

MODEL REGRESI TOBIT KONSUMSI SUSU CAIR PABRIK (Studi Kasus Rumah Tangga di Provinsi Bali)

I PUTU JERYANA¹, I PUTU EKA NILA KENCANA², G.K. GANDHIADI³

^{1,2,3}Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali
e-mail: ¹jerry_crock30@yahoo.com, ²i.putu.enk@gmail.com, ³gandhiadigk@yahoo.com

Abstract

Regression analysis is used to study the relationship between dependent (response) variable with one or more independent (causal) variables. While response data were censored, then Tobit regression model could be applied. According to Greene (2003), censored data were data with incomplete observation or the dependent variable has a value of zero, while for the other observations have particular value. This research aimed to model dairy milk's consumption from households at Bali Province. By using data from Survey Sosial Ekonomi Nasional (SUSENAS) or Social Economy's National Survey (SENS) for year 2012, 615 households were selected as sampling unit using simple random sampling technique, and found 123 households who consumed dairy milk. The independent variables in our model were last education level completed by head of household's (X_1), head of household's work (X_2), age of head of household's (X_3), amount of expenditure for food consumption's (X_4), number of household members (X_5), and household income (X_6), the response variable was budget for buying dairy milk (Y). From six independent variables, is found only last education level by head household and amount of expenditure for food consumption had significant effect on Y 's. The final Tobit regression model were obtained using AIC (Akaike Information Criterion) method is $Y = -3314724 + 565429,7 X_1 + 0,014278 X_4$ with pseudo R^2 as much as 16.79 per cent.

Keywords: Akaike Information Criterion, Censored Data, Pseudo R^2 , Tobit Regression

1. Pendahuluan

Analisis regresi merupakan studi mengenai ketergantungan variabel terikat (*dependent variable*) dengan satu atau lebih variabel bebas (*independent variable*). Tujuan analisis regresi adalah untuk menduga nilai variabel terikat berdasarkan nilai variabel bebas yang diketahui [1]. Dalam analisis regresi linier umumnya variabel terikat yang digunakan minimal berskala pengukuran interval. Dalam beberapa kasus terdapat variabel terikat dengan data campuran. Data campuran adalah data dimana variabel terikat memiliki nilai nol untuk sebagian pengamatan, sedangkan untuk pengamatan yang lain memiliki nilai yang bervariasi, sehingga data tersebut termasuk data tersensor [2]. Salah satu karakteristik data

tersensor adalah variabel tersebut mempunyai batas atas atau batas bawah. Pada data tersensor, beberapa pengamatan berada dalam batas atas ataupun batas bawah dan pengamatan yang lain berada dalam rentang yang cukup lebar di atas atau di bawah batas [3].

Apabila data tersensor tetap dianalisis menggunakan Metode Kuadrat Terkecil (MKT) dalam proses pendugaan koefisien dari parameter regresi akan menghasilkan interpretasi yang hasilnya bias (terlalu tinggi atau terlalu rendah) dan tidak konsisten (jika ada data baru, hasilnya tidak akan sama atau tidak sesuai dengan hasil semula) [4]. Sehingga pada kasus pemodelan data tersensor, dibutuhkan metode khusus untuk menganalisis data tersebut dan metode yang dapat digunakan adalah Regresi Tobit [3].

¹ Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

² Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

1.1 Analisis Regresi

Regresi merupakan metode statistika yang menganalisis hubungan antara dua atau lebih variabel kuantitatif sehingga satu variabel bisa diramalkan dari variabel-variabel lainnya [5]. Secara umum model regresi linier dapat dituliskan sebagai berikut:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \dots + \beta_{(p-1)} X_{t(p-1)} + \varepsilon_t$$

dengan:

Y_t merupakan nilai variabel terikat pada amatan ke- t

$\beta_0, \beta_1, \dots, \beta_{(p-1)}$ adalah parameter

$X_{t1}, X_{t2}, \dots, X_{t(p-1)}$ adalah nilai variabel bebas dari amatan ke- t .

ε_t saling bebas dan menyebar normal dengan mean 0 dan variansi σ^2

Jika ditetapkan $X_{t0} = 1$, model regresi (2.1) dapat dituliskan sebagai berikut:

$$Y_t = \beta_0 X_{t0} + \beta_1 X_{t1} + \beta_2 X_{t2} + \dots + \beta_{(p-1)} X_{t(p-1)} + \varepsilon_t \text{ dengan } X_{t0} = 1$$

$$\text{atau: } Y_t = \sum_{k=0}^{p-1} \beta_k X_{tk} + \varepsilon_t$$

Karena $E(\varepsilon_t) = 0$, maka nilai harapan model regresi adalah:

$$E(Y_t) = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \dots + \beta_{(p-1)} X_{t(p-1)}$$

Sehingga model regresi umum dengan sisaan menyebar normal berimplikasi bahwa amatan-amatan Y_t adalah variabel acak normal bebas, dengan rata-rata $E(Y_t)$ dengan ragam konstan σ^2 .

Dalam penggunaan metode regresi linier terdapat beberapa asumsi yang harus dipenuhi seperti kehomogenan ragam dan kenormalan sisaan. Asumsi kehomogenan ragam dan kenormalan sisaan mendasari penggunaan metode kuadrat terkecil (MKT) untuk menduga parameter model regresi. Disamping asumsi-asumsi tersebut, skala pengukuran variabel terikat juga harus diperhatikan, karena pada metode regresi linier variabel terikat setidaknya tidaknya diukur pada skala interval [5].

