

BAB III

METODE PENELITIAN

3.1 Objek dan Subjek Penelitian

Dalam penelitian ini yang menjadi objek penelitian adalah Kemiskinan (Y) merupakan variabel terikat (*dependent variable*) dan Pendidikan (X_1) dan Kesehatan (X_2) merupakan variabel bebas (*independent variable*). Sedangkan yang menjadi subjek penelitian adalah kepala rumah tangga di negara Indonesia. Unit analisis yang digunakan dalam penelitian ini adalah kepala rumah tangga. Alasan kepala rumah tangga yang dijadikan observasi diharapkan dapat mencerminkan bahwa apakah rumah tangga tersebut cenderung hidup dalam kemiskinan atau jauh dari kemiskinan.

3.2 Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah penelitian kuantitatif. Penelitian kuantitatif merupakan metode-metode tertentu untuk menguji teori-teori tertentu dengan cara meneliti hubungan antarvariabel. Variabel-variabel ini diukur dengan instrument-instrumen penelitian sehingga data terdiri dari angka-angka yang dapat dianalisis berdasarkan prosedur-prosedur statistik (Creswell, 2019). Dengan tujuan penelitiannya tergolong pada penelitian eksplanatori. Penelitian eksplanatori merupakan penelitian yang menjelaskan pengaruh antara variabel satu dengan variabel lainnya.

3.3 Desain Penelitian

3.3.1 Definisi Operasional

Berdasarkan ranah penelitian ini yakni Pendidikan dan Kesehatan sebagai variabel bebas dan Kemiskinan sebagai variabel terikat, maka untuk lebih memahami mengenai penggunaan variabel yang digunakan dalam penelitian ini, penulis membuat operasional variabel yang dapat dilihat pada Tabel 3.1

Tabel 3. 1 Definisi Operasional Variabel

Variabel	Konsep	Definisi Operasional	Sumber Data
Variabel Terikat			
Kemiskinan	Kemiskinan adalah ketidakmampuan individu dalam memenuhi kebutuhan minimal untuk hidup layak, baik makanan maupun nonmakanan (BPS).	Variabel kemiskinan diukur dengan membandingkan PCE (<i>Per Capita Consumption Expenditure</i>) dengan garis kemiskinan di tahun 2014 sebesar Rp312.328. Variabel ini berupa <i>dummy</i> dengan ketentuan apabila PCE di bawah garis kemiskinan, maka kepala keluarga tersebut termasuk dalam miskin dan diberi nilai 1 (satu). Sedangkan, PCE yang di atas garis kemiskinan, maka kepala rumah tangga tersebut termasuk tidak miskin dan diberi nilai 0 (nol).	Data PCE (<i>Per Capita Consumption Expenditure</i>) IFLS-5

Variabel	Konsep	Definisi Operasional	Sumber Data
Variabel Bebas			
Pendidikan	Pendidikan merupakan salah satu sumber dasar dalam modal manusia yang mempunyai pengetahuan dan keterampilan. Pengetahuan dan keterampilan produktif yang dapat meningkatkan produktivitas masyarakat (Tilak, 2002).	Variabel pendidikan dilihat dari <i>years of schooling</i> (lama sekolah) dalam tahun pada kepala rumah tangga.	IFLS-5 buku K; data ar16 (Apa tingkat pendidikan tertinggi yang pernah/sedang diikuti oleh Ibu/Bapak/Sdr?) dan data ar17 (Apa tingkat/kelas tertinggi yang pernah diselesaikan oleh Ibu/Bapak/Sdr pada sekolah ini?)
Kesehatan	Kesehatan adalah komponen langsung dari kesejahteraan manusia dan bentuk modal manusia yang dapat meningkatkan kemampuan individu (Bloom & Canning, 2003)	Variabel kesehatan dapat dilihat dari keadaan kesehatan pada kepala rumah tangga yang akan menjadi variabel <i>dummy</i> .	IFLS-5 buku 3B; data KK01 (Secara umum, bagaimana keadaan kesehatan Ibu/Bapak/Sdr saat ini?)

3.3.2 Populasi dan Sampel

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari data *Indonesian Family Life Survey* (IFLS) yang dikumpulkan oleh RAND corporation. Data tersebut berjumlah 15.900 rumah tangga dan 709

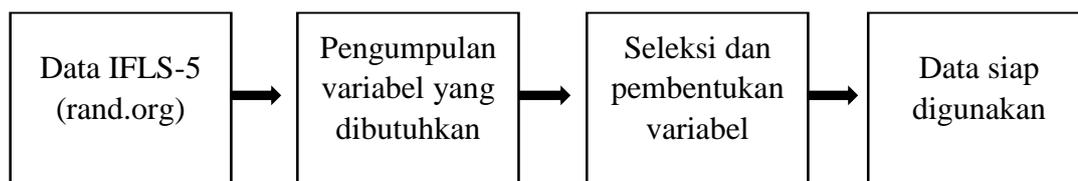
komunitas dengan jumlah individu dalam rumah tangga sebanyak 50.000 individu (Strauss et al., 2016). Data IFLS yang digunakan dalam penelitian ini adalah IFLS *wave 5* (gelombang ke-5) yang dilakukan pada tahun 2014. Data ini digunakan untuk melihat probabilitas untuk kecenderungan menjadi miskin atau tidak miskin.

