect* lebih baik bila dibandingkan dengan model *fixed effect*. Di bawah hipotesis nul, pendugaan parameter dengan menggunakan *random effect* adalah konsisten dan efisien, sedangkan pendugaan dengan *fixed effect* meskipun tetap konsisten, tetapi tidak lagi efisien. Di bawah hipotesis alternatif, estimasi dengan *random effect* menjadi tidak konsisten, sebaliknya estimasi dengan *fixed effect* tetap konsisten. Jika nilai statistik Hausman lebih besar daripada nilai kritis statistik *chi-square*, maka hipotesis nul akan ditolak, yang berarti estimasi yang

tepat untuk regresi data panel adalah metode *fixed effect* dari pada metode *random effect*

Sementara itu, Judge *et al.* dalam Gujarati (2003) memberikan sejumlah pertimbangan terkait pilihan apakah menggunakan model *fixed effect* (FE) ataukah model *random effect* (RE). Pertimbangan-pertimbangan itu adalah sebagai berikut:

1. Jika jumlah data *time series* (T) besar dan jumlah data *cross-section* (N) kecil, ada kemungkinan perbedaan nilai parameter yang diestimasi dengan FE dan RE cukup kecil. Karena itu, pilihan ditentukan berdasarkan kemudahan perhitungan. Dalam hal ini, adalah model FE.
2. Ketika N besar dan T kecil estimasi kedua metode dapat berbeda secara signifikan. Pada kondisi seperti ini, pilihan ditentukan berdasarkan keyakinan apakah individu yang diobservasi merupakan sampel acak yang diambil dari populasi tertentu atau tidak. Jika observasi bukan merupakan sampel acak, maka digunakan model FE. Jika sebaliknya, maka digunakan model RE.
3. Jika efek individu tidak teramati *ai* berkorelasi dengan satu atau lebih variabel bebas, maka estimasi dengan RE bias, sedangkan estimasi dengan FE tidak bias.
4. Jika N besar dan T kecil, serta semua asumsi yang disyaratkan oleh model RE terpenuhi, maka estimasi dengan menggunakan RE lebih efisien dibanding estimasi dengan FE.

Dalam penelitian ini, penentuan apakah model FE atau RE yang akan digunakan selain didasarkan pada sejumlah pertimbangan yang telah disebutkan, juga akan didasarkan pada kriteria ekonomi (*make sense* secara ekonomi). Dalam hal ini, adalah kesesuaian tanda hasil estimasi koefisien regresi setiap variabel di dalam model dengan teori dan kewajaran besaran nilai koefisien hasil estimasi tersebut.

Setelah menentukan spesifikasi model yang akan digunakan, tahapan selanjutnya adalah memilih metode estimasi (estimator) yang tepat sesuai dengan struktur varian kovarian residual.

Konsekuensi yang muncul ketika membangun model regresi dengan data panel adalah bertambahnya komponen residual, karena adanya dimensi *cross-section* dan *time-series* pada data. Kondisi ini menyebabkan matriks varian kovarian residual menjadi sedikit lebih kompleks bila dibandingkan dengan model regresi klasik yang hanya menggunakan data *cross-section* atau data *time-series*.

Pada model regresi klasik, pelanggaran terhadap asumsi klasik terkait residual, seperti heterokedastisitas dan autokorelasi merupakan masalah serius yang mengakibatkan penduga parameter regresi yang diestimasi dengan OLS tidak lagi bersifat BLUE (*best linier unbiased estimator*). Tindakan yang biasa dilakukan untuk mengatasi masalah ini adalah dengan melakukan penghitungan *robust standard error*. Dalam pemodelan regresi dengan data panel, terjadinya pelanggaran asumsi regresi linier klasik pada residual adalah hal yang sangat sulit dihindari, dan tidak

seperti pada regresi klasik, pelanggaran dapat diakomodasi untuk menentukan metode estimasi terbaik bagi spesifikasi model yang digunakan.

Terdapat beberapa kemungkinan struktur varian kovarian residual yang mungkin terjadi pada model regresi data panel. Berbagai kemungkinan yang dibahas pada bagian ini adalah yang biasa dijumpai pada estimasi model dengan *common effects* dan *fixed effect*. Karena itu, metode-metode estimasi yang dapat digunakan terkait struktur varian kovarian residual yang dipaparkan pada bagian ini hanya akan diterapkan pada model yang diestimasi dengan *common effects* atau *fixed effect*.