1.2 Data Tersensor

Data tersensor merupakan data dimana nilai dari variabel terikat yang diteliti tidak dapat diamati secara penuh atau tidak memberikan informasi yang lengkap. Hal ini disebabkan oleh batasan pengamatan yang dilakukan atau individu yang diamati keluar dari penelitian. Sedangkan variabel terikat yang memiliki nilai nol untuk sebagian pengamatan, sedangkan untuk pengamatan yang lain memiliki nilai tertentu yang bervariasi juga termasuk data tersensor [2]. Terdapat 3 macam tipe data tersensor [6], yaitu:

a. Sensor tipe I

Pada sensor tipe I, semua objek yang diteliti (n) masuk pengujian dalam waktu yang bersamaan, dan pengujian dihentikan setelah batas waktu t_0 yang ditentukan. Kelemahan dari sensor tipe I ini bisa terjadi sampai batas waktu t_0 yang ditentukan semua objek masih hidup sehingga tidak diperoleh data tahan hidup dari objek yang diuji.

b. Sensor tipe II

Pada sensor tipe II, semua objek yang diteliti (n) masuk pengujian dalam waktu yang bersamaan, dan pengujian dihentikan setelah mendapatkan r objek diantaranya mati, dengan $1 \leq r \leq n$. Kelemahan dari sensor tipe II ini waktu yang diperlukan untuk memperoleh r objek yang mati bisa jadi sangat panjang, tetapi pasti diperoleh data tahan hidup dari r objek tersebut.

c. Sensor Tipe III

Pada sensor tipe III, objek masuk dalam pengujian pada waktu yang tidak bersamaan selama periode waktu yang telah ditentukan. Beberapa objek yang mati/gagal sebelum pengamatan berakhir mempunyai data tahan hidup, sebagian lain masih tetap hidup sampai waktu pengujian berakhir, sebagian lagi ada yang masih hidup tetapi keluar dari pengujian (pada kasus objek berupa manusia/individu yang menjalani terapi tertentu).

Selain tipe data tersensor tersebut, terdapat juga data tersensor lainnya, diantaranya sensor kanan, sensor kiri, dan sensor interval. Semua tipe data tersensor tipe I, tipe II dan tipe III termasuk sensor kanan, sedangkan sensor kiri

terjadi jika kejadian yang diamati sudah terjadi pada suatu individu sebelum individu tersebut masuk dalam periode penelitian. Sedangkan sensor interval adalah sensor yang waktu terjadinya berada dalam suatu selang waktu tertentu yaitu antara selang antara waktu a dan waktu b.

1.3 Regresi Tobit

Regresi Tobit merupakan analisis regresi dimana nilai variabel terikatnya memiliki nilai masukan berupa sebagian data diskrit (bernilai nol) dan sebagian lagi data kontinu (bernilai tidak nol) [7]. Model stokastik yang mendasari Regresi Tobit dapat dinyatakan oleh hubungan berikut [8]:

$$y_i = \begin{cases} y_i^*, & y_i^* > 0 \\ 0, & y_i^* \leq 0 \end{cases}$$

dengan y_i^* adalah variabel laten dengan persamaan sebagai berikut:

$$y_i^* = x_i' \beta + u_i$$

dimana:

y_i^* : nilai variabel laten, $Y^* \sim N(x_i' \beta, \sigma^2)$,
 $U \sim N(0, \sigma^2)$

y_i : merupakan transformasi dari y_i^*

β : vektor dari parameter yang berukuran $k \times 1$

x_i : vektor dari bilangan yang diketahui berukuran $k \times 1$ untuk pengamatan i

x_i' : transpose dari x_i berukuran $1 \times k$

u_i : residual (*error*) yang berdistribusi normal independen dengan mean 0 dan variansi σ^2

i : 1, 2, 3, ..., n, n jumlah observasi

Jika Regresi Tobit hanya menggunakan observasi yang bernilai positif, atau $y_i = y_i^*$ dengan $y_i^* > 0$, atau $x_i' \beta + u_i > 0$, maka $\Pr(Y > 0) = \Pr(Y^* > 0)$

$$\begin{aligned} &= \Pr(U > -x_i' \beta) \\ &= \int_{-x_i' \beta}^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{u}{\sigma}\right)^2} du \end{aligned}$$

misalkan : $u = -\sigma t$ dengan $\sigma > 0$

$$: du = -\sigma dt$$

sehingga:

$$f(y_i > 0) = \int_{\frac{x_i' \beta}{\sigma}}^{-\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{t\sigma}{\sigma}\right)^2} -\sigma dt$$

$$= \int_{-\infty}^{\frac{x_i' \beta}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

$$f(y_i > 0) = \Phi\left(\frac{x_i' \beta}{\sigma}\right)$$

Sedangkan untuk c suatu konstanta

$$\begin{aligned} f(u|u > c) &= \frac{f(u)}{f(u > c)} \\ &= \frac{\frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{u}{\sigma}\right)^2}}{\int_c^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{u}{\sigma}\right)^2} du} \end{aligned}$$

