

**ANALISIS REGRESI TOBIT PADA PERMASALAHAN  
PENGELUARAN KONSUMSI ROKOK  
KOTA KEDIRI TAHUN 2011**

**SKRIPSI**

**Sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Sains dalam bidang Statistika**

oleh:  
**GILANG PERMANA**  
**0910950007-95**



**PROGRAM STUDI STATISTIKA  
JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS BRAWIJAYA  
MALANG  
2013**

**LEMBAR PENGESAHAN SKRIPSI**

**ANALISIS REGRESI TOBIT PADA PERMASALAHAN  
PENGELUARAN KONSUMSI ROKOK  
KOTA KEDIRI TAHUN 2011**

oleh:  
**GILANG PERMANA**  
0910950007-95

Setelah dipertahankan di depan Majelis Penguji  
pada tanggal 4 November 2013  
dan dinyatakan memenuhi syarat untuk memperoleh gelar Sarjana  
Sains dalam bidang Statistika

Pembimbing I

Pembimbing II

Dr. RahmaFitriani, S.Si., M.Sc  
NIP. 197603281999032001

Prof. Dr. Ir. Waego H. Nugroho  
NIP. 195212071979031003

Mengetahui,  
Ketua Jurusan Matematika  
Fakultas MIPA Universitas Brawijaya

Dr. Abdul Rouf Alghofari, M.Sc  
NIP. 196709071992031001

## LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Gilang Permana  
NIM : 0910950007-95  
Jurusan : Matematika  
Penulisan Skripsi berjudul :

### ANALISIS REGRESI TOBIT PADA PERMASALAHAN PENGELUARAN KONSUMSI ROKOK KOTA KEDIRI TAHUN 2011

Dengan ini menyatakan bahwa:

1. Isi dari Skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Skripsi.
2. Apabila di kemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan, maka saya bersedia menanggung segala risiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 4 November 2013  
Yang menyatakan

(Gilang Permana)  
0910950007-95

# ANALISIS REGRESI TOBIT PADA PERMASALAHAN PENGELUARAN KONSUMSI ROKOK KOTA KEDIRI TAHUN 2011

## ABSTRAK

Variabel laten merupakan variabel yang tidak dapat diukur secara langsung. Variabel pengeluaran rumah tangga untuk konsumsi rokok merupakan data tersensor yang digunakan sebagai variabel pengganti dari variabel laten yaitu keinginan untuk membeli atau mengkonsumsi. Salah satu metode yang digunakan untuk menganalisis data tersensor ini adalah regresi Tobit. Penelitian ini bertujuan untuk melihat kekonsistenan penduga yang dihasilkan Regresi Tobit relatif terhadap OLS tentang pengeluaran konsumsi kota Kediri tahun 2011. Data yang digunakan adalah data hasil SUSENAS tahun 2011 sebanyak 217 rumah tangga. Variabel respon yang digunakan dalam penelitian ini adalah pengeluaran rumah tangga untuk konsumsi rokok. Variabel prediktor yang digunakan antara lain umur, proporsi anggota rumah tangga dewasa (PART), pendapatan rumah tangga (PDT), harga barang komplementer (HBK). Dari 217 rumah tangga diambil sampel berukuran 30 secara acak dan dilakukan pengulangan sebanyak 100 kali, pendugaan parameter dilakukan pada masing - masing data. Berdasarkan indikator kesamaan tanda, signifikansi, uji serempak dan model yang terbaik menunjukkan bahwa regresi *Tobit* secara konsisten lebih baik dari pada regresi *OLS*.

**Kata-kata kunci** : variabel laten, data tersensor, regresi *tobit*,  
pengeluaran rumah tangga, konsumsi rokok

# **TOBIT REGRESSION ANALYSIS ON CIGARETTE CONSUMPTION EXPENDITURES PROBLEM IN KEDIRI CITY 2011**

## **ABSTRACT**

Latent variables are variables that can not be measured directly. Household expenditure variables for cigarette consumption is censored data is used as a proxy variable of latent variables that wishes to purchase or consume. One method used to analyze the data is censored Tobit regression. This study aimed to examine the consistency of the resulting estimator Tobit relative to the OLS regression of consumption expenditure in Kediri 2011. Source of data used is SUSENAS 2011 of 217 households. Response variables used in this study is the household expenditure on tobacco consumption. Predictor variables used include age, the proportion of adult household members (PART), household income (PDT), the price of complementary goods (HBK). Having taken 30 random data and be repeated 100 times, based on the similarity indicator sign, significance, simultaneously test and the best model suggests that Tobit regression consistently better than OLS regression.

**Key words** : latent variables, censored data, tobit regression, household expenditure, cigarette consumption



## KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala limpahan karunia dan kasih-Nya serta sholawat dan salam kepada Nabi Muhammad SAW atas segala ketauladanannya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “**Analisis Regresi Tobit Pada Permasalahan Pengeluaran Konsumsi Rokok Kota Kediri Tahun 2011**”. Tugas Akhir ini disusun sebagai salah satu syarat untuk menyelesaikan Program Sarjana Statistika, Fakultas MIPA, Universitas Brawijaya.

Dalam penyelesaian Tugas Akhir ini, tidak terlepas dari bantuan dan dukungan berbagai pihak. Maka dari itu, penulis ingin mengucapkan terima kasih kepada:

1. Ayah, ibu, kakak Yudhistira dan Putri sekeluarga yang telah memberi dukungan moril dan materi serta doa restu.
2. Ibu Dr. Rahma Fitriani S.Si., M.Sc selaku Dosen Pembimbing I dan Bapak Prof. Dr. Ir. Waego Hadi Nugroho selaku Dosen Pembimbing II atas segala masukan, nasehat dan bimbingan selama proses penyelesaian Tugas Akhir.
3. Ibu Eni Sumarminingsih, S.Si., MM. selaku Dosen Penguji atas saran dan masukan serta nasehat yang telah diberikan.
4. Bapak Dr. Abdul Rouf Alghofari, MSc. selaku Ketua Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang.
5. Bapak dan Ibu Dosen Statistika atas ilmu yang diberikan selama kuliah.
6. Teman-teman keluarga Statistika angkatan 2009 terutama untuk SMESH yang selalu memberikan bantuan dan semangat.
7. Semua pihak yang telah membantu penulisan skripsi ini yang tidak dapat penulis sebutkan satu-persatu.

Penulis menyadari keterbatasan dalam penulisan skripsi ini. Oleh karena itu saran ataupun kritik yang membangun akan sangat berguna bagi penulis untuk mengembangkan kemampuan menulis ilmiah. Semoga skripsi ini dapat bermanfaat bagi para pembaca.

Malang, 4 November 2013

Penulis

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>HALAMAN PENGESAHAN</b> .....	ii
<b>HALAMAN PERNYATAAN</b> .....	iii
<b>ABSTRAK</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>KATA PENGANTAR</b> .....	vi
<b>DAFTAR ISI</b> .....	vii
<b>DAFTAR GAMBAR</b> .....	ix
<b>DAFTAR TABEL</b> .....	x
<b>DAFTAR LAMPIRAN</b> .....	xi
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Tujuan Penelitian .....	2
1.4 Batasan Masalah .....	2
1.5 Manfaat Penelitian .....	2
<b>BAB II TINJAUAN PUSTAKA</b>	
2.1 Regresi Tobit .....	3
2.2 Distribusi Normal Tersensor .....	3
2.3 Data Tersensor .....	5
2.4 Pendugaan Parameter .....	6
2.5 Pengujian Pendugaan Parameter .....	10
2.6 Asumsi-asumsi Analisis Regresi Tobit .....	12
2.6.1 Kenormalan Galat .....	12
2.6.2 Multikolinieritas .....	13
2.6.3 Kesamaan Ragam Galat .....	15
2.6.4 Autokorelasi .....	16
2.7 Penentuan Keباikaaan Model .....	17
2.8 Tinjauan Non Statistik .....	18

### **BAB III METODOLOGI PENELITIAN**

3.1 Sumber Data.....	21
3.2 Variabel Penelitian.....	21
3.2.1 Variabel Respon .....	21
3.2.2 Variabel Prediktor .....	21
3.3 Metode Analisis .....	22
3.4 Diagram Alir .....	23

### **BAB IV HASIL DAN PEMBAHASAN**

4.1 Analisis Deskriptif Pengeluaran Rumah Tangga untuk Biaya Konsumsi Rokok.....	25
4.1.1 Deskripsi Variabel-variabel yang Mempengaruhi Pengeluaran Rumah Tangga untuk Biaya Konsumsi Rokok .....	25
4.2 Pemodelan Pengeluaran Rumah Tangga untuk Biaya Konsumsi Rokok .....	27
4.2.1 Pengujian Secara Serentak.....	28
4.2.2 Pengujian Secara Individu .....	31
4.2.3 Pengujian Asumsi.....	32

### **BAB V KESIMPULAN DAN SARAN**

5.1 Kesimpulan .....	37
5.2 Saran.....	38

<b>DAFTAR PUSTAKA</b> .....	39
<b>LAMPIRAN</b> .....	43



## DAFTAR GAMBAR

	Halaman
Gambar 2.1 Contoh Distribusi Normal Tersensor.....	4
Gambar 2.2 Pola Hipotesis Residual yang Ditaksir.....	15
Gambar 2.3 Statistik d Durbin -Watson.....	17
Gambar 3.1 Diagram Alir Analisis Regresi Tobit.....	24

UNIVERSITAS BRAWIJAYA

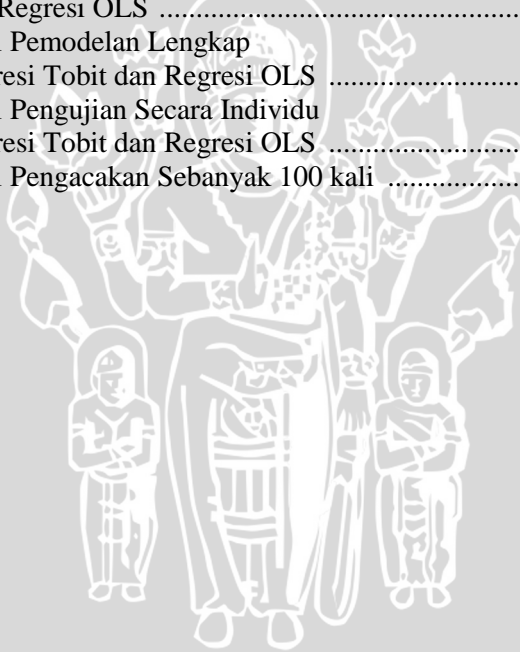


## DAFTAR TABEL

	Halaman
Tabel 2.1 Nilai kritis $D_n \alpha$ .....	13
Tabel 4.1 Deskripsi Pengeluaran Rumah Tangga untuk Konsumsi Rokok.....	25
Tabel 4.2 Deskripsi Variabel Prediktor.....	26
Tabel 4.3 Hasil Akhir Pengujian Serentak Variabel Prediktor Tobit.....	28
Tabel 4.4 Hasil Akhir Pengujian Serentak Variabel Prediktor OLS.....	28
Tabel 4.5 Hasil Pengujian Individu Tobit Antara Variabel Prediktor dengan Pengeluaran Rumah Tangga Untuk Konsumsi Rokok.....	31
Tabel 4.6 Hasil Pengujian Individu OLS Antara Variabel Prediktor dengan Pengeluaran Rumah Tangga Untuk Konsumsi Rokok.....	32
Tabel 4.7 Hasil Pengujian Asumsi antara Regresi Tobit dengan Regresi OLS.....	33
Tabel 4.8 Hasil <i>Resampling</i> Perbandingan antara Model RegresiTobit dengan Model Regresi OLS.....	33

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data Primer Hasil Susenas Pengeluaran Konsumsi Rokok Rumah Tangga Kota Kediri Tahun 2011 .....	43
Lampiran 2. Deskripsi Variabel Pengeluaran Rumah Tangga Untuk Biaya Konsumsi Rokok dengan Variabel Prediktor .....	44
Lampiran 3. Pengujian Asumsi Regresi Tobit dan Regresi OLS .....	45
Lampiran 4. Hasil Pemodelan Lengkap Regresi Tobit dan Regresi OLS .....	49
Lampiran 5. Hasil Pengujian Secara Individu Regresi Tobit dan Regresi OLS .....	50
Lampiran 6. Hasil Pengacakan Sebanyak 100 kali .....	54



# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Dalam penelitian bidang sosial, terdapat suatu variabel yang tidak dapat diukur secara langsung. Variabel semacam itu biasa disebut sebagai variabel laten. Beberapa contoh variabel laten adalah kepuasan konsumen, motivasi, kinerja dan lain sebagainya.

Salah satu kasus variabel laten di bidang ekonomi sosial yaitu keinginan membeli sesuatu atau konsumsi untuk memenuhi kebutuhan. Misalkan keinginan untuk mengkonsumsi rokok. Pada kasus tersebut variabel pengeluaran digunakan sebagai variabel proxy atau variabel pengganti dari variabel keinginan untuk membeli atau mengkonsumsi.

Regresi berganda banyak digunakan di penelitian bidang sosial. Pada umumnya regresi yang digunakan adalah regresi *ordinary least squares (OLS)*. Ketika digunakan variabel pengganti dari variabel laten, penggunaan regresi linier berganda dengan *OLS* dirasa kurang tepat karena sifat terbatas dari variabel pengganti. Jika *OLS* tetap digunakan maka akan dihasilkan penduga yang bias dan tidak konsisten. Untuk mengatasi hal tersebut seharusnya digunakan Regresi *Tobit*.

Regresi *Tobit* merupakan analisis regresi yang digunakan untuk mendekati masalah variabel tak bebas yang akibat sifat terbatasnya menjadi bernilai nol untuk beberapa pengamatan dan bernilai positif untuk selanjutnya. Pada kasus konsumsi rokok, yang dimaksud dengan sifat terbatas tersebut yaitu beberapa individu mempunyai pengeluaran untuk konsumsi rokok dan untuk beberapa individu lainnya tidak mempunyai pengeluaran untuk konsumsi rokok. Menurut Greene (2000) variabel tak bebas yang mempunyai sifat terbatas atau *mixture* (campuran) diskrit dan kontinyu, diskrit untuk yang bernilai nol dan kontinyu untuk yang tidak nol, seperti pada kasus konsumsi rokok ini, maka variabel tak bebas tersebut dikategorikan data tersensor.

Dalam menduga parameter regresi *Tobit* digunakan metode *Maximum Likelihood Estimation (MLE)*. Menurut Hosmer dan Lemeshow (2000) penggunaan metode ini menghasilkan penduga yang konsisten dan efisien untuk sampel yang berukuran besar. Dengan kondisi seperti itu, maka peneliti ingin melakukan penelitian untuk melihat kesamaan tanda penduga yang dihasilkan

Regresi *Tobit* relatif terhadap *OLS* tentang Pengeluaran Konsumsi Rokok pada data Pengeluaran Konsumsi Rokok Kota Kediri tahun 2011.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas maka rumusan masalah penelitian ini adalah :

1. Bagaimana memodelkan permasalahan pengeluaran konsumsi rokok untuk setiap anggota rumah tangga di kota Kediri yang bersifat laten dengan regresi *Tobit*?
2. Bagaimana metode regresi *Tobit* bisa memperbaiki metode regresi *OLS* pada data tersensor?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan di atas, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Membentuk model regresi *tobit* yang digunakan pada permasalahan pengeluaran konsumsi rokok rumah tangga kota Kediri yang bersifat laten.
2. Untuk menunjukkan perbaikan dari penggunaan metode regresi *OLS* ke metode regresi *Tobit* pada permasalahan pengeluaran konsumsi rokok rumah tangga kota Kediri.