**a. Pemilihan Estimator Struktur Homoskedastik atau Heteroskedastik dengan Uji Lagrange Multiplier (LM)**

Pada pengujian ini, hipotesis null ( $H_0$ ) yang digunakan adalah bahwa struktur varians-covarians residual bersifat homoskedastik. Sementara hipotesis alternatif ( $H_1$ ) adalah bahwa struktur varians-covarians residual bersifat heteroskedastik.

Secara matematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut :

$$LM = \frac{T}{2} \sum_{i=1}^n \left[ \frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right]^2$$

di mana  $T$  adalah jumlah periode waktu,  $n$  adalah jumlah individu,  $\hat{\sigma}_i^2$  adalah varians residual persamaan ke- $i$  pada kondisi homoskedastik,

dan  $\hat{\sigma}^2$  adalah *Sum Square Residual (SSR)* persamaan *system* pada kondisi homoskedastik.

Statistik uji LM ini mengikuti distribusi statistik *chi-square* dengan derajat bebas sebanyak  $n-1$ . Jika nilai statistik LM lebih besar dari nilai kritis statistik *chi-square*, maka hipotesis nul ditolak, yang berarti struktur varians-covarians residual bersifat heteroskedastik.

**b. Pemilihan Estimator Struktur Heteroskedastik dan Ada Cross Sectional Correlation**

Pengujian ini dilakukan apabila hasil pengujian LM menunjukkan bahwa struktur varians-covarians residual bersifat heteroskedastik. Pada pengujian ini, hipotesis nul ( $H_0$ ) yang digunakan adalah bahwa struktur varians-covarians residual bersifat heteroskedastik dan tidak ada *cross sectional correlation*. Sementara hipotesis alternatifnya ( $H_1$ ) adalah bahwa struktur varians-covarians residual bersifat heteroskedastik dan ada *cross sectional correlation (Seemingly Uncorrelated Regression/SUR)*. Secara matematis, statistik uji yang digunakan dapat dirumuskan sebagai berikut:

$$\lambda LM = T \sum_{i=2}^n \sum_{j=1}^{i-1} r_{ij}^2$$

di mana  $T$  adalah jumlah periode waktu,  $n$  adalah jumlah individu, dan  $r_{ij}$  adalah *residual correlation coefficient* antara persamaan ke- $i$  dan ke- $j$ .

Statistik uji LM ini mengikuti distribusi statistik *chi-square* dengan derajat bebas sebanyak  $n(n-1)/2$ . Jika nilai statistik LM lebih

besar dari nilai kritis statistik *chi-square*, maka hipotesis nul akan ditolak, yang berarti struktur varians-covarians residual bersifat heteroskedastik dan ada *cross sectional correlation* (*Seemingly Uncorrelated Regression/SUR*).

Apabila model yang digunakan adalah model *fixed effect*, maka pengujian autokorelasi dapat diabaikan.<sup>38</sup> Hal ini karena model *fixed effect* dapat membedakan efek individual dan efek waktu dan juga tidak perlu mengasumsikan bahwa komponen *error* tidak berkorelasi dengan variabel bebas yang mungkin sulit dipenuhi.<sup>39</sup>

Pengujian statistik juga dilakukan untuk mengetahui apakah model regresi non linier merupakan model yang tepat untuk menggambarkan hubungan antar variabel dan apakah ada hubungan yang signifikan diantara variabel-variabel dependen dengan variabel-variabel penjelas (seperti: uji statistik t dan uji statistik F) selain itu kita bisa melihat nilai hasil estimasi untuk  $R^2$  (koefisien determinasi).