Dengan memisalkan $u = \sigma t$ akan didapatkan

$$\begin{aligned} &= \frac{\frac{1}{\sigma} \phi\left(\frac{u}{\sigma}\right)}{\int_c^{\infty} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{t\sigma}{\sigma}\right)^2} \sigma dt} \\ &= \frac{\frac{1}{\sigma} \phi\left(\frac{u}{\sigma}\right)}{\int_{\frac{c}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt} \\ &= \frac{\frac{1}{\sigma} \phi\left(\frac{u}{\sigma}\right)}{1 - \int_{-\infty}^{\frac{c}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt} \\ &= \frac{\frac{1}{\sigma} \phi\left(\frac{u}{\sigma}\right)}{1 - \Phi\left(\frac{c}{\sigma}\right)} \end{aligned}$$

maka $f(u|u > 0) = \frac{\frac{1}{\sigma} \phi\left(\frac{u}{\sigma}\right)}{1 - \Phi(0)}$

sehingga dapat dituliskan:

$$E(U|u > c) = \sigma \frac{\phi\left(\frac{c}{\sigma}\right)}{1 - \Phi\left(\frac{c}{\sigma}\right)}$$

maka $E(U|u > 0) = \sigma \frac{\phi(0)}{1 - \Phi(0)}$

sehingga:

$$\begin{aligned} E(Y|y_i > 0) &= x_i' \beta + E(U|y_i > 0) \\ &= x_i' \beta + E(U|u > -x_i' \beta) \\ &= x_i' \beta + \sigma \frac{\phi\left(\frac{-x_i' \beta}{\sigma}\right)}{1 - \Phi\left(\frac{-x_i' \beta}{\sigma}\right)} \\ &= x_i' \beta + \sigma \frac{\phi\left(\frac{x_i' \beta}{\sigma}\right)}{1 - \Phi\left(\frac{-x_i' \beta}{\sigma}\right)} \end{aligned}$$

dimana $1 - \Phi\left(\frac{-x_i' \beta}{\sigma}\right) = \Phi\left(\frac{x_i' \beta}{\sigma}\right)$ sehingga:

$$E(Y|y_i > 0) = x'_i \beta + \sigma \frac{\phi\left(\frac{x'_i \beta}{\sigma}\right)}{\Phi\left(\frac{x'_i \beta}{\sigma}\right)}$$

$$= x'_i \beta + \sigma \lambda\left(\frac{x'_i \beta}{\sigma}\right)$$

dimana $\frac{\phi(w)}{\Phi(w)}$ disebut *Invers Mills Ratio*, dilambangkan dengan $\lambda(w)$

Nilai harapan dari model regresi Tobit:

$$E(Y) = \Pr(Y = 0) E(Y|y_i) + \Pr(Y > 0) E(Y|y_i > 0)$$

$$= \Pr(Y = 0).0 + \Pr(Y > 0)E(Y|y_i)$$

$$= \Pr(Y > 0)E(Y|y_i > 0)$$

Sehingga:

$$E(Y) = \Phi\left(\frac{x'_i \beta}{\sigma}\right) x'_i \beta + \sigma \phi\left(\frac{x'_i \beta}{\sigma}\right)$$

1.4 Fungsi Likelihood

Misalkan x_1, x_2, \dots, x_n merupakan sampel acak dengan fungsi peluang $f(x_i, \theta)$, $i = 1, 2, \dots, n$

Apabila L merupakan fungsi peluang bersama dari x_1, x_2, \dots, x_n dipandang sebagai fungsi dari θ dan x_1, x_2, \dots, x_n sebagai bilangan tertentu, maka

$L(\theta) = \prod_{i=1}^n f(x_i, \theta)$ disebut fungsi Likelihood. Fungsi Likelihood dari Regresi Tobit adalah:

$$L = \prod_0 \left[1 - \Phi\left(\frac{x'_i \beta}{\sigma}\right) \right] \prod_1 \sigma^{-1} \phi\left[\frac{y_i - x'_i \beta}{\sigma}\right]$$

1.5 Pendugaan Parameter

Metode yang digunakan untuk menduga parameter dalam regresi Tobit adalah Metode Kemungkinan Maksimum (*Method of Maximum Likelihood*). Prinsip dasar metode ini adalah untuk memperoleh penduga parameter dengan memaksimumkan fungsi likelihood sehingga diperoleh penduga yang konsisten dan efisien untuk sampel yang berukuran besar.

Berikut fungsi Likelihood Model Tobit:

$$L = \prod_0 \left[1 - \Phi\left(\frac{x'_i \beta}{\sigma}\right) \right] \prod_1 \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - x'_i \beta)^2} \right]$$

Fungsi ln Likelihood sebagai berikut:

$$\ln L = \sum_0 \ln \left[1 - \Phi\left(\frac{x'_i \beta}{\sigma}\right) \right] + \sum_1 \ln \left[\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - x'_i \beta)^2} \right]$$

Turunan pertama ln Likelihood terhadap β dan σ^2 adalah:

$$\frac{\partial \ln L}{\partial \beta} = -\frac{1}{\sigma} \sum_0 \frac{\phi\left(\frac{x'_i \beta}{\sigma}\right) x'_i}{1 - \Phi\left(\frac{x'_i \beta}{\sigma}\right)} + \frac{1}{\sigma^2} \sum_1 (y_i - x'_i \beta) x'_i$$

$$\frac{\partial \ln L}{\partial \sigma^2} = \frac{1}{2\sigma^3} \sum_0 \frac{(x'_i \beta) \phi\left(\frac{x'_i \beta}{\sigma}\right)}{1 - \Phi\left(\frac{x'_i \beta}{\sigma}\right)} - \frac{n_1}{2\sigma^2} +$$

$$\frac{1}{2\sigma^4} \sum_1 (y_i - x'_i \beta)^2$$

dengan $\frac{\partial \ln L}{\partial \beta} = 0$ dan $\frac{\partial \ln L}{\partial \sigma^2} = 0$ maka didapat nilai σ^2 dan β

$$\sigma^2 = \frac{\sum_1 (y_i - x'_i \beta) y_i}{n_1}$$

$$\beta = (X'_1 X_1)^{-1} X'_1 Y_1 - \sigma (X'_1 X_1)^{-1} X'_0 \gamma_0$$

$$= \beta_{LS} - \sigma (X'_1 X_1)^{-1} X'_0 \gamma_0$$

1.6 Pengujian Dugaan Paramater

Pengujian parameter dilakukan untuk mengetahui apakah variabel bebas yang dimasukkan dalam model regresi tobit mempunyai kontribusi nyata terhadap perubahan variasi dari variabel terikat. Pengujian ini meliputi uji serentak dan uji parsial [3].

a. Uji Serentak

Uji serentak bertujuan untuk mengetahui pengaruh variabel bebas secara bersama-sama terhadap variabel terikat, untuk mengujinya digunakan metode Likelihood Rasio.