## 1.4 Batasan Masalah

Pada penelitian ini, kajian teoritis tentang metode regresi *tobit* pada permasalahan pengeluaran konsumsi rokok rumah tangga kota Kediri pada data dengan variabel respon yang memiliki skala diskrit (bernilai nol) dan berskala kontinu (memiliki nilai tertentu).

## 1.5 Manfaat Penelitian

Manfaat yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mengembangkan wawasan keilmuan dan pengetahuan mengenai variabel respon yang tidak sepenuhnya mewakili *range* dari nilai yang diharapkan, untuk mengatasi masalah tersebut digunakan regresi *tobit*.
2. Memberikan alternatif, penggunaan metode selain metode regresi linier klasik, tentang data tersensor.
3. Memberikan informasi tentang model yang lebih baik diantara model regresi *Tobit* dan model regresi *OLS*.



## BAB II TINJAUAN PUSTAKA

### 2.1 Regresi *Tobit*

Analisis regresi digambarkan melalui sebuah model persamaan garis yang melibatkan dua variabel yakni variabel prediktor ( $X$ ) dan variabel respon ( $Y$ ). Analisis regresi digunakan untuk mengetahui pola hubungan antar variabel. Salah satu jenis dari analisis regresi adalah regresi *Tobit*.

Regresi tersensor atau model *Tobit* merupakan analisis regresi yang digunakan untuk variabel respon yang sebagian datanya memiliki skala pengukuran diskrit dan sebagian yang lain berskala kontinyu. Greene (1997) menyebutkan bahwa variabel respon yang bersifat *mixture* (campuran) memiliki struktur data dengan skala diskrit untuk yang bernilai nol, dan berskala kontinyu untuk yang tidak bernilai nol. Data tersebut disebut data tersensor. Sebaran data tersensor adalah sebaran normal tersensor, yang mengikuti asumsi  $N(\mu, \sigma^2)$ . Formulasi model *Tobit* dalam Tobin (1958) secara umum adalah sebagai berikut.

$$Y_i = \begin{cases} Y_i^*, & Y_i > 0 \\ 0, & Y_i \leq 0 \end{cases} \quad (2.1)$$

di mana  $i = 1, 2, \dots, T$  dan  $Y_i^*$  adalah variabel tak bebas dengan persamaan sebagai berikut.

$$Y_i^* = \beta' X_i + u_i \quad (2.2)$$

Dengan:

$Y_i^*$ : variabel tak bebas (variabel respon)

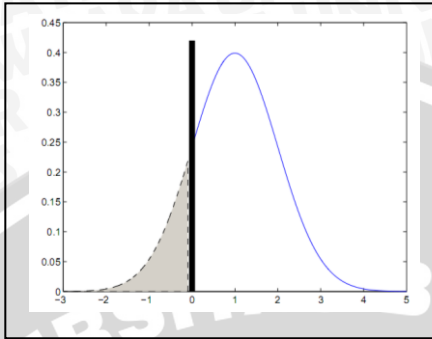
$X$ : vektor variabel bebas (variabel prediktor)

$\beta$ : vektor koefisien yang berukuran  $k \times 1$  yang tidak diketahui,  $k$  adalah banyaknya parameter

$u_i$ : residual model yang mengikuti sebaran normal tersensor ( $0, \sigma^2$ )

### 2.2 Distribusi Normal Tersensor

Variabel tersensor dalam model ini adalah variabel respon atau  $Y$ .  $Y_i^*$  mempunyai sebaran normal dengan mean  $X_i \beta$  dan varians  $\sigma^2$ . Berdasarkan Gambar 2.1 terlihat bahwa pada  $y_i = 0$  berkumpul menjadi satu pada titik nol. Ini merupakan batas sensor. Sedangkan  $y_i > 0$  membentuk kurva normal yang tidak utuh.



Gambar 2.1 Contoh Distribusi Normal Tersensor

Jika  $y_i > 0$ , maka fungsi kepadatan peluangnya adalah

$$f(y_i|x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \beta'x_i)^2\right) \\ = \frac{1}{\sigma} \phi\left(\frac{y_i - \beta'x_i}{\sigma}\right) \quad (2.3)$$

dimana

$$\phi(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \quad (2.4)$$

Persamaan (2.3) merupakan *pdf* untuk distribusi Normal standar. Sedangkan *cdf* dari distribusi Normal standar tersebut sebagai berikut.

$$\Phi(z) = \int_{-\infty}^z \phi(u)du \quad (2.5)$$

Untuk  $y_i = 0$  digunakan probabilitas sebagai berikut.

$$\begin{aligned} Pr(y_i = 0|x_i) &= Pr(y_i^* \leq 0|x_i) \\ &= Pr(\beta'x_i + u_i \leq 0|x_i) \\ &= Pr\left(\frac{u_i}{\sigma} \leq -\frac{\beta'x_i}{\sigma} \mid x_i\right) \\ &= \Phi\left(\frac{-\beta'x_i}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right) \end{aligned} \quad (2.6)$$

Meskipun pada Gambar 2.1 terlihat seperti distribusi normal yang tidak penuh, nilai probabilitasnya tetap satu. Dan nilai ekspektasi untuk  $y_i$  adalah sebagai berikut.

$$\begin{aligned} E(y_i|x_i) &= E(y_i|y_i > 0, x_i)Pr(y_i > 0|x_i) \\ &\quad + E(y_i|y_i = 0, x_i)Pr(y_i = 0|x_i) \end{aligned}$$

dimana :  $E(y_i|y_i = 0, x_i) = 0$

$$\begin{aligned} E(y_i|x_i) &= E(y_i|y_i > 0, x_i)Pr(y_i > 0|x_i) \\ &= E(y_i|y_i > 0, x_i)(1 - Pr(y_i = 0|x_i)) \\ &= E(y_i|y_i > 0, x_i)\Phi\left(\frac{\beta'x_i}{\sigma}\right) \end{aligned} \quad (2.7)$$

Untuk mendapatkan nilai  $E(y_i|y_i > 0, x_i)$  dapat dengan menggunakan bantuan sebagai berikut. Jika  $z \sim N(0,1)$  maka

$$1. E(z|z > c) = \frac{\phi(c)}{1-\Phi(c)}$$

$$2. \phi(-c) = \phi(c)$$

$$3. 1 - \Phi(-c) = \Phi(c)$$

Sehingga,

$$\begin{aligned} E(y_i|y_i > 0, x_i) &= E(\beta'x_i + u_i|\beta'x_i + u_i > 0, x_i) \\ &= \beta'x_i + E(u_i|\beta'x_i + u_i > 0, x_i) \\ &= \beta'x_i + \sigma E\left(\frac{u_i}{\sigma} \middle| \frac{u_i}{\sigma} > \frac{-\beta'x_i}{\sigma}, x_i\right) \\ &= \beta'x_i + \sigma \frac{\phi\left(\frac{\beta'x_i}{\sigma}\right)}{\Phi\left(\frac{\beta'x_i}{\sigma}\right)} \end{aligned} \quad (2.8)$$

Atau bisa ditulis

$$= \beta'x_i + \sigma\lambda\left(\frac{\beta'x_i}{\sigma}\right) \quad (2.9)$$

dimana  $\lambda(c) = \phi(c)/\Phi(c)$

Persamaan (2.8) di substitusi ke dalam persamaan (2.7) untuk mendapatkan nilai ekspektasi sebenarnya dari variabel  $y_i$ .

$$\begin{aligned} E(y_i|x_i) &= \left(\beta'x_i + \sigma \frac{\phi\left(\frac{\beta'x_i}{\sigma}\right)}{\Phi\left(\frac{\beta'x_i}{\sigma}\right)}\right)\Phi\left(\frac{\beta'x_i}{\sigma}\right) \\ &= \Phi\left(\frac{\beta'x_i}{\sigma}\right)(\beta'x_i) + \sigma\phi\left(\frac{\beta'x_i}{\sigma}\right) \end{aligned} \quad (2.10)$$

### 2.3 Data Tersensor

Pada kasus di bidang ekonomi banyak ditemukan permasalahan dengan variabel respon yang tersensor. Jika variabel respon tersensor, maka nilai dalam rentang tertentu ditransformasi ke dalam sebuah nilai tunggal.

Tobin (1958) memberikan contoh data tersensor berupa pengeluaran rumah tangga untuk konsumsi rokok. Rumah tangga dengan pendapatan yang rendah cenderung tidak mempunyai pengeluaran rumah tangga untuk konsumsi rokok.

Variabel laten merupakan suatu variabel yang tidak dapat diukur secara langsung. Variabel laten memiliki sifat terbatas. Yang dimaksud terbatas yaitu sebagian bernilai nol dan bernilai positif untuk selanjutnya. Karena sifat variabel laten yang terbatas maka digunakan regresi Tobit.

## 2.4 Pendugaan Parameter

Penaksiran parameter regresi Tobit menggunakan metode *Maximum Likelihood Estimation* (MLE). Berikut diperoleh fungsi likelihood (Fair, 1977).

$$L = \prod_{y_i > 0} f(y_i) \prod_{y_i = 0} F_{y_i}(0) \quad (2.11)$$

$$\ln L = \ln \left( \prod_{y_i > 0} f(y_i) \prod_{y_i = 0} F_{y_i}(0) \right)$$

$$\ln L = \sum_{y_i > 0} \ln f(y_i) + \sum_{y_i = 0} \ln F_{y_i} \quad (2.12)$$

dengan

$$\begin{aligned} F_{y_i}(0) &= P(y \leq 0) = P\left(Z \leq \frac{0 - \beta' x_i}{\sigma}\right) \\ &= P\left(Z \leq \frac{0 - \beta' x_i}{\sigma}\right) \\ &= \Phi\left(\frac{-\beta' x_i}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{\beta' x_i}{\sigma}\right) \quad (2.13) \end{aligned}$$

$$\begin{aligned} f(y_i) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \beta' x_i)^2\right) \\ &= \sigma^{-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \beta' x_i)^2\right) \quad (2.14) \end{aligned}$$

misal  $z = \frac{y_i - \beta' x_i}{\sigma}$  maka  $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(z)^2\right)$

sehingga  $\phi\left(\frac{y_i - \beta' x_i}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \beta' x_i}{\sigma}\right)^2\right)$

didapatkan

$$f(y_i) = \sigma^{-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \boldsymbol{\beta}'\mathbf{x}_i)^2\right) = \sigma^{-1} \phi\left(\frac{y_i - \boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)$$

$\phi(\cdot)$  merupakan fungsi probabilitas distribusi  $N(0,1)$ , sedangkan  $\Phi(\cdot)$  merupakan fungsi kepadatan probabilitas kumulatif distribusi  $N(0,1)$ . Fungsi ln-likelihood dapat ditulis sebagai berikut.

$$\ln L = \sum_{y_i > 0} \ln\left(\sigma^{-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)^2\right)\right) + \sum_{y_i=0} \ln\left(1 - \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)\right) \quad (2.15)$$

$$\ln L = \sum_{y_i > 0} \left(-\ln\sigma + \ln \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)^2\right)\right) + \sum_{y_i=0} \ln\left(1 - \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)\right)$$

$$\ln L = \sum_{y_i > 0} \left(-\ln\sigma - \frac{1}{2}\left(\frac{y_i - \boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)^2 + \left(-\frac{1}{2}\right) \ln(2\pi)\right)$$

$$+ \sum_{y_i=0} \ln\left(1 - \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)\right)$$

$$\ln L = \sum_{y_i > 0} \left(-\ln\sigma - \frac{1}{2} \ln(2\pi) - \frac{1}{2\sigma^2}(y_i - \boldsymbol{\beta}'\mathbf{x}_i)^2\right)$$

$$+ \sum_{y_i=0} \ln\left(1 - \Phi\left(\frac{\boldsymbol{\beta}'\mathbf{x}_i}{\sigma}\right)\right) \quad (2.16)$$



Selanjutnya dari persamaan 2.16 tersebut diturunkan terhadap  $\beta$  dan disama dengankan nol untuk memaksimumkan likelihood.

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta} &= \sum_{y_i > 0} \left( 0 + 0 - \frac{1}{2\sigma^2} (2)(-x_i)(y_i - \beta'x_i) \right) \\ &\quad + \sum_{y_i=0} \left( f\left(\frac{\beta'x_i}{\sigma}\right) \left( -\frac{1}{1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right)} \right) \left( \frac{x_i}{\sigma} \right) \right) = 0 \\ \frac{\partial \ln L}{\partial \beta} &= \sum_{y_i > 0} \left( \frac{(y_i - \beta'x_i)(x_i)}{\sigma^2} \right) \\ &\quad + \sum_{y_i=0} \left( \left( f\left(\frac{\beta'x_i}{\sigma}\right) \right) \left( -\frac{1}{1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right)} \right) \left( \frac{x_i}{\sigma} \right) \right) = 0 \end{aligned} \quad (2.17)$$

Persamaan 2.16 diturunkan terhadap  $\sigma^2$

$$\begin{aligned} \ln L &= \sum_{y_i > 0} \left( -\ln(\sigma^2)^{\frac{1}{2}} - \frac{1}{2} \ln(2\pi) - \frac{1}{2\sigma^2} (y_i - \beta'x_i)^2 \right) \\ &\quad + \sum_{y_i=0} \ln \left( 1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right) \right) \\ \ln L &= \sum_{y_i > 0} \left( -\frac{1}{2} \ln(\sigma^2) - \frac{1}{2} \ln(2\pi) - \frac{1}{2\sigma^2} (y_i - \beta'x_i)^2 \right) \\ &\quad + \sum_{y_i=0} \ln \left( 1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right) \right) \\ \frac{\partial \ln L}{\partial \sigma^2} &= \sum_{y_i > 0} \left( \left( -\frac{1}{2} \right) \left( \frac{1}{\sigma^2} \right) + 0 - \frac{1}{2} \sigma^{-2} (y_i - \beta'x_i)^2 \right) \\ &\quad + \sum_{y_i=0} \left( \left( -\frac{1}{1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right)} \right) \left( -\frac{\beta'x_i}{2\sigma^3} \right) \left( f\left(\frac{\beta'x_i}{\sigma}\right) \right) \right) = 0 \\ \frac{\partial \ln L}{\partial \sigma^2} &= \sum_{y_i > 0} \left( \left( -\frac{1}{2} \right) \left( \frac{1}{\sigma^2} \right) - \left( \frac{1}{2} \right) (-2) \left( \frac{1}{\sigma^3} \right) (y_i - \beta'x_i)^2 \right) \\ &\quad + \sum_{y_i=0} \left( \left( -\frac{1}{1 - \Phi\left(\frac{\beta'x_i}{\sigma}\right)} \right) \left( -\frac{\beta'x_i}{2\sigma^3} \right) \left( f\left(\frac{\beta'x_i}{\sigma}\right) \right) \right) = 0 \end{aligned}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = \sum_{y_i > 0} \left( \left( -\frac{1}{2\sigma^2} \right) + \frac{(y_i - \beta' x_i)^2}{\sigma^3} \right) + \sum_{y_i = 0} \left( \left( -\frac{1}{1 - \Phi\left(\frac{\beta' x_i}{\sigma}\right)} \right) \left( -\frac{\beta' x_i}{2\sigma^3} \right) \left( f\left(\frac{\beta' x_i}{\sigma}\right) \right) \right) = 0 \quad (2.18)$$

Kemudian persamaan 2.17 dikalikan dengan  $\beta'/2\sigma^2$  dan dijumlahkan hasil dari persamaan 2.18, menghasilkan persamaan berikut.

$$\sigma^2 = \frac{1}{R} \sum_{i=1}^R (Y_i - \beta' X_i) Y_i = \frac{Y'(Y - X\beta)}{R}, \text{ kalikan dengan } \sigma - \bar{X}'\bar{\gamma} + \frac{1}{\sigma} X'(Y - X\beta) = 0, \text{ dimana } \bar{\gamma} \text{ adalah vector } 1 \times S. \text{ Diperoleh nilai } \beta \text{ sebagai berikut.}$$

$$\hat{\beta} = (X'X)^{-1}X'Y - \sigma(X'X)^{-1}\bar{X}'\bar{\gamma} = \beta_R^{LS} - \sigma(X'X)^{-1}\bar{X}'\bar{\gamma} \quad (2.19)$$

Dengan  $Y'$  : vektor  $1 \times R$

$X'$  : vektor  $1 \times R$

$\bar{X} = (X_{R+1}, X_{R+2}, \dots, X_1)$

$\bar{\gamma} = \frac{\phi}{1-\phi}$  dimana  $\phi$  adalah pdf dari distribusi normal standar.