Sampel yang digunakan dibatasi pada kepala rumah tangga yang berusia 15 tahun dan di atas 15 tahun. Dari 14.238 observasi dari seluruh Indonesia, 501 observasi yang dikategorikan sebagai miskin dan 13.737 dikategorikan tidak miskin.

3.3.3 Teknik dan Alat Pengumpulan Data

Pengumpulan data pada penelitian ini dengan teknik dokumentasi atau penelitian yang menggunakan data sekunder yang diperoleh dari publikasi lembaga survey yaitu *Indonesia Family Life Survey-5* (IFLS-5) tahun 2014 berupa data *longitudinal*. Pengolah data yang dilakukan menggunakan stata.

Berikut langkah-langkah pengumpulan data:



Gambar 3.1 Langkah Pengumpulan Data

1. Pengumpulan variabel yang dibutuhkan

Dalam IFLS-5 terdapat buku pedoman pertanyaan (*IFLS Household Survey Questionnaires/HH*) tahun 2014. Pada buku ini terdapat berbagai topik kuesioner yang dapat dijadikan komponen variabel. Penelitian ini menggunakan buku K untuk variabel pendidikan dan buku 3B untuk variabel kesehatan yang akan menjadi variabel bebas. Sedangkan, variabel terikat yakni kemiskinan bersumber dari data set khusus pada IFLS *Consumption; Expenditure Aggregates* yakni data PCE (*Per Capita Consumption Expenditure*)

2. Seleksi dan pembentukan variabel

Selanjutnya dilakukan seleksi variabel dari seksi-seksi pertanyaan kuesioner. Pada variabel pendidikan dilihat dari lama sekolah dengan menggunakan seksi kuesioner AR dengan kolom pertanyaan AR16 dan

akan dibandingkan dengan AR17. Selain itu, variabel kesehatan menggunakan seksi kuesioner 3B dengan kolom pertanyaan KK01 dijadikan variabel *dummy*. Variabel terikat pada penelitian ini berbentuk *dummy* yang dibentuk dari perbandingan PCE (*Per Capita Consumption Expenditure*) yang didapatkan dari data set khusus dengan garis kemiskinan di tahun 2014.

3.3.4 Teknik Analisis Data

a. Spesifikasi Model

Penelitian ini menggunakan spesifikasi model regresi logit (*binary logistic regression*). Model ini merupakan pemodelan yang variabel dependennya bersifat biner atau sering disebut *dummy*. Kategori dari variabel dependen dalam model tersebut adalah angka 1 dan 0. Model ini bertujuan untuk menemukan probabilitas dari sebuah kejadian. Berikut spesifikasi model dalam penelitian ini:

$$L_1 = \beta_0 + \beta_1 X_1 + \beta_2 D_1 + \beta_3 D_2 + \beta_4 D_3 + u \dots \dots \dots (1)$$

$$L_2 = \beta_0 + \beta_1 X_1 + \beta_2 D_1 + u \dots \dots \dots (2)$$

Persamaan (1), L_i merupakan nilai variabel dependen atau logit, yakni probabilitas kejadian untuk menjadi miskin oleh kepala rumah tangga di Indonesia. X_1 merupakan pendidikan diproksi dalam lama sekolah, X_2 merupakan variabel kesehatan yang diproksi menjadi variabel *dummy group* dengan empat kategori, tiga kategori diantaranya (Sangat Sehat, Cukup Sehat, Kurang Sehat) dijadikan variabel *dummy* dan satu kategori (Tidak Sehat) menjadi kategori dasar. D_1 merupakan kategori sangat sehat (1=Sangat Sehat, 0=Lainnya), D_2 merupakan kategori cukup sehat (1=Cukup Sehat, 0=Lainnya), D_3 merupakan kategori kurangs sehat (1=Kurang Sehat, 0=Lainnya).

Persamaan (2) merupakan modifikasi model persamaan penelitian 1 dimana letak perbedaannya ada pada variabel kesehatan yang hanya dijadikan dua kategori, yakni Sehat dan Sakit sebagai kategori dasar. L_i merupakan nilai variabel dependen atau logit, yakni probabilitas kejadian

untuk menjadi miskin oleh kepala rumah tangga di Indonesia. X_1 merupakan pendidikan yang diturunkan dalam lama sekolah, D_1 merupakan kesehatan (1= Sehat, 0=Sakit)