Misalkan m merupakan seluruh parameter $\beta_1, \beta_2, \dots, \beta_m$, sedangkan r merupakan parameter yang akan diuji dengan $r \leq m$.

Hipotesis yang digunakan:

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

$$H_1 : \exists \beta_i \neq 0, \text{ minimal untuk satu nilai } i, i = 1, 2, \dots, r$$

Statistik uji yang digunakan:

$$\chi^2 = -2 \ln \left(\frac{L(0, \hat{\beta}_{r+1}, \hat{\beta}_{r+2}, \dots, \hat{\beta}_m)}{L(\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_m)} \right) \sim \chi^2(r)$$

H_0 ditolak jika $\chi^2 > \chi^2_{(r)}$

b. Uji Parsial

Uji Parsial bertujuan untuk mengetahui pengaruh variabel bebas secara individu terhadap variabel terikat. Untuk mengujinya digunakan *Wald test*.

Hipotesis yang digunakan:

$H_0 : \beta_j = 0$ (koefisien β_j tidak signifikan secara statistik)

$H_1 : \beta_j \neq 0$ (koefisien β_j signifikan secara statistik), $j = 1, 2, \dots, p$

Statistik uji yang digunakan:

$$W = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)}$$

dimana:

$\hat{\beta}_j$: penduga parameter β_j

$SE(\hat{\beta}_j)$: standar *error* dari β_j

H_0 ditolak jika nilai statistik uji $|W| < -Z_{\alpha/2}$ atau jika p -value $< \alpha$ yang berarti β_j berpengaruh signifikan terhadap variabel terikat.

1.7 Metode Akaike Information Criterion (AIC)

Metode yang digunakan untuk menduga parameter dalam regresi Tobit adalah Metode Kemungkinan Maksimum (*Method of Maximum Likelihood*), namun dalam sebagian besar aplikasi, nilai maksimum dari fungsi likelihood tidak diketahui, sehingga bisa digunakan langkah-langkah *heutiristik alternatif* seperti metode AIC (*Akaike Information Criterion*) [9].

Metode AIC adalah metode yang dikembangkan oleh Hirotugu Akaike pada tahun 1974 [9]. AIC tidak melakukan pengujian untuk satu model saja, tetapi dengan membandingkan antara beberapa model sehingga diperoleh model yang terbaik [10].

Bentuk umum metode AIC:

$$AIC = -2 \ln(L) + 2k$$

dimana:

L : nilai maksimum dari fungsi Likelihood

k : banyaknya parameter dalam model

Model yang terbaik menurut nilai AIC adalah model yang memberikan nilai AIC terkecil diantara model yang lain.

1.8 Pseudo R^2

Pengukuran *goodness of fit* yang konvensional dengan menggunakan nilai R^2 tidak dapat menjelaskan keragaman variabel terikat oleh variabel bebasnya. Ukuran lain yang serupa dengan R^2 yaitu *Pseudo- R^2* yang dapat melihat kemampuan model regresi Tobit dalam menerangkan variasi perubahan variabel terikat. Dalam program *E-views* nilai *Pseudo- R^2* berbentuk McFadden R^2 .

Berikut persamaan *Pseudo- R^2* untuk regresi Tobit [11]:

$$Pseudo R^2 = 1 - \frac{\sum_{i=1}^n \hat{u}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

dengan $\hat{u}_i = Y_i - \beta' X_t - \sigma \left(\frac{\beta' X_i}{\sigma} \right)$

1.9 Pola Konsumsi Susu

Susu merupakan sumber yang kaya mineral seperti kalsium, kalium, natrium, fosfor, vitamin, dan nutrisi lainnya. Susu dan produk-produk olahannya merupakan asupan nutrisi yang penting bagi perkembangan pertumbuhan manusia.

Semakin berkembangnya industri susu didunia mendorong pertumbuhan produk olahan susu cair pabrik, salah satunya adalah susu UHT. Susu UHT merupakan susu cair segar yang diolah menggunakan pemanasan dengan suhu tinggi dan dalam waktu yang sangat singkat untuk membunuh mikroba dan dikemas secara higienis dengan menggunakan kemasan aseptik [12].

Kesadaran masyarakat akan pentingnya mengonsumsi susu masih sangat rendah, hal ini mengakibatkan rumah tangga tidak mengalokasikan pengeluaran untuk susu (*Zero Consumption*). *Zero Consumption* adalah kondisi dimana ada rumah tangga yang tidak mengonsumsi susu, sedangkan rumah tangga yang lain mengonsumsi dengan jumlah yang bervariasi.

2. Metode Penelitian

Data yang digunakan pada penelitian ini adalah data sekunder. Data sekunder dalam penelitian ini merupakan data yang diperoleh melalui Survei Sosial Ekonomi Nasional

(SUSENAS) tahun 2012 yang dilakukan oleh Badan Pusat Statistik (BPS) di Provinsi Bali dengan jumlah sampel sebanyak 615 rumah tangga, meliputi rumah tangga yang mengonsumsi susu cair pabrik dan rumah tangga yang tidak mengonsumsi susu cair pabrik.