$\bar{Y}' = (Y_{R+1}, Y_{R+2}, \dots, Y_1)$

$\sigma^2 = \frac{Y'(Y - X\beta)}{R}$

$\beta_R^{LS}$  : penduga OLS pada pengamatan yang tidak sama dengan nol.

Penduga parameter seperti pada persamaan (2.19) merupakan penduga yang menunjukkan adanya bias antara regresi OLS dengan regresi *Tobit*. Dalam penerapannya, penduga parameter regresi *Tobit* menghasilkan sebuah persamaan non linier seperti yang dijelaskan dalam Fair (1977). Untuk menyelesaikan persamaan non linier tersebut digunakan salah satu metode iterasi, yaitu iterasi *Newton Raphson*. Prinsip dari metode iterasi ini adalah menentukan nilai parameter secara berulang-ulang dengan member nilai awal tertentu sampai dicapai nilai yang konvergen.

## 2.5 Pengujian Pendugaan Parameter

Untuk menguji taksiran parameter, statistik uji yang biasa digunakan adalah uji *Wald*, *Likelihood Ratio* (LR). Uji *Wald* dan LR test sering digunakan untuk pengujian taksiran parameter dalam model *Tobit* (Robinson, Bera and Jarque, 1985).

Langkah-langkah dalam uji parameter regresi adalah sebagai berikut.

### 1. Uji Serentak

Uji serentak digunakan untuk menguji parameter secara keseluruhan atau bersama-sama. Pengujian menggunakan metode *likelihood ratio* atau uji *G*. Misalkan  $Y_1, Y_2, \dots, Y_I$  adalah variabel random yang saling bebas sebanyak  $I$ , yang masing-masing mempunyai fungsi sebaran probabilitas  $f(Y_i; \beta_1, \beta_2, \dots, \beta_p)$  untuk  $i = 1, 2, \dots, I$ . Himpunan yang terdiri dari semua parameter titik  $(\beta_1, \beta_2, \dots, \beta_p)$  dinotasikan dengan  $\Omega$  dan  $\omega$  subset dari  $\Omega$ .

$$L(\omega) = \prod_{i=1}^I f(y_i; \beta_0) \text{ dengan } \omega = \{\beta_0\}$$

$$L(\Omega) = \prod_{i=1}^I f(y_i; \beta_1, \beta_2, \dots, \beta_p), \text{ dengan } \Omega = \{\beta_1, \beta_2, \dots, \beta_p\}$$

Hipotesis yang digunakan adalah sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{Paling tidak terdapat satu } \beta_i \neq 0, i = 1, 2, \dots, p$$

Statistik Uji yang digunakan adalah uji *G* (*Likelihood Ratio Test*) sebagai berikut.

$$G = -2 \ln \left( \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad (2.20)$$

di mana :

$L_{\omega}$  = nilai maksimum *Likelihood* tanpa variabel prediktor tertentu;

$L_{\Omega}$  = nilai maksimum *Likelihood* dengan variabel prediktor tertentu.

Nilai G didapatkan dari rumus sebagai berikut.

$$L(\omega) = L(\beta_0, \sigma^2) = \prod_{y_i > 0} \frac{1}{\sigma} \phi\left(\frac{y_i - \beta_0}{\sigma}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\beta_0}{\sigma}\right)\right)$$

$$L(\Omega) = L(\beta_0, \beta_1, \beta_2, \dots, \beta_p, \sigma^2) = \prod_{y_i > 0} \frac{1}{\sigma} \phi\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right)\right)$$

Sehingga rasio antara  $L(\hat{\Omega})$  dan  $L(\hat{\omega})$  dapat ditulis sebagai berikut.

$$\begin{aligned} \Lambda &= \frac{L(\hat{\omega})}{L(\hat{\Omega})} = \frac{\prod_{y_i > 0} \frac{1}{\hat{\sigma}_\omega} \phi\left(\frac{y_i - \hat{\beta}_0}{\hat{\sigma}_\omega}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\hat{\beta}_0}{\hat{\sigma}_\omega}\right)\right)}{\prod_{y_i > 0} \frac{1}{\hat{\sigma}} \phi\left(\frac{y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)\right)} \\ &= \frac{\prod_{y_i > 0} \phi\left(\frac{y_i - \hat{\beta}_0}{\hat{\sigma}_\omega}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\hat{\beta}_0}{\hat{\sigma}_\omega}\right)\right)}{\prod_{y_i > 0} \phi\left(\frac{y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)\right)} \end{aligned}$$

Didapatkan nilai G sebagai berikut.

$$\begin{aligned} G &= -2 \ln \left( \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = -2 \ln \left( \frac{\prod_{y_i > 0} \phi\left(\frac{y_i - \hat{\beta}_0}{\hat{\sigma}_\omega}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\hat{\beta}_0}{\hat{\sigma}_\omega}\right)\right)}{\prod_{y_i > 0} \phi\left(\frac{y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) \prod_{y_i = 0} \left(1 - \Phi\left(\frac{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)\right)} \right) \\ &= -2 \left( \sum_{y_i > 0} \phi\left(\frac{y_i - \hat{\beta}_0}{\hat{\sigma}_\omega}\right) + \sum_{y_i = 0} \left(1 - \Phi\left(\frac{\hat{\beta}_0}{\hat{\sigma}_\omega}\right)\right) - \left( \sum_{y_i > 0} \phi\left(\frac{y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) + \sum_{y_i = 0} \left(1 - \Phi\left(\frac{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)\right) \right) \right) \\ &= 2 \sum_{y_i > 0} \left( \phi\left(\frac{y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) - \phi\left(\frac{y_i - \hat{\beta}_0}{\hat{\sigma}_\omega}\right) \right) + 2 \sum_{y_i = 0} \left( \Phi\left(\frac{\hat{\beta}_0}{\hat{\sigma}_\omega}\right) - \Phi\left(\frac{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) \right) \end{aligned}$$

Statistik uji ini mengikuti sebaran *chi-square* dengan derajat bebas  $p$  yaitu banyaknya parameter dalam model. Kesimpulan yang dapat diambil adalah tolak  $H_0$  jika uji  $G > \chi^2_{\alpha, p}$  berarti terdapat minimal satu  $\beta_i$  yang mempunyai peran berarti terhadap model.

## 2. Uji Parsial

Uji ini dilakukan untuk menguji setiap  $\beta_i$  secara individual untuk menunjukkan apakah suatu variabel bebas layak untuk masuk dalam model. Hipotesa yang digunakan adalah sebagai berikut.

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

Statistik uji yang digunakan adalah statistik uji *Wald*, yaitu sebagai berikut.

$$w^2 = \frac{\hat{\beta}_i^2}{\text{se}(\hat{\beta}_i)^2} \quad (2.21)$$

Dengan  $\text{Se}(\hat{\beta}_i)^2 = \text{varian}$

$\hat{\beta}_i$  = nilai koefisien dugaan variabel prediktor

$w^2$  mengikuti sebaran  $\chi^2$  dengan keputusan yang dapat diambil adalah tolak  $H_0$  jika  $w^2 > \chi^2_{\alpha}$  yang dapat disimpulkan  $\beta_i$  mempunyai peran berarti terhadap model. Pada dasarnya statistik uji *Wald* ini sama dengan statistik uji  $t$ .

## 2.6 Asumsi-asumsi Analisis Regresi *Tobit*

### 2.6.1 Asumsi Kenormalan Galat

Uji Kolmogorov Smirnov merupakan suatu tes *goodness of fit* yang berdasarkan uji Chi-Square dan mampu memberi pendekatan nilai eksak dengan nilai maksimum = 1 dan nilai minimum = 0. Uji ini menggunakan sebaran kumulatif contoh dan sebaran kumulatif sebaran normal.

$$\text{Hipotesis: } H_0 : F(y) = F_0(y) \quad \text{vs} \quad H_1 : F(y) \neq F_0(y)$$

Statistik uji ini adalah jarak tegak max (nilai max) antara dua fungsi sebaran sebagai ukuran jarak antara kedua fungsi sebaran atau disebut juga  $D_n$ .

$$D_n = \text{Maks} [F_n(y) - F_0(y)] \quad (2.22)$$

dimana:

$D_n$  = jarak tegak maksimum antara fungsi sebaran empiris dan fungsi sebaran normal

$F_n(y)$  = sebaran kumulatif contoh

$F_0(y)$  = sebaran kumulatif sebaran normal



Tabel 2.1 Nilai kritis  $D_n \alpha$

A	0.01	0.05	0.1
$D_n \alpha$	$1.63/\sqrt{n}$	$1.36/\sqrt{n}$	$1.22/\sqrt{n}$

Kriteria keputusan berdsarkan uji ini adalah tolak  $H_0$  jika  $D_n > D_n \alpha$ , begitu pula sebaliknya. Uji ini memberikan informasi adanya ketidaksamaan model (*lack of fit*) bila menolak  $H_0$ . Selain itu pengambilan keputusan juga dapat dilakukan berdasarkan nilai  $p$ . Pengambilan keputusan berdasarkan perbandingan antara nilai  $p$  dengan  $\alpha$  yaitu jika nilai  $p > \alpha$  maka terima  $H_0$  begitu pula sebaliknya.

### 2.6.2 Asumsi Non-Multikolinearitas (kebebasan antar variabel prediktor)

Multikolinearitas adalah adanya hubungan linier antara variabel bebas dalam model regresi. Untuk regresi  $k$  variabel, meliputi variabel bebas  $X_1, X_2, \dots, X_k$  (dimana  $X_1=1$  untuk semua pengamatan untuk memungkinkan unsur intersep), suatu hubungan linier yang pasti dikatakan ada apabila memenuhi kondisi sebagai berikut:

$$\lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_k X_k = 0 \quad (2.23)$$

di mana  $\lambda_1, \lambda_2, \dots, \lambda_k$  adalah konstanta sedemikian sehingga tidak semuanya secara simultan sama dengan nol.

Multikolinieritas sempurna: satu variabel prediktor adalah fungsi linier dari variabel prediktor yang lain

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

$$X_3 = \delta_1 + \delta_2 X_2$$

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 (\delta_1 + \delta_2 X_2) + u$$

$$Y = (\beta_1 + \beta_3 \delta_1) + (\beta_2 + \beta_3 \delta_2) X_2 + u$$

$$Y = v_1 + v_2 X_2 + u$$

Sampel dipakai untuk menduga koefisien  $v_1$  dan  $v_2$ . Penduga  $\beta_1$  dan  $\beta_2$  didapat dari solusi persamaan :

$$\hat{v}_1 = \hat{\beta}_1 + \hat{\beta}_3 \delta_1$$

$$\hat{v}_2 = \hat{\beta}_2 + \hat{\beta}_3 \delta_2$$

Tidak ada solusi unik bagi penduga parameter populasi karena 2 persamaan untuk 3 peubah.

Multikolinieritas tak sempurna terjadi jika terdapat hubungan linier yang tidak sempurna antar variabel prediktor. Efek dari multikolinieritas tak sempurna :

- Penduga MLE tetap dapat diduga
  - Penduga MLE tetap bersifat BLUE
  - Ragam dan peragam dari penduga MLE relatif besar
  - Selang kepercayaan menjadi lebih besar (lebih banyak menerima hipotesis nol (koefisien tidak nyata))
  - Statistik uji  $w$  dari satu atau beberapa koefisien menjadi tidak nyata (walaupun  $R^2$  secara keseluruhan besar)
  - Tanda bagi penduga koefisien berkebalikan dengan teori a priori
- Mendeteksi multikolinieritas dapat dilakukan dengan melihat nilai VIF (*Variance Inflation Factor*)

$$VIF = \frac{1}{1-r_{23}^2} \quad (2.24)$$

$$r_{23}^2 = \frac{(\sum x_{2i}x_{3i})^2}{(\sum x_{2i}^2)(\sum x_{3i}^2)}$$

Untuk mencari  $var(\hat{\beta}_2)$  dan  $var(\hat{\beta}_3)$  yaitu :

$$var(\hat{\beta}_2) = \frac{\sigma^2}{(\sum x_{2i}^2)(1-r_{23}^2)} \quad var(\hat{\beta}_3) = \frac{\sigma^2}{(\sum x_{3i}^2)(1-r_{23}^2)}$$

$$var(\hat{\beta}_2) = \frac{\sigma^2}{(\sum x_{2i}^2)} VIF \quad var(\hat{\beta}_3) = \frac{\sigma^2}{(\sum x_{3i}^2)} VIF$$

Semakin besar multikolinieritas maka semakin besar VIF, semakin besar VIF maka semakin besar ragam penduga MLE.

Untuk regresi lebih dari 2 variabel definisi dari VIF :

$$VIF_j = \frac{1}{1-R_j^2} \quad (2.25)$$

Dimana  $R_j^2$  : koefisien determinasi dari *auxiliary regression*

*Auxiliary regression* merupakan regresi dengan  $X_j$  sebagai variabel respon, dan  $X$  selainnya sebagai variabel prediktor.

Digunakan kriteria dengan nilai 10 dikarenakan apabila  $R_j^2$  bernilai 0.9 maka VIF yang didapatkan bernilai 10

$$VIF = \frac{1}{1-0.9} = 10$$

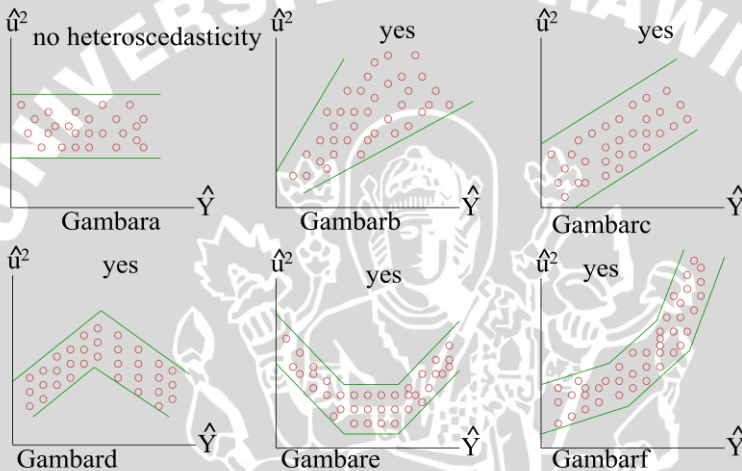
VIF yang lebih dari 10 adalah bukti cukup untuk multikolinieritas. Apabila nilai  $VIF \leq 10$  disimpulkan tidak terjadi multikolinieritas.