## 2. Uji Hipotesis

Dalam pengujian hipotesis ini digunakan uji statistik, yang meliputi uji t, uji F dan melihat koefisien determinasi ( $R^2$ ):

### a. Uji Parsial (uji t)

Uji parsial (uji t) ini digunakan untuk mendeteksi seberapa baik variabel bebas (*independent variable*) dapat menjelaskan variabel tidak

---

<sup>38</sup> Wahyu Ario Pratomo dan Paidi Hidayat, *Pedoman Praktis Penggunaan Eviews dalam Ekonometrika Edisi Kedua* (Medan: USU Press, 2010)

<sup>39</sup> Nachrowi dkk, op. cit.

bebas (*dependent variable*) secara individu. Langkah-langkah yang harus dilakukan dengan uji t sebagai berikut:

a) Merumuskan hipotesa

$H_0: \beta_i = 0$  Artinya secara individual tidak ada pengaruh yang signifikan dari variabel bebas ke-i terhadap variabel tidak bebas

$H_1: \beta_i \neq 0$  Artinya secara individual ada pengaruh yang signifikan dari variabel bebas ke-i terhadap variabel tidak bebas.

b) Menentukan tingkat confident of level ( $\alpha = 5\%$ ) dan *degree of freedom* (df) dengan rumus  $(nT-n-k)$  untuk mengetahui  $t_{tabel}$ .

c) Mencari nilai  $t_{hitung}$  dengan rumus:

$$t_{hitung} = \frac{\beta_i}{SE(\beta_i)}$$

Dimana:

$\beta_i$  : penafsiran koefisien  $\beta_i$

$SE(\beta_i)$  : standar error  $\beta_i$

$i$  : angka 1, 2, 3

d) Membandingkan hasil  $t_{hitung}$  dengan  $t_{tabel}$  dengan kriteria sebagai berikut:

- $t_{hitung} < t_{tabel}$ , maka  $H_0$  diterima dan  $H_1$  ditolak, artinya tidak ada pengaruh antara variabel independen terhadap variabel dependen.
- $t_{hitung} > t_{tabel}$ , maka  $H_0$  ditolak dan  $H_1$  diterima, artinya ada pengaruh antara variabel independen terhadap variabel dependen.

### b. Uji Serempak (F)

Uji F merupakan pengujian koefisien regresi secara keseluruhan. Pengujian ini menunjukkan apakah semua variabel bebas (*independent variable*) yang dimasukkan ke dalam model mempunyai pengaruh secara bersama-sama terhadap variabel tidak bebas (*dependent variabel*). Pengujian semua koefisien penaksiran regresi secara bersama-sama dilakukan dengan uji F dengan langkah-langkah sebagai berikut:

a) Merumuskan hipotesa

$H_0: \beta_1 = \beta_2 = 0$ , artinya bersama-sama tidak ada pengaruh yang signifikan dari variabel bebas terhadap variabel tidak bebas.

$H_1: \beta_1 \neq \beta_2 \neq 0$ , artinya secara bersama-sama ada pengaruh yang signifikan dari variabel bebas terhadap variabel tidak bebas.

b) Menentukan tingkat *confident of level* ( $\alpha = 5\%$ ) dan *degree of freedom* (df) dengan rumus ( $N_1 = n-k-1$  dan  $N_2 = nT-n-k$ ) untuk mengetahui  $F_{\text{tabel}}$ .

c) Mencari nilai  $F_{\text{hitung}}$  dengan rumus:

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

Dimana:

$R^2$  : koefisien determinasi

$n$  : banyaknya observasi

$k$  : banyaknya variabel-variabel bebas

d) Membandingkan hasil  $F_{hitung}$  dengan  $F_{tabel}$  dengan kriteria sebagai berikut:

- $F_{hitung} < F_{tabel}$ , berarti  $H_0$  diterima dan  $H_1$  ditolak (tidak signifikan).
- $F_{hitung} > F_{tabel}$ , berarti  $H_0$  ditolak dan  $H_1$  diterima (signifikan).

**c. Koefisien Determinasi ( $R^2$ )**

$R^2$  digunakan untuk mengukur kebaikan atau kesesuaian suatu model persamaan regresi.

Besaran  $R^2$  dihitung dengan rumus:

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

Dan  $R^2$  *adjusted* dihitung dengan rumus:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{nT - 1}{nT - n - k}$$

Dimana:

*ESS*: Jumlah kuadrat yang dijelaskan

*RSS*: Jumlah kuadrat residual.

*TSS*: Jumlah kuadrat total.

*n*: Jumlah observasi (propinsi)

*T*: Jumlah periode waktu.

*k*: Banyaknya variabel bebas tanpa intersep.