Adapun variabel-variabel yang digunakan dalam penelitian ini adalah:

1. Variabel Terikat
(y_i): jumlah pengeluaran untuk konsumsi susu cair pabrik dimana y_i bernilai 0 untuk rumah tangga yang tidak mengonsumsi susu dan bernilai y_i^* untuk rumah tangga yang mengonsumsi susu cair pabrik.
2. Variabel Bebas
 x_1 : Pendidikan terakhir Kepala Rumah Tangga.
 x_2 : Pekerjaan Kepala Rumah Tangga
 x_3 : Umur Kepala Rumah Tangga
 x_4 : Jumlah Pengeluaran untuk konsumsi makanan
 x_5 : Jumlah anggota rumah tangga
 x_6 : Pendapatan rumah tangga

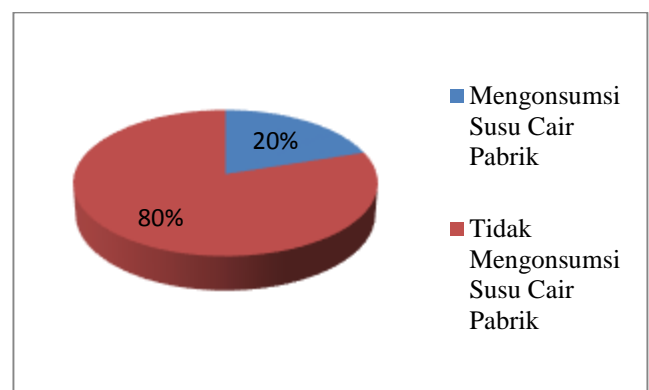
Pada penelitian ini, pemodelan akan dilakukan dengan menggunakan regresi Tobit. Proses analisis data pada penelitian ini menggunakan *software SPSS 20* dan *E-views 5.1*. Langkah-langkah analisis untuk memperoleh model regresi tobit adalah:

1. Membentuk model regresi Tobit dari $y = f(x_1, x_2, x_3, x_4, x_5, x_6)$
2. Melakukan uji secara serentak menggunakan metode uji rasio likelihood.
3. Menguji signifikansi parameter penduga dari model menggunakan *Wald test*. Apabila terdapat parameter penduga yang tidak signifikan, maka variabel tersebut dikeluarkan dari model, selanjutnya kembali melakukan langkah 1 dengan menggunakan parameter penduga yang signifikan.
4. Membentuk model-model regresi Tobit dari banyaknya penduga parameter yang signifikan.

5. Menghitung nilai AIC dari masing-masing model regresi Tobit yang dapat terbentuk dari banyaknya penduga parameter yang signifikan.
6. Memilih nilai AIC terkecil dari semua model sebagai model regresi Tobit terbaik.
7. Menghitung nilai *Pseudo-R²* dari model dengan nilai AIC terkecil untuk melihat besar keragaman variabel terikat yang mampu dijelaskan oleh variabel bebas.

3. Hasil dan Pembahasan

Analisis deskriptif digunakan untuk memberikan gambaran umum mengenai data yang telah diperoleh. Variabel bebas pada penelitian ini adalah faktor-faktor yang memengaruhi jumlah pengeluaran untuk konsumsi susu cair pabrik yaitu pendidikan terakhir Kepala Rumah Tangga (KRT) (x_1), pekerjaan KRT (x_2), umur KRT (x_3), jumlah pengeluaran untuk konsumsi makanan (x_4), jumlah anggota rumah tangga (x_5), dan pendapatan rumah tangga (x_6), sedangkan sebagai variabel terikat adalah jumlah pengeluaran untuk konsumsi susu cair pabrik (y). Data tersebut diperoleh dari hasil Survei Sosial Ekonomi Nasional (SUSENAS) pada tahun 2012 yang dilakukan oleh Badan Pusat Statistik (BPS) Provinsi Bali. Data SUSENAS tersebut terdiri dari 615 rumah tangga, dimana jumlah rumah tangga yang mengonsumsi susu cair pabrik sebanyak 123 rumah tangga dan 492 rumah tangga yang tidak mengonsumsi susu cair pabrik.



Gambar 3.1 Diagram Pie Persentase Rumah Tangga untuk Konsumsi Susu Cair Pabrik.

Gambar 3.1 menggambarkan persentase rumah tangga yang mengonsumsi susu cair pabrik sebesar 20% sedangkan rumah tangga yang tidak mengonsumsi susu cair pabrik sebesar 80%, sehingga dapat dikatakan konsumsi untuk susu cair pabrik oleh rumah tangga di Provinsi Bali masih sangat rendah.

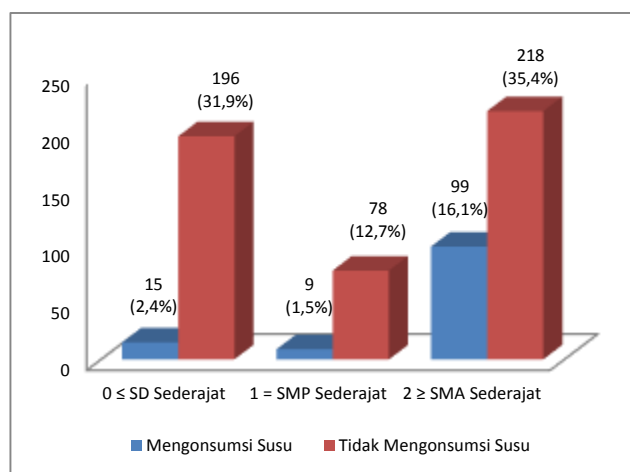
Tabel 3.1 Deskripsi Variabel Bebas yang Memengaruhi Rumah Tangga Mengonsumsi Susu Cair Pabrik (Data Kontinu)