### 2.6.3 Asumsi Kesamaan Ragam Galat (Homoskedastisitas)

Salah satu asumsi penting dari model analisis regresi adalah bahwa variabel *disturbance* atau galat  $u_i$  yang muncul adalah homoskedastik. Homoskedastisitas, *scedasticity* (penyebaran) dan *homos* (sama) yaitu ragam yang sama. Artinya, variabel pengganggu memiliki ragam yang sama.

$$E(u_i^2) = \sigma^2, \quad i = 1, 2, \dots, N$$

Untuk mendeteksi asumsi homokedastisitas ini dapat dilakukan dengan melihat plot residual.



Gambar 2.2 Pola hipotesis residual yang ditaksir (Gujarati, 1991).

Dari Gambar 2.2 dapat ditunjukkan bahwa pada gambar a tidak terdapat sifat heterokedastisitas karena tidak terdapat pola tertentu antara dua variabel tersebut. Sedangkan gambar b sampai gambar f menunjukkan pola tertentu sehingga terdapat sifat heterokedastisitas.

Selain dengan melihat plot residual, untuk mendeteksi asumsi homoskedastisitas dapat dilakukan dengan uji Breusch-Pagan LM test. Langkah-langkah uji Breusch-Pagan LM adalah :

1. Menduga regresi Tobit dan menduga residual  $u_i$

$$\hat{u}_i = \hat{Y}_i - Y_i$$

2. Menduga *auxiliary regression*, di mana variabel prediktor yang digunakan adalah variabel-variabel yang mungkin mempengaruhi ragam galat

Variabel prediktor ( $X$ )

$$\hat{u}_i^2 = a_1 + a_2 X_{2i} + \dots + a_p X_{pi} + v_i$$

3. Hipotesis  
 $H_0$  : Tidak ada hubungan antara variabel  $X$  dengan residual (homokedastisitas)  
 $H_1$  : Terdapat hubungan antara variabel  $X$  dengan residual (heterokedastisitas)
4. Didapatkan statistik uji berdasarkan koefisien determinasi dari *auxiliary regression*  $R^2$   
 $LM = nR^2 \sim \chi_{p-1}^2$   
 di mana derajat bebas adalah jumlah  $X$  yang digunakan di dalam *auxiliary regression*
5. Tolak  $H_0$  jika ada bukti yang nyata dari statistik uji (Gujarati,1991).

#### 2.6.4 Asumsi Non Autokorelasi (Kebebasan galat)

Satu dari asumsi penting dari model regresi *Tobit* adalah bahwa kesalahan atau gangguan  $u_i$  yang masuk kedalam fungsi regresi populasi adalah random atau tak berkorelasi. Jika asumsi ini dilanggar, maka terdapat problem autokorelasi.

Istilah autokorelasi dapat didefinisikan sebagai “korelasi antara anggota serangkaian observasi yang diurutkan menurut waktu (seperti dalam data deret waktu) atau ruang (seperti dalam data cross-sectional) atau juga dapat dikatakan korelasi antara 2 deretan waktu seperti  $u_1, u_2, \dots, u_{10}$  dan  $u_2, u_3, \dots, u_{11}$ , di mana yang pertama adalah seri yang terakhir tetapi ketinggalan dengan satu periode waktu. Dalam konteks regresi mengasumsikan bahwa autokorelasi seperti itu tidak terdapat dalam gangguan  $u_i$ . Dengan menggunakan lambang

$$E(u_i u_j) = 0 \quad i \neq j$$

Secara sederhana dapat dikatakan model regresi mengasumsikan bahwa unsur gangguan yang berhubungan dengan observasi tidak dipengaruhi oleh unsur gangguan yang berhubungan dengan pengamatan lain yang manapun.

Mendeteksi autokorelasi dilakukan dengan Uji *Durbin-Watson*.

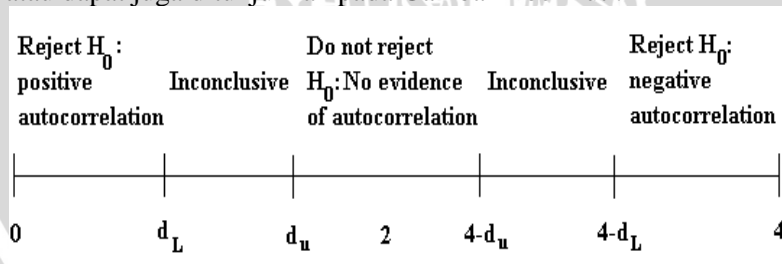
Statistik  $d$  dari *Durbin-Watson* adalah

$$d = \frac{\sum_{t=2}^{t=N} (e_t - e_{t-1})^2}{\sum_{t=1}^{t=N} e_t^2} \quad (2.26)$$

Keuntungan besar dari statistik  $d$  adalah bahwa statistik tadi didasarkan pada residual yang ditaksir, yang secara rutin dihitung dalam analisis regresi.

Mekanisme tes Durbin-Watson adalah sebagai berikut:

1. Menduga regresi Tobit dan menduga residual  $e_i$ .
  2. Hipotesis
    - $H_0$  : Tidak ada autokorelasi positif atau negatif
    - $H_1$  : Terdapat autokorelasi positif atau negatif
  3. Hitung  $d$ .
  4. Untuk ukuran sampel tertentu dan banyaknya variabel yang menjelaskan tertentu, dapatkan nilai kritis  $d_L$  dan  $d_U$ .
  5. Jika Hipotesis  $H_0$  adalah bahwa tidak ada serial korelasi positif, jika
    - $d < d_L$  : menolak  $H_0$
    - $d > d_U$  : tidak menolak  $H_0$
    - $d_L \leq d \leq d_U$  : pengujian tidak meyakinkan
  6. Jika Hipotesis  $H_0$  adalah bahwa tidak ada serial korelasi negatif, jika
    - $d < 4 - d_L$  : menolak  $H_0$
    - $d > 4 - d_U$  : tidak menolak  $H_0$
    - $4 - d_L \leq d \leq 4 - d_U$  : pengujian tidak meyakinkan
  7. Jika  $H_0$  adalah ujung-ujung, yaitu bahwa tidak ada serial korelasi positif atau negatif, maka jika
    - $d < d_L$  : menolak  $H_0$
    - $d > 4 - d_L$  : menolak  $H_0$
    - $d_u < d < 4 - d_u$  : tidak menolak  $H_0$
    - $d_L \leq d \leq d_u$  atau  $4 - d_u \leq d \leq 4 - d_L$  : pengujian tidak meyakinkan
- atau dapat juga ditunjukkan pada Gambar 2 berikut :



Gambar 2.3 Statistik  $d$  Durbin-Watson (Gujarati, 1991).

## 2.7 Penentuan Kebaikan Model

Penentuan kebaikan model dalam analisis regresi dapat menggunakan *Akaike's Information Criterion (AIC)*. *Akaike's Information Criterion (AIC)* pertama kali diperkenalkan oleh



Akaike, seorang statistikawan asal Jepang pada tahun 1973. Metode ini merupakan salah satu metode yang dapat digunakan untuk memilih model regresi terbaik didasarkan pada metode *Maximum Likelihood Estimation* (MLE). Semakin kecil nilai AIC maka semakin baik model regresi yang digunakan.

Untuk menghitung nilai AIC digunakan persamaan sebagai berikut:

$$AIC = -2 \log (\text{maximum likelihood}) + 2 (\text{number of parameters}) \quad (2.27)$$

## 2.8 Tinjauan Non Statistik

Menurut Laventhal dan Cleary (2000) faktor yang mempengaruhi konsumsi rokok terbagi dalam dua bagian:

### 1. Faktor Psikologis

Pada umumnya terbagi dalam lima bagian yaitu:

#### a. Kebiasaan

Perilaku merokok adalah sebuah perilaku yang harus tetap dilakukan tanpa adanya motif yang bersifat positif ataupun negatif. Seseorang merokok hanya untuk meneruskan perilakunya tanpa tujuan tertentu.

#### b. Reaksi emosi yang positif

Kebiasaan merokok digunakan untuk menghasilkan reaksi yang positif, misalnya rasa senang, relaksasi dan kenikmatan rasa. Merokok juga dapat menunjukkan kejantanan (kebanggaan diri) dan menunjukkan kedewasaan.

#### c. Reaksi untuk penurunan emosi

Merokok ditunjukkan untuk mengurangi rasa tegang, kecemasan biasa ataupun kecemasan yang ditimbulkan karena adanya interaksi dengan orang lain.

#### d. Alasan sosial

Merokok ditunjukkan untuk mengikuti kebiasaan merokok, identifikasi perokok lain, dan menentukan *image* diri seseorang.

#### e. Kecanduan dan ketagihan

Seseorang merokok telah mengalami kecanduan karena kandungan *nikotin* dalam rokok sehingga akan menimbulkan ketagihan.

## 2. Faktor Biologis

Faktor ini menekankan pada kandungan *nikotin* yang ada di dalam rokok yang dapat mempengaruhi ketergantungan seseorang pada rokok secara biologis.

Selain itu faktor yang mempengaruhi konsumsi rokok biasanya dari faktor lingkungan. Faktor lingkungan bisa saja dari faktor keluarga, tempat tinggal atau bahkan lingkungan pergaulan. Remaja cenderung merokok karena memiliki teman-teman atau keluarga yang merokok. Secara umum perilaku merokok merupakan fungsi dari lingkungan dan individu. Artinya perilaku merokok selain disebabkan oleh faktor-faktor dari dalam diri, juga disebabkan faktor lingkungan.



UNIVERSITAS BRAWIJAYA



## BAB III METODE PENELITIAN

### 3.1 Sumber Data

Data yang akan digunakan dalam penelitian ini adalah data hasil survei Sosial Ekonomi Nasional (Susenas) Kota Kediri yang dilaksanakan Badan Pusat Statistik (BPS) tahun 2011. Unit pengamatan pada penelitian ini adalah Rumah Tangga (RT), di mana sampel yang diamati adalah 217 RT.

### 3.2 Variabel Penelitian

Penelitian ini mengamati satu variabel respon dan beberapa variabel prediktor.

#### 3.2.1 Variabel Respon

Variabel respon (Y) yang diamati dari hasil Susenas adalah pengeluaran rumah tangga untuk konsumsi rokok selama satu bulan dalam rupiah.

$$Y_i = \begin{cases} 0 & , \text{tidak mengeluarkan biaya konsumsi rokok (Rp)} \\ y_i^* & , \text{mengeluarkan biaya konsumsi rokok (Rp)} \end{cases}$$

#### 3.2.2 Variabel Prediktor

Variabel prediktor yang digunakan dalam penelitian ini adalah:

Umur = Umur kepala rumah tangga (tahun).

PART = Proporsi anggota rumah tangga dewasa dalam rumah tangga (persen). Yaitu banyaknya anggota rumah tangga yang berusia  $\geq 20$  tahun dibagi seluruh anggota rumah tangga dikali seratus.

PDT = Pendapatan rumah tangga dengan pendekatan pengeluaran (rupiah). Dalam hal ini didekati oleh data pengeluaran rumah tangga selama satu bulan.

HBK = Harga Barang Komplementer

Harga barang komplementer adalah harga barang yang melengkapi utilitas seseorang dalam mengkonsumsi rokok, yaitu korek api. Dalam hal ini harga barang komplementer diukur dengan satuan rupiah (Rp).

### 3.3 Metode Analisis

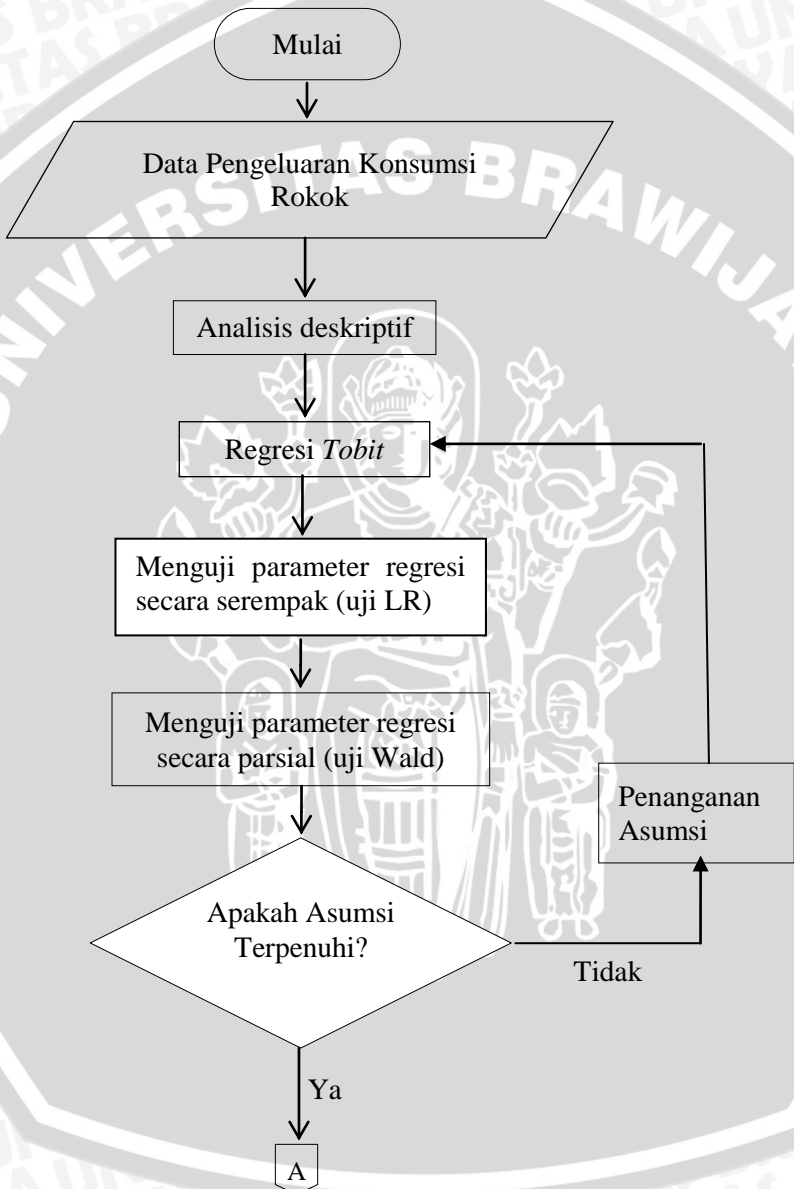
Metode analisis yang digunakan dalam penelitian ini ada dua yakni analisis statistik deskriptif dan juga analisis regresi *Tobit*.