Kemudian kedua persamaan tersebut terdapat variabel kontrol yang berupa *dummy* terdiri dari jenis kelamin (1=Pria, 0=Wanita), asal daerah (1=Kota, 0=Desa), asal wilayah (1=Jawa, 0= Luar Jawa), status pernikahan (1=Menikah, 0=Tidak Menikah), dan usia yang dibuat dalam *variabel dummy group* dengan empat kategori, tiga kategori diantaranya ada usia <40 tahun, 40-49 tahun, dan 50-59 tahun., sedangkan usia >60 tahun menjadi kategori dasar. Usia <40 tahun (1= <40 tahun, 0=Lainnya), usia 40-49 tahun (1=40-49 tahun, 0=Lainnya), usia 50-59 tahun (1=50-59 tahun, 0=Lainnya).

b. Pengujian Hipotesis

Berdasarkan Gujarati & Porter (2012; hlm. 198), OLS (*Ordinary Least Square*) dan WLS (*Weighted Least Squares*) tidak digunakan dalam model logit, melainkan dengan menggunakan estimasi nonlinear dengan metode *maximum likelihood*. Metode *maximum likelihood* digunakan untuk sampel besar dan estimasi standar error bersifat asimptotik. Sehingga ada beberapa hal yang perlu diperhatikan untuk menginterpretasi hasilnya, yakni:

(1) Pengujian hipotesis secara parsial (Uji z)

Nilai t dalam analisis regresi dengan OLS tidak digunakan di analisis regresi logit, sebagai gantinya, digunakan nilai statistic z, yang karakteristiknya mirip dengan nilai statistic t (Winarno, 2015; hlm. 6.14). Hal ini juga sesuai Gujarati & Porter (2012; hlm.198) karena signifikansi secara statistik untuk koefisien tidak dievaluasi menggunakan statistik t, tetapi menggunakan statistik z.

Pada uji statistik z, keputusan dalam hal menolak hipotesis nol jika nilai absolut statistik z hitung lebih besar dari nilai z kritis dari distribusi tabel z. Sementara, pada prosedur uji nilai p hanya membandingkan *p-value* < 0,05 maka menolak H_0 dan menerima H_a , sebaliknya jika *p-value* > 0,05 maka menerima H_0 dan menolak H_a . Hipotesis yang akan diuji berdasarkan penelitian ini adalah sebagai berikut.

- a) $H_0 : \beta_1 \geq 0$: pendidikan dan kesehatan tidak berpengaruh negatif terhadap kemiskinan
- b) $H_a : \beta_1 < 0$: pendidikan dan kesehatan berpengaruh negatif terhadap kemiskinan

(2) Pengujian hipotesis secara simultan (Uji F)

Uji ini digunakan untuk menguji apakah semua variabel independen dalam model estimasi secara serentak mempengaruhi variabel dependen. Statistik *Likelihood Ratio* (LR) berfungsi sebagai uji F pada model regresi logit. Uji ini bertujuan untuk menunjukkan signifikansi keseluruhan variabel independen yang mempengaruhi variabel dependen.

$$LR = 2(\mathcal{L}_{ur} - \mathcal{L}_r)$$

(Wooldridge, 2019)

Jika nilai probabilitas LR statistic $< \alpha = 5$ persen, maka seluruh variabel independen signifikan menjelaskan variabel dependen. Sebaliknya, jika nilai probabilitas LR statistic $> \alpha = 5$ persen, maka seluruh variabel independen tidak signifikan menjelaskan variabel dependen.

(3) *Goodness of fit* (Uji Koefisien Determinasi/Pseudo R^2)

Menurut Gujarati & Porter (2012; hlm. 199), R^2 (koefisien determinasi) pada umumnya tidak dapat digunakan dalam model dengan variabel dependen yang biner. Sehingga, ukuran yang serupa dengan R^2 yaitu *pseudeo- R^2* . Ada banyak macam ukuran dari *pseudeo- R^2* diantaranya ada *R^2 McFadden*, dimana sama seperti R^2 yang juga bernilai 0 dan 1. Ukuran *goodness of fit* lainnya yang tergolong sederhana dibandingkan dengan ukuran lainnya adalah *count R^2* (R^2 hitung) yang didefinisikan sebagai berikut.

$$R^2 \text{ Hitung} = \frac{\text{jumlah prediksi yang benar}}{\text{total jumlah yang diobservasi}}$$

Nilai pseudo R^2 berkisar antara 0 dan 1 ($0 < R^2 < 1$) dengan ketentuan sebagai berikut.

- a) Jika R^2 semakin mendekati angka 1, maka hubungan antara variabel independen dengan variabel dependen semakin erat atau dekat, dengan kata lain model tersebut dapat dinilai baik.

b) Jika R^2 semakin menjauhi angka 1, maka hubungan antara variabel independen dengan variabel dependen jauh atau tidak dekat, dengan kata lain model tersebut dinilai kurang baik.

c. *Marginal Effect*

Besarnya pengaruh masing-masing variabel akan dijelaskan melalui estimasi nilai *marginal effect*. Nilai dari *marginal effect* lebih informatif dibandingkan koefisien estimasi. Nilai *marginal effect* memberikan tingkat perubahan probabilitas untuk setiap unit perubahan regressor (Gujarati & Porter, 2012; hlm.202).