| No | | | Rumah Tangga | |
|----|---|-------------|----------------|----------------|
| | | | Konsumsi Susu | Tidak Konsumsi |
| 1 | Umur Kepala Rumah Tangga (Tahun) | Minimum | 26 | 18 |
| | | Maksimum | 73 | 91 |
| | | Rata-rata | 42,62 | 44,59 |
| | | St. Deviasi | 9,09 | 11,41 |
| 2 | Pengeluaran untuk Konsumsi Makanan (Rupiah) | Minimum | 26.820.856,92 | 14.820.000 |
| | | Maksimum | 296.217.256,90 | 465.548.571,3 |
| | | Rata-rata | 100.984.384,11 | 71.142.989,21 |
| | | St. Deviasi | 46.909.861,07 | 39.895.611,92 |
| 3 | Anggota Rumah Tangga (Orang) | Minimum | 3 | 3 |
| | | Maksimum | 10 | 8 |
| | | Rata-rata | 4,5 | 4,19 |
| | | St. Deviasi | 1,31 | 1,13 |
| 4 | Pendapatan (Rupiah) | Minimum | 54.565.110,6 | 24.420.378,84 |
| | | Maksimum | 1.039.636.991 | 3.822.799.809 |
| | | Rata-rata | 273.064.151,02 | 164.722.143,15 |
| | | St. Deviasi | 183.072.082,69 | 235.419.897,89 |

Berdasarkan Tabel 3.1 rata-rata umur kepala rumah tangga yang mengonsumsi susu cair pabrik adalah 42,62 tahun dengan nilai standar deviasi sebesar 9,09, sedangkan yang tidak mengonsumsi rata-rata berumur 44,59 tahun dengan nilai standar deviasi sebesar 11,41. Berdasarkan nilai standar deviasi keragaman umur kepala rumah tangga pada rumah tangga yang tidak mengonsumsi susu lebih besar daripada rumah tangga yang mengonsumsi. Untuk rumah tangga yang mengonsumsi susu cair pabrik rata-rata mengeluarkan biaya untuk mengonsumsi makanan sebesar Rp 100.984.384,11 dan untuk rumah tangga yang tidak mengonsumsi susu cair pabrik rata-rata pengeluaran untuk

mengonsumsi makanan sebesar Rp 71.142.989,21 dengan pengeluaran tertinggi sebesar Rp 465.548.571,30, ini mengindikasikan bahwa rumah tangga lebih memprioritaskan pengeluarannya untuk memenuhi kebutuhan pokok.

Dilihat dari jumlah anggota rumah tangga, baik rumah tangga yang mengonsumsi susu cair pabrik maupun yang tidak mengonsumsi memiliki jumlah rata-rata anggota keluarga sebanyak 4 orang, dengan pendapatan keluarga per tahun rata-rata sebesar Rp 273.064.151,02 pada rumah tangga yang mengonsumsi susu cair pabrik dan Rp 164.722.143,15 per tahun pada rumah tangga yang tidak mengonsumsi susu cair pabrik.



Gambar 3.2 Grafik Persentase Rumah Tangga Dilihat Dari Pendidikan Kepala Rumah Tangga (Data Kategorik)

Berdasarkan Gambar 3.2 terlihat untuk rumah tangga yang mengonsumsi susu cair pabrik pada tingkat pendidikan setara atau lebih tinggi dari SMA sebesar 16,1%, ini mengindikasikan bahwa kepala rumah tangga yang pendidikannya setara atau lebih tinggi dari SMA kecenderungan mengonsumsi susu cair pabrik lebih besar. Gambar 3.2 menunjukkan bahwa keinginan rumah tangga di Provinsi Bali untuk mengonsumsi susu masih sangat rendah, terlihat dari persentase rumah tangga yang tidak mengonsumsi susu terbesar pada rumah tangga yang tingkat pendidikan kepala rumah tangganya setara atau lebih tinggi dari SMA yaitu 35,4%.

Tabel 3.2 mendeskripsikan rumah tangga yang mengonsumsi maupun yang tidak mengonsumsi susu cair pabrik berdasarkan pekerjaan kepala rumah tangga. Berdasarkan bidang pekerjaan kepala rumah tangga persentase rumah tangga yang tidak mengonsumsi susu lebih besar dibandingkan dengan yang mengonsumsi susu yaitu pada kepala rumah tangga yang tidak bekerja, kepala rumah tangga yang bekerja pada bidang pertanian, industri dan perdagangan. Sementara pada bidang pekerjaan yang lain persentase rumah tangga yang mengonsumsi susu lebih besar.

Tabel 3.2 Persentase Rumah Tangga Dilihat dari Bidang Pekerjaan (Data Kategorik)

| No | Pekerjaan | Rumah Tangga | | | | Total |
|----|-------------------|---------------|------|----------------|------|-------|
| | | Konsumsi Susu | | Tidak Konsumsi | | |
| | | N | % | N | % | |
| 1 | 0 = Tidak Bekerja | 6 | 4,9 | 27 | 5,5 | 33 |
| 2 | 1 = Pertanian | 4 | 3,3 | 42 | 8,5 | 46 |
| 3 | 2 = Industri | 6 | 4,9 | 103 | 20,9 | 109 |
| 4 | 3 = Perdagangan | 5 | 4,1 | 57 | 11,6 | 62 |
| 5 | 4 = Jasa | 65 | 52,8 | 158 | 32,1 | 223 |
| 6 | 5 = PNS & BUMN | 11 | 8,9 | 33 | 6,7 | 44 |
| 7 | 6 = Lain-lain | 26 | 21,1 | 72 | 14,6 | 98 |
| | Total | 123 | 100 | 492 | 100 | 615 |