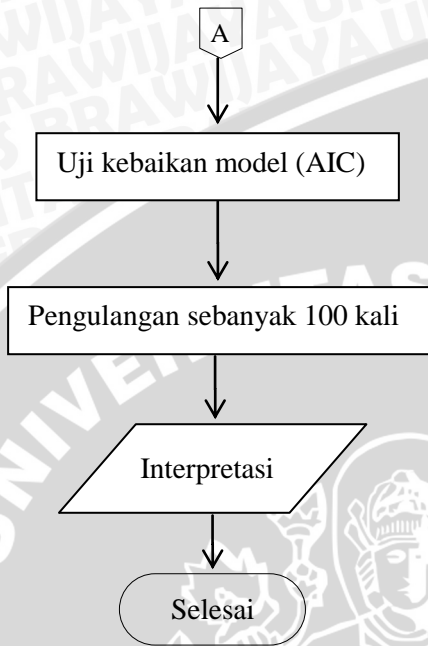
1. Tahap pengkajian terhadap karakteristik pengeluaran biaya konsumsi rokok rumah tangga kota Kediri dengan menghitung ukuran penyebaran dan pemusatan data pada variabel respon dan variabel prediktor
2. Tahap analisis pola hubungan pengeluaran biaya konsumsi rokok rumah tangga kota Kediri dengan faktor-faktor yang mempengaruhi.
  - a. Mergresikan variabel  $y$  terhadap semua variabel prediktor hingga didapatkan persamaan  $y_i^* = x_i\beta + u_i$ .
  - b. Mencari nilai penduga parameter dengan metode *MLE* dari model yang didapat pada langkah a dengan rumus (2.19).
  - c. Pengujian terhadap penduga parameter yang telah didapat pada langkah b menggunakan *LR test* dengan rumus (2.20) untuk keseluruhan model dan uji *wald* dengan rumus (2.21) untuk menguji secara individu setiap penduga parameter.
  - d. Dilakukan pengujian asumsi
  - e. Dilakukan uji kebaikan model dengan menghitung nilai *Akaike's Information Criterion* (AIC) model dengan rumus (2.27)
  - f. Langkah terakhir dilakukan resampling terhadap data. Diambil 30 ukuran sampel secara acak dan diulang sebanyak 100 kali kemudian digunakan untuk melakukan perbandingan antara OLS dan *Tobit* untuk melihat kebaikan regresi *Tobit*



### 3.4 Diagram Alir

Berikut ini merupakan diagram alir Analisis *Regresi Tobit* :



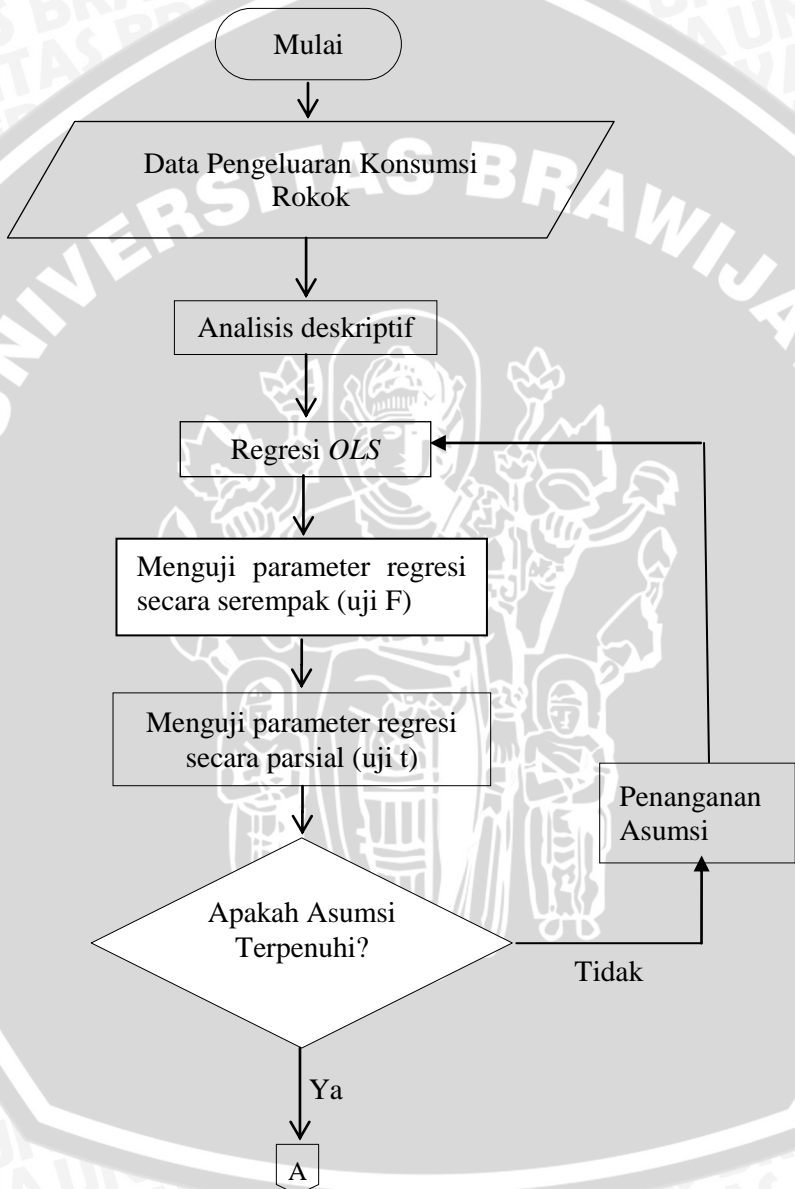


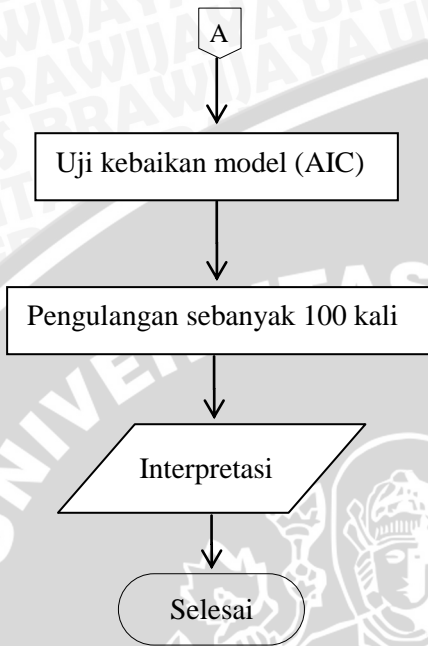
Gambar 3.1. Diagram Alir Analisis Regresi *Tobit*

Diagram alir dari metode penelitian disajikan pada Gambar Analisis data dilakukan dengan menggunakan bantuan software statistika, yaitu: SPSS 20, Minitab dan Gretl.

### 3.5 Diagram Alir

Berikut ini merupakan diagram alir Analisis *Regresi OLS* :





Gambar 3.1. Diagram Alir Analisis Regresi *OLS*

Diagram alir dari metode penelitian disajikan pada Gambar Analisis data dilakukan dengan menggunakan bantuan software statistika, yaitu: SPSS 20, Minitab dan Gretl.

## BAB IV HASIL DAN PEMBAHASAN

### 4.1 Analisis Deskriptif Pengeluaran Rumah Tangga untuk Biaya Konsumsi Rokok

Analisis deskriptif ini dilakukan untuk mengetahui gambaran mengenai karakteristik variabel-variabel yang mempengaruhi biaya konsumsi rokok rumah tangga kota Kediri. Analisis deskriptif dapat dilihat dari nilai rata-rata, varians, nilai maksimum, dan nilai minimum dari setiap variabel. Pengeluaran rumah tangga untuk konsumsi rokok memiliki berbeda-beda untuk setiap rumah tangga. Bahkan ada beberapa rumah tangga yang tidak memiliki pengeluaran untuk konsumsi rokok tersebut. Hal ini menyebabkan besarnya variasi pada pengeluaran ini.

Pada Tabel 4.1 dapat dilihat nilai rata-rata, varians, nilai maksimum, dan nilai minum dari variabel pengeluaran rumah tangga untuk konsumsi rokok. Nilai rata-rata pengeluaran rumah tangga untuk konsumsi rokok sebesar 126287.07 rupiah. Terlihat pula bahwa range yang terbentuk dari nilai minimum dan nilai maksimum sangat jauh. Deskripsi yang lebih lengkap disajikan pada Tabel 4.1.

**Tabel 4.1** Deskripsi Pengeluaran Rumah Tangga untuk Konsumsi Rokok

Deskripsi	Nilai
N	217
Mean	Rp 126.287,07
Varian	Rp 6.401.055.516,865
Min	0
Max	Rp 315.000

#### 4.1.1 Deskripsi Variabel-variabel yang Mempengaruhi Pengeluaran Rumah Tangga untuk Biaya Konsumsi Rokok

Banyak faktor yang mempengaruhi rumah tangga untuk biaya konsumsi rokok. Berikut diberikan deskripsi mengenai variabel prediktor dalam penelitian ini, yaitu umur kepala rumah tangga (tahun), proporsi anggota rumah tangga dewasa dalam rumah tangga (persen), pendapatan rumah tangga (Rp), harga barang komplementer (Rp).

Deskripsi variabel prediktor yang mempengaruhi pengeluaran untuk konsumsi rokok dideskripsikan berdasarkan nilai rata-rata,



varians, nilai maksimum, serta nilai minimum. Tabel 4.2 memberikan deskripsi keempat variabel prediktor yang berskala kontinu secara lengkap.

**Tabel 4.2** Deskripsi Variabel Prediktor

Var	Mean	Varian	Min	Max
Umur	49.79 th	182.489 th	17 th	83 th
PART	64.7221%	632.880%	0%	100%
PDT	Rp2.221.402,60	Rp937.656.205.345,112	Rp574.417	Rp7.095.900
HBK	Rp 2.241,94	Rp 854.390,681	0	Rp 5.000

Variabel prediktor yang dianggap mempengaruhi pembelian rokok adalah umur kepala rumah tangga. Umur juga dianggap memberikan pengaruh terhadap pengeluaran untuk pembelian rokok. Hal ini disebabkan oleh kedewasaan berpikir dan kedewasaan dalam menentukan pilihan apakah barang itu penting atau tidak dan menyehatkan atau tidak.

Berdasarkan Tabel 4.2 diketahui bahwa dari 217 rumah tangga yang menjadi sampel penelitian, tiap rumah tangga rata-rata memiliki kepala rumah tangga berumur 49 tahun. Umur dalam anggota rumah tangga yang paling muda adalah 17 tahun dan yang paling tua adalah 83 tahun. Umur kepala rumah tangga memiliki varians yang sangat besar yaitu 183. Dari nilai varians ini dapat dilihat bahwa tiap rumah tangga sampel memiliki umur yang berbeda-beda untuk kepala rumah tangga.

Variabel selanjutnya dijelaskan bahwa rata-rata proporsi anggota rumah tangga dewasa dalam tiap rumah tangga sebesar 64.7221%. Selain itu, proporsi terkecil anggota rumah tangga dewasa dalam tiap rumah tangga adalah sebesar 0%. Sedangkan proporsi terbesar anggota rumah tangga dewasa dalam tiap rumah tangga adalah sebesar 100%. Proporsi anggota rumah tangga dewasa dalam tiap rumah tangga memiliki varians yang besar yaitu 632.880%. Dari nilai varians ini dapat dilihat bahwa tiap rumah tangga sampel memiliki proporsi bervariasi.

Pendapatan rumah tangga dianggap memberikan pengaruh untuk menentukan pembelian konsumsi rokok. Variabel ini dihitung dalam satuan rupiah. Pada Tabel 4.2 diketahui bahwa dari 217 rumah tangga yang menjadi sampel penelitian, memiliki rata-rata untuk pendapatan sebesar Rp 2.221.402,60. Selain itu, pendapatan terkecil yang diterima oleh rumah tangga adalah sebesar Rp

574.417. Sedangkan pendapatan terbesar yang diterima oleh rumah tangga adalah sebesar Rp 7.095.900. Pendapatan rumah tangga memiliki varians yang sangat besar yaitu Rp 937.656.205.345,112. Dari nilai varians ini dapat dilihat bahwa tiap rumah tangga sampel memiliki pendapatan yang bervariasi.

Pengeluaran rumah tangga untuk barang pelengkap (komplementer) memiliki rata-rata sebesar Rp 2.241,94. Selain itu, pengeluaran untuk barang pelengkap (komplementer) yang dikeluarkan oleh rumah tangga paling kecil adalah sebesar Rp 0. Sedangkan pengeluaran terbesar untuk barang pelengkap (komplementer) yang dikeluarkan oleh rumah tangga adalah sebesar Rp 5.000. Nilai paling minimum atau nilai nol didapatkan untuk pengeluaran barang pelengkap (komplementer) disebabkan oleh tidak semua rumah tangga yang mengeluarkan biaya untuk konsumsi rokok membeli barang pelengkap selain rokok. Dapat juga disebabkan adanya rumah tangga yang tidak mengeluarkan biaya untuk konsumsi rokok sehingga tidak perlu mengeluarkan biaya untuk membeli barang pelengkap. Pengeluaran rumah tangga untuk barang pelengkap memiliki varians yang sangat besar yaitu Rp 854.390,681. Dari nilai varians ini dapat dilihat bahwa tiap rumah tangga sampel memiliki pengeluaran yang nilainya bervariasi untuk barang pelengkap.

#### **4.2 Pemodelan Pengeluaran Rumah Tangga untuk biaya Konsumsi Rokok**

Pemodelan pengeluaran biaya konsumsi rumah tangga kota Kediri dilakukan dalam dua tahap, yakni pengujian secara individu dan secara serentak. Pada tahap pengujian secara individu dilihat variabel apa saja yang berpengaruh signifikan terhadap model. Jika variabel tersebut berpengaruh signifikan secara individu terhadap model, maka dilanjutkan tahap pengujian serentak. Langkah pertama yang dilakukan adalah memasukkan semua variabel prediktor dalam model. Kemudian dicari penduga yang tidak signifikan, berdasarkan nilai  $p$  yang besar atau lebih dari  $\alpha$ . Nilai  $p$  yang besar atau lebih dari  $\alpha$  menunjukkan bahwa variabel tersebut tidak mempunyai pengaruh yang nyata terhadap model. Oleh karena itu variabel tersebut dapat dihilangkan. Selanjutnya dilakukan pengulangan hingga semua variabel berpengaruh signifikan. Hasil akhir dari pengujian serentak tersebut yang dijadikan sebagai model regresi *Tobit* pengeluaran biaya konsumsi rokok rumah tangga kota Kediri.

#### 4.2.1 Pengujian Secara Serentak

Untuk mengetahui model regresi *Tobit* dari pengeluaran rumah tangga untuk konsumsi rokok dilakukan pengujian secara serentak. Pengujian secara serentak ini dilakukan dengan memasukkan semua variabel prediktor ke dalam model regresi *Tobit*. Pengujian hipotesis untuk pengujian serentak sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_4 = 0$$

$$H_1: \text{Paling tidak ada satu } \beta_j \neq 0, j = 1, 2, \dots, 4$$

Pengujian serentak ini menggunakan statistik uji G. Hasil dari pengujian serentak diberikan pada Tabel 4.3 untuk *Tobit* dan 4.4 untuk OLS.