*Adjusted  $R^2$*  digunakan karena sudah menghilangkan pengaruh penambahan variabel bebas dalam model, karena nilai  $R^2$  akan terus naik

seiring dengan penambahan variabel bebas. Karena itu kita harus berhati-hati dalam menggunakan nilai  $R^2$  ketika menilai kebaikan dan kesesuaian suatu model persamaan regresi. Penggunaan *adjusted*  $R^2$  sudah memperhitungkan jumlah derajat bebas.

Koefisien determinasi merupakan angka yang menunjukkan besarnya derajat kemampuan yang dapat menerangkan variabel bebas terhadap variabel terikat.

Nilai  $R^2$  berkisar antara 0 – 1 ( $0 < R^2 < 1$ ) yang berarti semakin mendekati satu, maka semakin dekat pula hubungan antara variabel bebas dan variabel terikat serta dapat dikatakan model tersebut adalah model terbaik.

### 3. Uji Asumsi Klasik

Untuk membangun persamaan regresi panel yang terbaik dari kriteria ekonometrika, perlu dilakukan penyelidikan dan penanganan adanya masalah-masalah yang berkaitan dengan pelanggaran asumsi dasar. Berikut ini adalah asumsi-asumsi yang diperlukan dalam analisis regresi:

#### a. Normalitas

Pengujian normalitas dilakukan untuk mengetahui apakah residual berdistribusi normal atau tidak. Berdasarkan dalil limit pusat (*central limit theorem*), ada kecenderungan residual yang terjadi sebenarnya menyebar secara normal. Jika residual merupakan jumlah residual dari beberapa sumber, maka apapun sebaran peluang masing-masing residual itu, akan mendekati sebaran normal bila komponen residual semakin banyak. Salah

satu metode yang dapat digunakan untuk menguji Normalitas adalah *Jarque-Bera test*. Uji statistik ini dapat dihitung dengan rumus berikut:

$$JB = n \left[ \frac{\mu_3^2}{6\mu_2^3} + \frac{(\frac{\mu_4}{\mu_2} - 3)^2}{24} \right]$$

di mana:

$n$  = jumlah sampel       $\mu_2$  = varians

$\mu_3$  = skewness           $\mu_4$  = kurtosis

*Jarque-Bera test* mempunyai distribusi *chi square* dengan derajat bebas dua. Jika hasil *Jarque-Bera test* lebih kecil dari nilai *chi square* pada  $\alpha=5\%$ , maka terima hipotesis nul yang berarti berdistribusi normal. Jika hasil *Jarque-Bera test* lebih besar dari nilai *chi square* pada  $\alpha=5\%$ , maka tolak hipotesis nul yang berarti *error term* tidak berdistribusi normal.

## b. Multikolinieritas

Multikolinieritas berarti adanya korelasi antar variabel bebas, yang terjadi karena variabel-variabel bebas tersebut memiliki hubungan pada populasi atau hanya pada sampel. Cara mendeteksi adanya kolinieritas:

- Uji formal untuk mendeteksi keberadaan dari multikolinieritas yaitu dengan melihat nilai *Variance Inflation Factor* (VIF). Nilai VIF ini menunjukkan bagaimana varians dari sebuah estimator akan meningkat akibat adanya multikolinieritas. Nilai VIF diperoleh dengan formula berikut:

$$VIF_k = \frac{1}{1 - R_k^2}$$

di mana  $k = 1, 2, \dots, p - 1$  dan  $R_k^2$  merupakan koefisien determinasi dari regresi berganda ketika  $X_k$  diregresikan dengan  $p-2$  variabel lainnya dalam model.

Jika nilai VIF lebih dari 10, maka hal tersebut dapat berindikasi bahwa multikolinieritas bersifat serius dan akan memengaruhi estimasi yang menggunakan OLS karena meskipun estimator tetap bersifat *unbiased* namun sudah tidak lagi memiliki varians yang minimum. Selain itu, keberadaan multikolinieritas juga akan membuat estimator bersifat sensitif untuk perubahan yang kecil pada data, sehingga akan mengakibatkan kesalahan (*misleading*) dalam menginterpretasikan suatu model regresi.

Cara mengatasi adanya multikolinieritas antara lain melepas satu atau lebih variabel yang memiliki korelasi yang tinggi, mentransformasi model, atau memperbesar jumlah sampel.