3.1 Model Dugaan

Model awal regresi Tobit dibentuk dari 6 variabel bebas yang terdiri dari Kepala Rumah Tangga (KRT) (x_1), pekerjaan KRT (x_2), umur KRT (x_3), jumlah pengeluaran untuk konsumsi makanan (x_4), jumlah anggota rumah tangga (x_5), pendapatan rumah tangga (x_6) sedangkan variabel terikat adalah jumlah pengeluaran untuk konsumsi susu oleh rumah tangga (y). Dengan model yang akan dibentuk dijabarkan pada persamaan berikut:

$$\hat{y}^* = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon$$

Dengan bantuan *software E-views* diperoleh model seperti persamaan berikut:

$$\hat{y}^* = -3226050 + 499586,8 x_1 + 59993,54 x_2 - 9875,311 x_3 + 0,012558 x_4 + 63651,03 x_5 + 0,000397 x_6$$

3.2 Uji Serentak

Setelah memperoleh model, akan dilakukan uji secara serentak yang bertujuan untuk mengetahui pengaruh variabel bebas secara bersama-sama dan diuji menggunakan metode likelihood rasio. Hipotesis yang digunakan:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_6 = 0$$

H_1 : minimal ada salah satu β_i yang tidak sama dengan nol, $i = 1, 2, \dots, 6$

H_0 ditolak jika $\chi^2 > \chi^2_{(6)}$ atau $p\text{-value} < \alpha$ yang berarti ada salah satu atau lebih β_i yang berpengaruh signifikan terhadap variabel tak bebas.

Tabel 3.3 menunjukkan nilai χ_{hitung}^2 (*log likelihood ratio*) sebesar 108,8757 dengan probabilitas sebesar 0,000000. Apabila dibandingkan dengan nilai tabel distribusi Chi-kuadrat dengan taraf signifikansi $\alpha = 0,05$ dan derajat bebas 6, maka nilai *log likelihood ratio* lebih besar dari nilai tabel Chi-kuadrat ($\chi_{(1-0,05);6}^2 = 12,592$). Keputusan yang diambil adalah menolak H_0 , sehingga minimal ada satu variabel bebas yang memengaruhi jumlah pengeluaran untuk konsumsi susu cair pabrik

Tabel 3.3 Uji Serentak

| Redundant Variabel : $x_1, x_2, x_3, x_4, x_5, x_6$ | | | |
|---|----------|----------------------|------|
| F-hitung | 22,60218 | Prob. F (6,607) | 0,00 |
| Rasio Log Likelihood | 108,8757 | Prob. Chi-Square (6) | 0,00 |
| Data Sensor Kiri | 492 | Data Sensor Kanan | 0 |
| Data Tak Tersensor | 123 | Total Data | 615 |

3.3 Uji Signifikansi Parameter

Setelah melakukan uji serentak, selanjutnya dilakukan uji signifikansi parameter model dengan *Wald test*. *Wald test* dilakukan untuk melihat peranan masing-masing variabel bebas terhadap variabel terikat. Hipotesis yang digunakan:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j= 1, 2, \dots, 6$$

H_0 ditolak jika nilai statistik uji $|W| < -Z_{\alpha/2}$ atau jika p-value $< \alpha$ yang berarti β_j berpengaruh signifikan terhadap variabel terikat.

Tabel 3.4 Uji Signifikansi Parameter Model

| Variabel | Koefisien | Std. Error | z-hitung | Prob. |
|----------|-----------|------------|-----------|--------|
| c | -3226050 | 638107,8 | -5,055650 | 0,0000 |
| x_1 | 499586,8 | 128009,4 | 3,902735 | 0,0001 |
| x_2 | 59993,54 | 61027,78 | 0,983053 | 0,3256 |
| x_3 | -9875,311 | 9666,469 | -1,021605 | 0,3070 |
| x_4 | 0,012558 | 0,002308 | 5,442217 | 0,0000 |
| x_5 | 63651,03 | 77814,99 | 0,817979 | 0,4134 |
| x_6 | 0,000397 | 0,000396 | 1,003858 | 0,3154 |

Berdasarkan Tabel 3.4 diperoleh Z_{hitung} beserta nilai probabilitas Z_{hitung} untuk masing-masing variabel bebas. Dengan mengambil α sebesar 0,05, nilai probabilitas $Z_{hitung} > \alpha$ untuk koefisien $\beta_2, \beta_3, \beta_5, \beta_6$ sehingga H_0 diterima atau koefisien tidak signifikan. Oleh karena itu variabel bebas yang tidak signifikan dikeluarkan dari model, sehingga perlu dilakukan kembali uji secara serentak dengan variabel bebas yang signifikan dengan probabilitas $Z_{hitung} < \alpha$ yaitu variabel (x_1) dan (x_4). Model yang terbentuk adalah sebagai berikut:

$$\hat{y}^* = -3314724 + 565429,7 x_1 + 0,014278 x_4$$

Tabel 3.5 menunjukkan nilai χ_{hitung}^2 (log likelihood ratio) sebesar 105,109 dengan probabilitas sebesar 0,000000. Apabila dibandingkan dengan nilai tabel distribusi Chi-kuadrat dengan taraf signifikansi $\alpha = 0,05$ dan derajat bebas 2, maka nilai log likelihood ratio lebih besar dari nilai tabel Chi-kuadrat ($\chi_{(1-0,05);2}^2 = 5,991$). Keputusan yang diambil adalah menolak H_0 , sehingga minimal ada satu variabel bebas yang memengaruhi jumlah pengeluaran untuk konsumsi susu cair pabrik.

Selanjutnya kembali dilakukan uji signifikansi parameter model dengan menggunakan *Wald test*.