**Tabel 4.3** Hasil Akhir Pengujian Serentak Penduga Koefisien Variabel Prediktor Tobit

Variabel	Nilai Duga Koefisien	Standard Error	P-value
Intersep	-83.967,7	25.300,2	0.00090***
Umur	-679,642	335,48	0.04278**
PART	789,493	177,613	<0.00001***
PDT	0,0167134	0,00518852	0.00128***
HBK	65,5285	7,20418	<0.00001***
Akaike criterion		4785.390	

**Tabel 4.4** Hasil Akhir Pengujian Serentak Penduga Koefisien Variabel Prediktor OLS

Variabel	Nilai Penduga	Standard Error	P-value
Intersep	-29147.8	20.505,8	0.15666
Umur	-582.485	294,257	0.04905**
PART	721.717	161,688	0.00001***
PDT	0.0219704	0,00466656	<0.00001***
HBK	39.6633	4,95172	<0.00001***
Akaike criterion		5378.254	

Pada Tabel 4.3 dan 4.4 diketahui bahwa semua penduga koefisien dari variabel prediktor berpengaruh signifikan terhadap model secara serentak. Hasil pengujian tersebut menghasilkan penduga koefisien dari variabel-variabel prediktor yang bertanda positif dan bertanda negatif. Penduga koefisien yang bertanda positif

memiliki arti bahwa variabel tersebut menyebabkan peningkatan terhadap pengeluaran rumah tangga untuk konsumsi rokok. Untuk penduga koefisien yang bertanda negatif memiliki arti bahwa variabel tersebut akan menyebabkan penurunan terhadap pengeluaran rumah tangga untuk konsumsi rokok. Secara matematis, model regresi *Tobit* yang dihasilkan untuk pengeluaran rumah tangga terhadap konsumsi rokok adalah sebagai berikut.

$$\hat{y} = -83967.7 - 679.642Umur + 789.493PART \\ + 0.0167134PDT + 65.5285HBK$$

Sedangkan model regresi OLS yang dihasilkan untuk pengeluaran rumah tangga terhadap konsumsi rokok adalah sebagai berikut.

$$\hat{y} = -29147.8 - 582.485Umur + 721.717PART \\ + 0.0219704PDT + 39.6633HBK$$

dengan

$\hat{y}$  : penduga pengeluaran konsumsi rokok

Umur : umur kepala rumah tangga (tahun)

PART : proporsi anggota rumah tangga dewasa dalam rumah tangga (persen)

PDT : pendapatan rumah tangga (Rp)

HBK : harga barang komplementer (Rp)

Hasil pemodelan yang telah didapatkan diperoleh kesimpulan bahwa model *Tobit* memiliki nilai *Akaike criterion* lebih kecil daripada model OLS. Model yang didapatkan diinterpretasikan dalam sub bab ini. Dari model *Tobit* dan OLS yang diperoleh, didapatkan variabel-variabel dengan penduga koefisien yang bertanda negatif dan bertanda positif. Variabel dengan penduga koefisien yang bertanda negatif yaitu umur kepala rumah tangga. Hal ini menunjukkan bahwa variabel tersebut memberikan penurunan terhadap pengeluaran rumah tangga untuk biaya konsumsi rokok. Hal ini sesuai karena semakin tua seseorang akan mengurangi pengeluaran biaya untuk rokok agar tetap dapat menjaga kesehatannya dan mengerti tentang bahaya rokok.

Dari Tabel 4.3 dan 4.4 dapat dilihat bahwa pada model regresi *Tobit* dan pada model regresi OLS ada empat variabel dengan penduga koefisien yang signifikan terhadap pengeluaran biaya konsumsi rokok rumah tangga. Pada kedua model variabel yang signifikan sama baik di regresi *Tobit* maupun di regresi OLS. Namun walaupun terlihat sama baik, namun pada regresi OLS diragukan hasilnya karena terdapat dua asumsi yang tidak terpenuhi.



Umur kepala rumah tangga baik pada regresi OLS maupun regresi *Tobit* mempunyai pengaruh terhadap pengeluaran biaya konsumsi rokok rumah tangga. Penduga koefisien dari variabel Umur kepala rumah tangga berpengaruh secara signifikan dan bertanda negatif. Hal ini menunjukkan bahwa bertambahnya umur kepala rumah tangga akan mengurangi biaya pengeluaran konsumsi rokok. Hal ini disebabkan bertambahnya pengetahuan dan kedewasaan dalam hal kesehatan dan bahaya merokok.

Penduga Koefisien dari variabel Proporsi anggota rumah tangga dewasa (PART) mempunyai pengaruh positif dan signifikan pada model regresi *Tobit* maupun regresi OLS. Hal ini dimungkinkan terjadi karena ketika ada anggota rumah tangga dewasa yang ada di dalam rumah tangga tersebut yang kebanyakan atau cenderung merokok pasti mengalokasikan pengeluaran untuk biaya rokok lebih. Misalnya jika tidak ada anggota rumah tangga yang merokok, biasanya membeli rokok satu pack, ketika ada anggota rumah tangga yang merokok sekarang membeli rokok menjadi dua pack.

Pengaruh penduga koefisien dari variabel pendapatan rumah tangga (PDT) pada pengeluaran biaya konsumsi rokok pada model regresi *Tobit* dan regresi OLS adalah positif dan signifikan. Hal ini menunjukkan bahwa semakin besar pendapatan rumah tangga semakin besar pengeluaran biaya konsumsi rokok. Ini dapat terjadi dikarenakan rumah tangga yang memiliki pendapatan lebih mempunyai banyak pilihan untuk membeli sesuatu sebagai pemenuhan kebutuhannya termasuk membeli rokok.

Pada dasarnya biaya untuk membeli barang pelengkap atau komplementer (HBK) dikeluarkan pada saat membeli rokok. Pada model regresi *Tobit* maupun regresi OLS terlihat bahwa penduga koefisien dari variabel harga barang komplementer berpengaruh signifikan dan positif. Hal ini wajar terjadi dikarenakan semakin besar atau banyak barang pelengkap yang dibeli, semakin besar pula pengeluaran untuk biaya konsumsi rokok.

Berdasarkan nilai *Akaike criterion* antara model regresi *Tobit* dengan model regresi OLS pada pengeluaran biaya konsumsi untuk rokok, nilai *Akaike criterion* model regresi *Tobit* lebih kecil dibandingkan model regresi OLS. Berdasarkan hal tersebut dapat disimpulkan bahwa model dari regresi *Tobit* lebih baik dari model regresi OLS.



#### 4.2.2 Pengujian Secara Individu

Pengujian secara individu digunakan untuk mengetahui variabel-variabel prediktor yang berpengaruh signifikan secara individu terhadap model regresi *Tobit*. Pengujian secara individu dilakukan dengan meregresikan satu per satu variabel prediktor dengan pengeluaran rumah tangga untuk konsumsi rokok. Pengujian secara individu menggunakan hipotesis sebagai berikut.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 4$$

Statistik uji yang digunakan pada pengujian secara individu ini adalah statistik uji Wald, dalam perhitungan menggunakan program keputusan didasarkan pada: Tolak  $H_0$  jika  $P\text{-value} < \alpha$  (10%). Nilai penduga koefisien untuk masing-masing variabel prediktor diberikan pada Tabel 4.5 untuk *Tobit* dan Tabel 4.6 untuk OLS.

**Tabel 4.5** Hasil Pengujian Individu Penduga Koefisien Tobit Antara Variabel Prediktor dengan Pengeluaran Rumah Tangga Untuk Konsumsi Rokok

Model Dugaan	Variabel	Penduga Koefisien	<i>P-value</i>
$y = 156796 - 1011.263Umur$	Intersep	156.796	<0.00001***
	Umur	-1.011,263	0.01659**
$y = 54550.1 + 1031.89PART$	Intersep	54.550,1	0.00087***
	PART	1.031,89	0.00001***
$y = 25569.1 + 0.0432781PDT$	Intersep	25.569,1	0.05850*
	PDT	0,0432781	<0.00001***
$y = -66582 + 81.4459HBK$	Intersep	-66.582	0.00016***
	HBK	81,4459	<0.00001***

Keterangan: \* Signifikan pada taraf  $\alpha$  (10%)  
 \*\* Signifikan pada taraf  $\alpha$  (5%)  
 \*\*\* Signifikan pada taraf  $\alpha$  (1%)

**Tabel 4.6** Hasil Pengujian Individu Penduga Koefisien OLS Antara Variabel Prediktor dengan Pengeluaran Rumah Tangga Untuk Konsumsi Rokok

Model Dugaan	Variabel	Penduga Koefisien	P-value
$y = 157995 - 636.794Umur$	Intersep Umur	157.995 -636,794	<0.00001*** 0.11426
$y = 68261.9 + 896.533PART$	Intersep PART	68.261,9 896,533	<0.00001*** 0.00003***
$y = 37784.5 + 0.0398408PDT$	Intersep PDT	37.784,5 0,0398408	0.00181*** <0.00001***
$y = 4403.7 + 54.3652HBK$	Intersep HBK	4.403,7 54,3652	0.69289 <0.00001***

Keterangan: \* Signifikan pada taraf  $\alpha$  (10%)  
 \*\* Signifikan pada taraf  $\alpha$  (5%)  
 \* \*\* Signifikan pada taraf  $\alpha$  (1%)

Berdasarkan Tabel 4.5 dan 4.6 dapat diketahui bahwa perbandingan hasil pengujian secara individu penduga koefisien antara *Tobit* dan OLS untuk memodelkan pengeluaran rumah tangga untuk konsumsi rokok, penduga koefisien dari variabel-variabel prediktor yang tidak berpengaruh signifikan setelah dilakukan pengujian individu pada OLS adalah umur kepala rumah tangga (Umur). Sedangkan pada *Tobit* penduga koefisien dari variabel-variabel prediktor semua berpengaruh signifikan. Hasil pengujian individu penduga koefisien tersebut menunjukkan ada penduga koefisien yang bertanda negatif pada OLS maupun *Tobit*, yaitu penduga koefisien dari variabel umur kepala rumah tangga (Umur). Hal ini berarti variabel tersebut menyebabkan penurunan terhadap pengeluaran rumah tangga untuk konsumsi rokok.

### 4.2.3 Pengujian Asumsi

Pada regresi *Tobit* maupun OLS sebelum dilakukan pengujian secara individu maupun secara serentak terlebih dahulu dilakukan pengujian asumsi. Pengujian asumsi ini terdiri dari normalitas, heteroskedastisitas, multikolinieritas dan autokorelasi.

Tabel 4.7 menyajikan ringkasan beberapa uji asumsi yang telah dilakukan, yang menunjukkan bahwa pada regresi *Tobit* semua asumsi terpenuhi sedangkan pada regresi OLS terdapat dua asumsi

yang tidak terpenuhi yaitu normalitas dan heterokedastisitas (hasil pengujian asumsi selengkapannya dapat dilihat pada Lampiran...). Hal ini menunjukkan bahwa regresi *Tobit* dapat memperbaiki asumsi yang tidak terpenuhi dalam regresi OLS.

**Tabel 4.7** Hasil Pengujian Asumsi antara Regresi *Tobit* dengan Regresi OLS

Asumsi	OLS	<i>Tobit</i>
Uji Normalitas	Tidak terpenuhi	Terpenuhi
Uji Heterokedastisitas	Tidak terpenuhi	Terpenuhi
Uji Multikolinieritas	Terpenuhi	Terpenuhi
Uji Autokorelasi	Terpenuhi	Terpenuhi

Untuk melihat apakah regresi *Tobit* lebih baik dari regresi OLS secara kesamaan tanda, signifikansi dan model terbaik, pada data pengeluaran biaya untuk konsumsi rokok tersebut dilakukan pengulangan sebanyak 100 kali dan diambil sampel sebanyak 30 rumah tangga secara acak.

**Tabel 4.8** Hasil *Resampling* Perbandingan antara Model Regresi *Tobit* dengan Model Regresi OLS

Variabel	Kesamaan Tanda Penduga Koefisien		Signifikansi Penduga Koefisien	
	<i>Tobit</i>	OLS	<i>Tobit</i>	OLS
Intersep	91 negatif	70 negatif	49 nyata	16 nyata
	9 positif	30 positif	51 tidak nyata	84 tidak nyata
Umur	73 negatif	80 negatif	30 nyata	23 nyata
	27 positif	20 positif	70 tidak nyata	77 tidak nyata
PART	15 negatif	15 negatif	43 nyata	40 nyata
	85 positif	85 positif	57 tidak nyata	60 tidak nyata

**Tabel 4.8**(lanjutan)

PDT	17 negatif	10 negatif	46 nyata	49 nyata
	83 positif	90 positif	54 tidak nyata	51 tidak nyata
HBK	0 negatif	0 negatif	96 nyata	95 nyata
	100 positif	100 positif	4 tidak nyata	5 tidak nyata

Tabel 4.8 menunjukkan seberapa konsisten dan kestabilan dari masing-masing penduga koefisien pada saat dilakukan pengulangan sebanyak 100 kali. Ada beberapa indikator yang menjadi acuan untuk menentukan seberapa konsisten. Indikator tersebut yaitu kesamaan tanda, signifikansi, uji serempak dan model yang terbaik.

Pada indikator kesamaan tanda, setelah dilakukan pengulangan sebanyak 100 kali dihasilkan 91 penduga koefisien Intersep yang bernilai negatif dan 9 penduga koefisien intersep yang bernilai positif pada regresi *Tobit*, sedangkan diperoleh 70 penduga koefisien intersep yang bernilai negatif dan 30 penduga koefisien intersep yang bernilai positif pada regresi OLS. Hal ini menunjukkan bahwa secara konsisten dan kestabilan regresi *Tobit* lebih baik daripada regresi OLS karena memiliki nilai penduga koefisien negatif yang lebih besar sehingga kesamaan tanda sama dengan model awal yang didapatkan. Pada indikator signifikansi dihasilkan 49 penduga koefisien intersep yang nyata dan 51 penduga koefisien intersep yang tidak nyata pada regresi *Tobit*, sedangkan diperoleh 16 penduga koefisien intersep yang nyata dan 84 penduga koefisien intersep yang tidak nyata pada regresi OLS. Hal ini menunjukkan bahwa regresi *Tobit* lebih baik dari regresi OLS karena memiliki nilai signifikansi penduga koefisien nyata yang lebih stabil sesuai dengan model awal yang didapatkan. Sehingga regresi *Tobit* dikatakan lebih baik dari regresi OLS pada indikator kesamaan tanda dan signifikansi.

Pada indikator kesamaan tanda, setelah dilakukan pengulangan sebanyak 100 kali dihasilkan 73 penduga koefisien umur yang bernilai negatif dan 27 penduga koefisien umur yang bernilai positif pada regresi *Tobit*, sedangkan diperoleh 80 penduga koefisien umur yang bernilai negatif dan 20 penduga koefisien umur yang bernilai positif pada regresi OLS. Hal ini menunjukkan regresi OLS lebih baik dari regresi *Tobit* karena memiliki nilai penduga koefisien



negatif yang lebih stabil sehingga kesamaan tanda sama dengan model awal yang didapatkan. Pada indikator signifikansi dihasilkan 30 penduga koefisien umur yang nyata dan 70 penduga koefisien umur yang tidak nyata pada regresi *Tobit*, sedangkan diperoleh 23 penduga koefisien umur yang nyata dan 77 penduga koefisien umur yang tidak nyata pada regresi OLS. Hal ini menunjukkan bahwa regresi *Tobit* lebih baik dari regresi OLS karena memiliki nilai signifikansi penduga koefisien nyata yang lebih stabil sesuai dengan model awal yang didapatkan. Jadi walaupun pada indikator kesamaan tanda menunjukkan regresi OLS lebih baik dari regresi *Tobit* namun pada indikator signifikansi regresi *Tobit* lebih baik dari regresi OLS.

Pada indikator kesamaan tanda, setelah dilakukan pengulangan sebanyak 100 kali dihasilkan 15 penduga koefisien PART yang bernilai negatif dan 85 penduga koefisien PART yang bernilai positif pada regresi *Tobit* maupun pada regresi OLS. Hal ini menunjukkan regresi *Tobit* maupun regresi OLS sama baiknya pada indikator kesamaan tanda. Pada indikator signifikansi dihasilkan 43 penduga koefisien PART yang nyata dan 57 penduga koefisien PART yang tidak nyata pada regresi *Tobit*, sedangkan diperoleh 40 penduga koefisien PART yang nyata dan 60 penduga koefisien PART yang tidak nyata pada regresi OLS. Hal ini menunjukkan bahwa regresi *Tobit* lebih baik dari regresi OLS karena memiliki nilai signifikansi penduga koefisien nyata yang lebih stabil sesuai dengan model awal yang didapatkan. Jadi pada pengujian variabel PART regresi *Tobit* lebih baik dari regresi OLS.

Pada indikator kesamaan tanda, setelah dilakukan pengulangan sebanyak 100 kali dihasilkan 17 penduga koefisien PDT yang bernilai negatif dan 83 penduga koefisien PDT yang bernilai positif pada regresi *Tobit*, sedangkan diperoleh 10 penduga koefisien PDT yang bernilai negatif dan 90 penduga koefisien PDT yang bernilai positif pada regresi OLS. Hal ini menunjukkan regresi OLS lebih baik dari regresi *Tobit* karena memiliki nilai penduga koefisien positif yang lebih stabil sehingga kesamaan tanda sama dengan model awal yang didapatkan. Pada indikator signifikansi dihasilkan 46 penduga koefisien PDT yang nyata dan 54 penduga koefisien PDT yang tidak nyata pada regresi *Tobit*, sedangkan diperoleh 49 penduga koefisien PDT yang nyata dan 51 penduga koefisien PDT yang tidak nyata pada regresi OLS. Hal ini menunjukkan bahwa regresi OLS lebih baik dari regresi *Tobit* karena memiliki nilai



signifikansi penduga koefisien nyata yang lebih stabil sesuai dengan model awal yang didapatkan. Jadi pada pengujian variabel PDT regresi OLS lebih baik dari regresi *Tobit*.

Pada indikator kesamaan tanda, setelah dilakukan pengulangan sebanyak 100 kali dihasilkan 0 penduga koefisien HBK yang bernilai negatif dan 100 penduga koefisien HBK yang bernilai positif pada regresi *Tobit* maupun pada regresi OLS. Hal ini menunjukkan regresi *Tobit* maupun regresi OLS sama baiknya pada indikator kesamaan tanda. Pada indikator signifikansi dihasilkan 96 penduga koefisien HBK yang nyata dan 4 penduga koefisien HBK yang tidak nyata pada regresi *Tobit*, sedangkan diperoleh 95 penduga koefisien HBK yang nyata dan 5 penduga koefisien HBK yang tidak nyata pada regresi OLS. Hal ini menunjukkan bahwa regresi *Tobit* lebih baik dari regresi OLS karena memiliki nilai signifikansi penduga koefisien nyata yang lebih stabil sesuai dengan model awal yang didapatkan. Jadi pada pengujian variabel HBK regresi *Tobit* lebih baik dari regresi OLS.

Pada indikator uji serempak, setelah dilakukan pengulangan sebanyak 100 kali, pada regresi *Tobit* maupun pada regresi OLS memiliki nilai signifikansi penduga koefisien nyata sama besarnya sehingga dapat dikatakan regresi *Tobit* maupun OLS sama baiknya pada indikator uji serempak.

Pada indikator model terbaik, setelah dilakukan pengulangan sebanyak 100 kali didapatkan nilai AIC dari regresi *Tobit* lebih kecil daripada regresi OLS sehingga dapat dikatakan untuk pemilihan model terbaik, model regresi *Tobit* secara konsisten dan stabil lebih baik daripada model OLS.

## BAB V KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Berdasarkan hasil analisis data dan pembahasan yang telah dilakukan, dapat diambil kesimpulan sebagai berikut.

- Rata-rata pengeluaran rumah tangga untuk biaya konsumsi rokok sebesar 126.287 rupiah. Nilai maksimum pengeluaran rumah tangga untuk biaya konsumsi rokok sebesar 315.000 rupiah. Sedangkan nilai minimum pengeluaran rumah tangga untuk biaya konsumsi rokok sebesar 0 rupiah atau dapat dikatakan rumah tangga tersebut tidak mengeluarkan biaya untuk membeli rokok.
- Secara matematis, model regresi *Tobit* yang dihasilkan untuk pengeluaran rumah tangga terhadap konsumsi rokok adalah sebagai berikut.

$$\hat{y} = -83967.7 - 679.642Umur + 789.493PART \\ + 0.0167134PDT + 65.5285HBK$$

Sedangkan model regresi OLS yang dihasilkan untuk pengeluaran rumah tangga terhadap konsumsi rokok adalah sebagai berikut.

$$\hat{y} = -29147.8 - 582.485Umur + 721.717PART \\ + 0.0219704PDT + 39.6633HBK$$

- Penduga koefisien dari variabel yang bertanda negatif yaitu variabel umur kepala rumah tangga. Hal ini menunjukkan bahwa variabel tersebut memberikan penurunan terhadap pengeluaran rumah tangga untuk biaya konsumsi rokok. Hal ini sesuai karena semakin tua seseorang akan mengurangi pengeluaran biaya untuk rokok agar tetap dapat menjaga kesehatannya. Sedangkan penduga koefisien dari variabel yang bertanda positif yaitu Proporsi anggota rumah tangga dewasa, pendapatan rumah tangga dan harga barang komplementer. Hal ini menunjukkan bahwa variabel tersebut memberikan peningkatan terhadap pengeluaran rumah tangga untuk biaya konsumsi rokok.
- Pada pengulangan sebanyak 100 kali, indikator uji serempak menunjukkan bahwa regresi *Tobit* maupun regresi OLS sama baiknya. Untuk indikator kesamaan tanda pada variabel Umur, regresi OLS lebih baik daripada regresi *Tobit* namun

pada indikator signifikansi regresi *Tobit* lebih baik daripada regresi OLS.

- Pada variabel PART dan variabel HBK untuk indikator kesamaan tanda regresi *Tobit* maupun OLS sama baiknya. Sedangkan untuk indikator signifikansi regresi *Tobit* lebih baik daripada regresi OLS.
- Pada variabel PDT untuk indikator kesamaan tanda maupun signifikansi regresi OLS lebih baik daripada regresi *Tobit*.
- Pada indikator pemilihan model terbaik, model regresi *Tobit* lebih baik daripada model regresi OLS karena memiliki nilai AIC yang lebih kecil. Jadi secara keseluruhan regresi *Tobit* secara konsisten dan stabil lebih baik daripada regresi OLS.

## 5.2 Saran

Berdasarkan analisis yang telah dilakukan masih terdapat kendala yang dapat dijadikan saran perbaikan untuk penelitian selanjutnya. Kendala ini terletak pada penentuan variabel-variabel prediktor yang masih terbatas sehingga hal ini dapat berakibat pada hasil model yang dihasilkan serta informasi yang diberikan. Untuk penelitian selanjutnya disarankan menambahkan variabel-variabel prediktor ke dalam perhitungan model.

## DAFTAR PUSTAKA

- Akaike, H. 1978. *A Bayesian Analysis of The Minimum AIC Procedure*. Annals of the Institute of Statistical Mathematics, Part A, Hal. 9-14.  
[http://www.ism.ac.jp/editsec/aism/pdf/030\\_1\\_0009.pdf](http://www.ism.ac.jp/editsec/aism/pdf/030_1_0009.pdf). Tanggal Akses: 17 Agustus 2013.
- Bierens, H.J. 2004. *The Tobit Model*. <http://econ.la.psu.edu/./Tobit.PDF>. Diakses pada 27 Maret 2013.
- BPS KotaKediri Jawa Timur. 2011. *Hasil Survey Sosial ekonomi Profil Rumah Tangga Pengeluaran Konsumsi Rokok 2011*.
- DeMaris. 2004. *Regression with Social Data: Modeling Continuous and Limited Response Variables*. Hoboken, NJ: Wiley.
- Draper, N.R dan H, Smith. 1981. *Applied Regression Analysis, Second Edition*. John Wiley & Sons, Inc.
- Fair, R. C. 1977. A Note on the Computation of The Tobit Estimator. *Jurnal Econometrica*, Vol. 45, No. 7.
- Greene, W.H. 1990. *Econometric Analysis*, Macmillan Publishing Company, New York.
- Gujarati, D. 1991. *Basic Econometrics*, Mc-Graw-Hill, Inc.
- Halifah. 2012. Analisis Faktor yang Mempengaruhi Permintaan Rokok Masyarakat Di Kota Makassar Tahun 2012. *Universitas Hasanuddin*.
- Hanief, I.U. 2010. *Analisis Regresi Tobit terhadap Faktor-faktor yang Mempengaruhi Pengeluaran Biaya Kesehatan Rumah Tangga (RT) di Wilayah Perkotaan dan Pedesaan di Propinsi Jawa Timur*. Jurusan Statistika, Institut Teknologi Sepuluh Nopember : Surabaya

Hosmer, D. W dan Lemeshow, S. 2000. *Applied Logistic Regression*. New York : John Wiley and Sons, Inc.

Kindrana, S. S. 2011. *Analisis Faktor-Faktor yang Mempengaruhi Remitan TKI Jawa Timur dengan Menggunakan Regresi Tobit*. Surabaya: Program Sarjana, ITS Surabaya.

Laily, U. 2010. *Analisis Faktor – Faktor Yang Mempengaruhi Pengeluaran Konsumsi Untuk Makanan Berprotein Dengan Menggunakan Regresi Tobit*. Surabaya: Program Sarjana, ITS Surabaya

Laventhal, H. dan Cleary. 2000. *The Smoking Problem: A Review of The Research and Theory in Behavioral Risk Modification*. *Psychological Bulletin*, Vol.88, No.2, 370-405.

Long, J. S. 1997. 'Regression Models For Categorical and Limited Dependent variables'. Thousand Oaks, CA: SAGE Publication. Inc.

Maddala, G, S. 1983. *Limited dependent and qualitative variables in econometrics*. Cambridge University Press. New York.

Marnique, J. dan H. H, Jensen. 2013. 'Consumption of Tobacco and Alcoholic Beverages among Spanish Consumers', *Southwest Economic Review, University of Houston, Iowa State Univesity*, dalam [www.ser.tcu.edu/2004%20Marnique%20Jensen%2041-56.pdf](http://www.ser.tcu.edu/2004%20Marnique%20Jensen%2041-56.pdf) didownload pada tanggal 05 Maret 2013.

Robinson, P.M., A.K, Bera. and C.M, Jarque. 1985. *Test for Serial Dependence in Limited Dependent Variable Model*. Economics Departement of the University of Pennsylvania. Pennsylvania.

Tauchman, H., G, S., T, Requate. dan M.C, Schmidt. 2008. '*Tobacco and Alcohol: Complements or substitute? A Structural Approach*', *Forschungsinstitut zur Zukunft der Arbeit Institute for thr Study of labour*, Jerman.



Tobin, J. 1958. Estimation of Relationship for Limited Dependent Variables. *Econometrica*. Vol. 26, No. 1 : 24-36.

Yen, T. S. 2003. 'Estimating Demand for Cigarettes and Alcohol With Zero Observation: A Censored System', *Departement of Agriculture Economics University of Tennessee*

Zain, I. 1997. *Model Regresi Tobit dan Aplikasinya*, Lembaga penelitian Institut Teknologi Sepuluh Nopember, Surabaya



UNIVERSITAS BRAWIJAYA



Lampiran 1. Data Primer Hasil Susenas Pengeluaran Konsumsi  
Rokok Rumah Tangga Kota Kediri 2011

Y (Rupiah)	Umur (tahun)	PART (persen)	PDT (Rupiah)	HBK (Rupiah)
122100	37	75.00	1311900	2000
85700	46	25.00	3527646	2500
0	56	50.00	1472700	0
42800	66	40.00	2749011	2500
77100	81	42.86	1537301	2500
98571	69	50.00	2042900	2500
102800	55	25.00	2381264	2500
0	62	50.00	1506000	0
122100	59	33.33	3192000	2500
122100	38	100.00	1558100	2000
68500	70	50.00	1391400	2500
115700	38	33.33	2639200	2500
54000	76	66.67	1109100	2500
47500	49	16.67	1228856	2500
210000	54	28.57	2786300	2500
0	17	0.00	1587400	0
115700	36	100.00	2009500	2000
114000	45	25.00	2342917	2500
51400	68	50.00	988100	2500
128500	53	33.33	1583900	2500
225000	35	100.00	2745900	2000
154200	66	66.67	2061017	2500
137100	71	66.67	1977200	2500
195000	67	60.00	2096934	2500
38500	65	50.00	2838020	2500
102800	61	50.00	1802281	2500
...	...	...	...	...
0	59	75.00	1343600	0

Lampiran 2. Deskripsi Variabel Pengeluaran Rumah Tangga Untuk Biaya Konsumsi Rokok dengan Variabel Prediktor

**Descriptive Statistics**

	N	Minimum	Maximum	Mean	Variance
Y	217	0	315000	126287.07	6401055516.865
Umur	217	17	83	49.79	182.489
PART	217	.00	100.00	64.7221	632.880
PDT	217	574417	7095900	2221402.60	937656205345.112
HBK	217	0	5000	2241.94	854390.681



### Lampiran 3. Pengujian Asumsi Regresi Tobit dan Regresi OLS

#### Uji Asumsi Tobit

- Uji Normalitas

Test for normality of residual -

Null hypothesis: error is normally distributed

Test statistic: Chi-square(2) = 2.31556

with p-value = 0.314183

- Uji Multikolinieritas

Variance Inflation Factors

Minimum possible value = 1.0

Values > 10.0 may indicate a collinearity problem

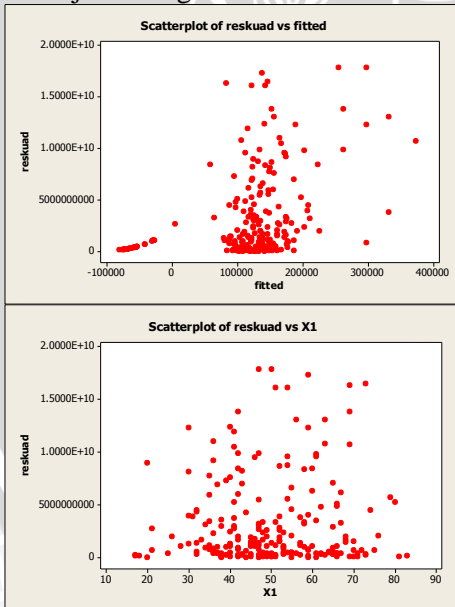
X1 1.027

X3 1.076

X4 1.328

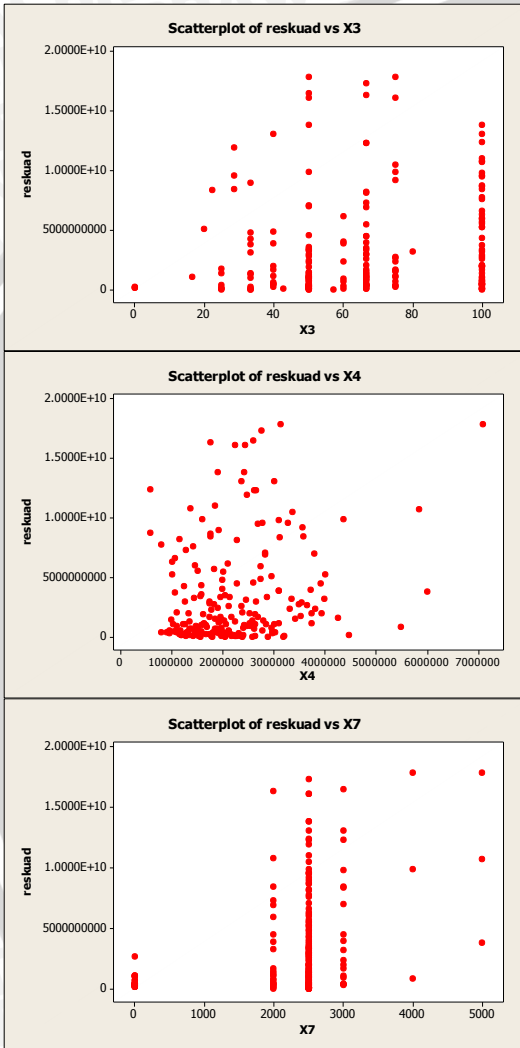
X7 1.362

- Uji Heterogenitas



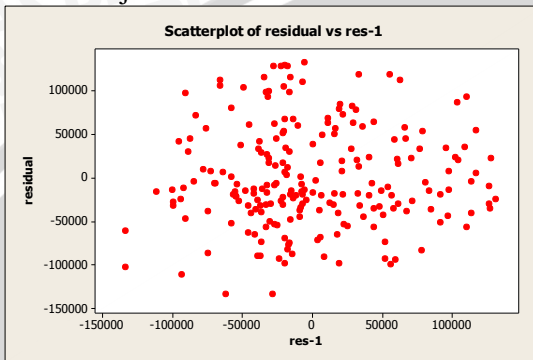


### Lampiran 3. Pengujian Asumsi Regresi Tobit dan Regresi OLS (Lanjutan)



### Lampiran 3. Pengujian Asumsi Regresi Tobit dan Regresi OLS (Lanjutan)

- Uji Autokorelasi



#### Uji Asumsi OLS

- Uji Normalitas

Test for normality of residual -

Null hypothesis: error is normally distributed

Test statistic: Chi-square(2) = 8.21615

with p-value = 0.0164394

- Uji Heterokedastisitas

Breusch-Pagan test for heteroskedasticity -

Null hypothesis: heteroskedasticity not present

Test statistic: LM = 21.0699

with p-value =  $P(\text{Chi-square}(4) > 21.0699) = 0.000306728$

- Uji Multikolinieritas

Variance Inflation Factors

Minimum possible value = 1.0

Values > 10.0 may indicate a collinearity problem

X1 1.027

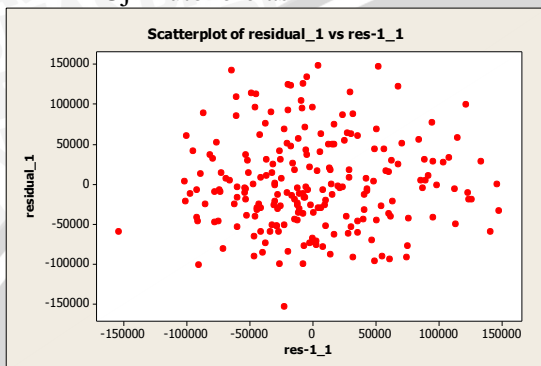
X3 1.076

X4 1.328

X7 1.36

### Lampiran 3. Pengujian Asumsi Regresi Tobit dan Regresi OLS (Lanjutan)

- Uji Autokorelasi



Lampiran 4. Hasil Pemodelan Lengkap Regresi Tobit dan Regresi OLS

Model 2: Tobit, using observations 1-217

Dependent variable: Y

Standard errors based on Hessian

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
Const	-83967.7	25300.2	-3.3189	0.00090	***
X1	-679.642	335.48	-2.0259	0.04278	**
X3	789.493	177.613	4.4450	<0.00001	***
X4	0.0167134	0.00518852	3.2212	0.00128	***
X7	65.5285	7.20418	9.0959	<0.00001	***
Chi-square(4)		189.3974	p-value		7.14e-40
Log-likelihood		-2386.695	Akaike criterion		4785.390
Schwarz criterion		4805.669	Hannan-Quinn		4793.582
sigma = 61485.3 (3204.05)					
Left-censored observations: 26					
Right-censored observations: 0					

Model 3: OLS, using observations 1-217

Dependent variable: Y

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	-29147.8	20505.8	-1.4214	0.15666	
X1	-582.485	294.257	-1.9795	0.04905	**
X3	721.717	161.688	4.4636	0.00001	***
X4	0.0219704	0.00466656	4.7080	<0.00001	***
X7	39.6633	4.95172	8.0100	<0.00001	***
Mean dependent var	126287.1	S.D. dependent var		80006.60	
Sum squared resid	7.04e+11	S.E. of regression		57639.60	
R-squared	0.490584	Adjusted R-squared		0.480973	
F(4, 212)	51.04075	P-value(F)		4.72e-30	
Log-likelihood	-2684.127	Akaike criterion		5378.254	
Schwarz criterion	5395.153	Hannan-Quinn		5385.080	

Lampiran 5. Hasil Pengujian Secara Individu Regresi Tobit dan Regresi OLS

Model 1: OLS, using observations 1-217

Dependent variable: Y

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	157995	20714.8	7.6271	<0.00001	***
X1	-636.794	401.572	-1.5858	0.11426	
Mean dependent var	126287.1	S.D. dependent var		80006.60	
Sum squared resid	1.37e+12	S.E. of regression		79727.56	
R-squared	0.011561	Adjusted R-squared		0.006963	
F(1, 215)	2.514615	P-value(F)		0.114265	
Log-likelihood	-2756.047	Akaike criterion		5516.095	
Schwarz criterion	5522.855	Hannan-Quinn		5518.826	

Model 3: OLS, using observations 1-217

Dependent variable: Y

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	68261.9	14445.8	4.7254	<0.00001	***
X3	896.533	208.099	4.3082	0.00003	***
Mean dependent var	126287.1	S.D. dependent var		80006.60	
Sum squared resid	1.27e+12	S.E. of regression		76940.12	
R-squared	0.079468	Adjusted R-squared		0.075187	
F(1, 215)	18.56062	P-value(F)		0.000025	
Log-likelihood	-2748.325	Akaike criterion		5500.650	
Schwarz criterion	5507.410	Hannan-Quinn		5503.380	



Lampiran 5. Hasil Pengujian Secara Individu Regresi Tobit dan Regresi OLS (Lanjutan)

Model 4: OLS, using observations 1-217

Dependent variable: Y

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	37784.5	11958.1	3.1597	0.00181	***
X4	0.0398408	0.00493651	8.0707	<0.00001	***
Mean dependent var	126287.1	S.D. dependent var		80006.60	
Sum squared resid	1.06e+12	S.E. of regression		70253.62	
R-squared	0.232514	Adjusted R-squared		0.228944	
F(1, 215)	65.13540	P-value(F)		4.91e-14	
Log-likelihood	-2728.596	Akaike criterion		5461.192	
Schwarz criterion	5467.952	Hannan-Quinn		5463.923	

Model 7: OLS, using observations 1-217

Dependent variable: Y

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
Const	4403.7	11135.3	0.3955	0.69289	
X7	54.3652	4.59341	11.8355	<0.00001	***
Mean dependent var	126287.1	S.D. dependent var		80006.60	
Sum squared resid	8.37e+11	S.E. of regression		62400.82	
R-squared	0.394501	Adjusted R-squared		0.391684	
F(1, 215)	140.0788	P-value(F)		3.25e-25	
Log-likelihood	-2702.874	Akaike criterion		5409.749	
Schwarz criterion	5416.509	Hannan-Quinn		5412.480	

Lampiran 5. Hasil Pengujian Secara Individu Regresi Tobit dan Regresi OLS (Lanjutan)

Model 8: Tobit, using observations 1-217

Dependent variable: Y

Standard errors based on Hessian

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
Const	156796	23370	6.7093	<0.00001	***
X1	-1011.263	453.831	-4.5672	0.01659	**
Chi-square(1)	2.456245		p-value		0.117058
Log-likelihood	-2478.126		Akaike criterion		4962.252
Schwarz criterion	4972.392		Hannan-Quinn		4966.348

sigma = 88596.5 (4317.96)

Left-censored observations: 26

Right-censored observations: 0

Test for normality of residual -

Null hypothesis: error is normally distributed

Test statistic: Chi-square(2) = 6.96154

with p-value = 0.0307837

Model 10: Tobit, using observations 1-217

Dependent variable: Y

Standard errors based on Hessian

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
Const	54550.1	16386.9	3.3289	0.00087	***
X3	1031.89	234.731	4.3960	0.00001	***
Chi-square(1)	19.32506		p-value		0.000011
Log-likelihood	-2469.977		Akaike criterion		4945.954
Schwarz criterion	4956.094		Hannan-Quinn		4950.050

sigma = 85365.2 (4190.71)

Left-censored observations: 26

Right-censored observations: 0

Test for normality of residual -

Null hypothesis: error is normally distributed

Test statistic: Chi-square(2) = 3.52505

with p-value = 0.171611

Lampiran 5. Hasil Pengujian Secara Individu Regresi Tobit dan Regresi OLS (Lanjutan)

Model 11: Tobit, using observations 1-217

Dependent variable: Y

Standard errors based on Hessian

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
Const	25569.1	13515.1	1.8919	0.05850	*
X4	0.0432781	0.0055387	7.8138	<0.00001	***
Chi-square(1)	61.05510		p-value		5.55e-15
Log-likelihood	-2452.237		Akaike criterion		4910.474
Schwarz criterion	4920.614		Hannan-Quinn		4914.570

sigma = 77894.9 (3820.5)

Left-censored observations: 26

Right-censored observations: 0

Test for normality of residual -

Null hypothesis: error is normally distributed

Test statistic: Chi-square(2) = 6.88094

with p-value = 0.0320497

Model 14: Tobit, using observations 1-217

Dependent variable: Y

Standard errors based on Hessian

	<i>Coefficient</i>	<i>Std. Error</i>	<i>z</i>	<i>p-value</i>	
Const	-66582	17618.9	-3.7790	0.00016	***
X7	81.4459	6.97539	11.6762	<0.00001	***
Chi-square(1)	136.3332		p-value		1.69e-31
Log-likelihood	-2400.561		Akaike criterion		4807.122
Schwarz criterion	4817.261		Hannan-Quinn		4811.218

sigma = 66592.3 (3451.92)

Left-censored observations: 26

Right-censored observations: 0

Test for normality of residual -

Null hypothesis: error is normally distributed

Test statistic: Chi-square(2) = 27.8841

with p-value = 8.81145e-007

Lampiran 6. Hasil Pengacakan Sebanyak 100 kali

- Intersep

**kesamaan tanda OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	70	70.0	70.0	70.0
Valid Positif	30	30.0	30.0	100.0
Total	100	100.0	100.0	

**kesamaan tanda Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	91	91.0	91.0	91.0
Valid Positif	9	9.0	9.0	100.0
Total	100	100.0	100.0	

**Signifikansi OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	16	16.0	16.0	16.0
Valid Tidak Nyata	84	84.0	84.0	100.0
Total	100	100.0	100.0	

**Signifikansi Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	49	49.0	49.0	49.0
Valid Tidak Nyata	51	51.0	51.0	100.0
Total	100	100.0	100.0	

Lampiran 6. Hasil Pengacakan Sebanyak 100 kali (Lanjutan)

- Variabel  $X_1$  (Umur Kepala Rumah Tangga)

**kesamaan tanda OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	80	80.0	80.0	80.0
Valid Positif	20	20.0	20.0	100.0
Total	100	100.0	100.0	

**kesamaan tanda Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	73	73.0	73.0	73.0
Valid Positif	27	27.0	27.0	100.0
Total	100	100.0	100.0	

**Signifikansi OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	23	23.0	23.0	23.0
Valid Tidak Nyata	77	77.0	77.0	100.0
Total	100	100.0	100.0	

**Signifikansi Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	30	30.0	30.0	30.0
Valid Tidak Nyata	70	70.0	70.0	100.0
Total	100	100.0	100.0	



Lampiran 6. Hasil Pengacakan Sebanyak 100 kali (Lanjutan)

- Variabel  $X_2$  (Proporsi Anggota Rumah Tangga Dewasa)

**kesamaan tanda OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	15	15.0	15.0	15.0
Valid Positif	85	85.0	85.0	100.0
Total	100	100.0	100.0	

**kesamaan tanda Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	15	15.0	15.0	15.0
Valid Positif	85	85.0	85.0	100.0
Total	100	100.0	100.0	

**Signifikansi OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	40	40.0	40.0	40.0
Valid Tidak Nyata	60	60.0	60.0	100.0
Total	100	100.0	100.0	

**Signifikansi Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	43	43.0	43.0	43.0
Valid Tidak Nyata	57	57.0	57.0	100.0
Total	100	100.0	100.0	

Lampiran 6. Hasil Pengacakan Sebanyak 100 kali (Lanjutan)

- Variabel  $X_3$  (Pendapatan Rumah Tangga)

**kesamaan tanda OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	10	10.0	10.0	10.0
Valid Positif	90	90.0	90.0	100.0
Total	100	100.0	100.0	

**kesamaan tanda Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Negatif	17	17.0	17.0	17.0
Valid Positif	83	83.0	83.0	100.0
Total	100	100.0	100.0	

**Signifikansi OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	49	49.0	49.0	49.0
Valid Tidak Nyata	51	51.0	51.0	100.0
Total	100	100.0	100.0	

**Signifikansi Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	46	46.0	46.0	46.0
Valid Tidak Nyata	54	54.0	54.0	100.0
Total	100	100.0	100.0	

Lampiran 6. Hasil Pengacakan Sebanyak 100 kali (Lanjutan)

- Variabel  $X_4$  (Harga Barang Komplementer)

**kesamaan tanda OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Positif	100	100.0	100.0	100.0

**kesamaan tanda Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Positif	100	100.0	100.0	100.0

**Signifikansi OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	95	95.0	95.0	95.0
Valid Tidak Nyata	5	5.0	5.0	100.0
Total	100	100.0	100.0	

**Signifikansi Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Nyata	96	96.0	96.0	96.0
Valid Tidak Nyata	4	4.0	4.0	100.0
Total	100	100.0	100.0	

Lampiran 6. Hasil Pengacakan Sebanyak 100 kali (Lanjutan)

- Uji Serempak

**OLS**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Nyata	100	100.0	100.0	100.0

**Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Nyata	100	100.0	100.0	100.0

- Pemilihan Model Terbaik

**OLS vs Tobit**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Tobit	100	100.0	100.0	100.0