Tabel 3.5 Uji Serentak Model Kedua

| Redundant Variabel : x_1, x_4 | | | |
|---------------------------------|----------|----------------------|------|
| F-hitung | 62,34103 | Prob. F (2,611) | 0,00 |
| Rasio Log Likelihood | 105,109 | Prob. Chi-Square (2) | 0,00 |
| Data Sensor Kiri | 492 | Data Sensor Kanan | 0 |
| Data Tak Tersensor | 123 | Total Data | 615 |

Dengan nilai $\alpha = 0,05$ dari Tabel 3.6 dapat dilihat probabilitas $Z_{hitung} < \alpha$ untuk koefisien β_1 dan β_2 , sehingga H_0 ditolak dan variabel (x_1) dan (x_4) signifikan.

Tabel 3.6 Uji Signifikansi Parameter Model Kedua

| Variabel | Koefisien | Std. Error | z-hitung | Prob. |
|----------|-----------|------------|----------|--------|
| c | -3314724 | 310516,5 | 10,67487 | 0,0000 |
| x_1 | 565429,7 | 120926,8 | 4,675801 | 0,0000 |
| x_4 | 0,014278 | 0,001922 | 7,429007 | 0,0000 |

3.4 Pemilihan Model dengan Metode AIC

Kriteria kesesuaian model pada regresi Tobit dengan menggunakan metode AIC dilakukan dengan membandingkan semua kemungkinan model yang bisa terbentuk dari variabel bebas yang membentuk model. Model yang terbaik menurut metode AIC adalah model dengan nilai AIC terendah.

Nilai AIC yang ditunjukkan pada Tabel 3.7 memperlihatkan semua kemungkinan model yang dapat dibentuk dari variabel bebas. Dari keempat model yang didapat, model yang terbaik menurut metode AIC adalah model 4, hal ini ditunjukkan oleh nilai AIC untuk model 4 terendah (6,751572) dibandingkan dengan nilai AIC untuk model yang lain.

Berdasarkan pengujian yang dilakukan diperoleh model regresi Tobit terbaik untuk konsumsi susu cair pabrik oleh rumah tangga yaitu:

$$\hat{y}^* = -3314724 + 565429,7 x_1 + 0,014278 x_4$$

dimana $\hat{y} = 0$ jika $\hat{y}^* \leq 0$ dan $\hat{y} = \hat{y}^*$ jika $\hat{y}^* > 0$

Tabel 3.7. Nilai AIC dari seluruh Kemungkinan Model

| No | Koefisien | | Pseudo- R^2 | AIC |
|----|-----------|----------|-----------------|-----------------|
| 1 | c | -1640136 | 0,001822 | 6,915977 |
| 2 | c | -2739493 | 0,028748 | 6,836334 |
| | x_1 | 863622,1 | | |
| 3 | c | -2769129 | 0,130810 | 6,790412 |
| | x_4 | 0,016823 | | |
| 4 | c | -3314724 | 0,167965 | 6,751572 |
| | x_1 | 565429,7 | | |
| | x_4 | 0,014278 | | |

3.5 Pseudo R^2

Seperti analisis regresi pada umumnya, analisis regresi Tobit juga memiliki nilai untuk melihat seberapa besar varian data mampu dijelaskan oleh model. Menurut Bierens (2004), pengukuran *goodness of fit* pada regresi Tobit menggunakan *Pseudo- R^2* , pada *software E-views* nilai *Pseudo- R^2* berbentuk McFadden R^2 yang juga bernilai diantara 0 dan 1.

Pada Tabel 3.7 dapat dilihat model terbaik dari analisis regresi Tobit memiliki nilai *Pseudo- R^2* sebesar 0,167965, yang berarti bahwa model yang diperoleh hanya mampu menjelaskan faktor-faktor rumah tangga untuk mengonsumsi susu cair pabrik sebesar 16,79%.

4. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah dilakukan, adapun simpulan yang dapat diambil dari penelitian ini, yaitu:

1. Berdasarkan analisis yang telah dilakukan, didapatkan model terbaik dari analisis regresi Tobit pada konsumsi susu cair pabrik oleh rumah tangga di Provinsi Bali, yaitu: $\hat{y}^* = -3314724 + 565429,7 x_1 + 0,014278 x_4$

2. Dari model terbaik yang dihasilkan, dapat dijelaskan bahwa faktor-faktor yang memengaruhi jumlah pengeluaran untuk konsumsi susu cair pabrik pada rumah tangga adalah pendidikan terakhir kepala rumah tangga (x_1) dan jumlah pengeluaran untuk konsumsi makanan (x_4).
3. Dengan nilai *Pseudo- R^2* pada model masih sangat kecil yaitu sebesar 16,79%, model yang didapat belum dapat mewakili faktor-faktor yang memengaruhi konsumsi susu cair pada rumah tangga.

Daftar Pustaka

- [1] Gujarati, D.N. 1995. *Basic Econometric*. 3rd Edition. McGraw-Hill International Edition, Economic Series.
- [2] Greene, W. H. 2003. *Econometrics Analysis*, 5th Edition. Prentice Hall. New Jersey.
- [3] Tobin, J. 1958. Estimation of Relationships for Limited Dependent Variabel. *Journal Econometrica*, Vol.26, No.1, pp 24-36.
- [4] Suhardi, I.Y. dan R. Llewelyn. 2001. Penggunaan Model Regresi Tobit untuk Menganalisa Faktor-Faktor yang Berpengaruh Terhadap Kepuasan Konsumen untuk Jasa Pengangkutan Barang. *Jurnal Manajemen & Kewirausahaan Vol. 3*, hal. 106-112.
- [5] Neter, John.,W.Wasserman dan M.H.Kutner. 1997. *Model Linear Terapan Buku II. Analisis Regresi Ganda*. Diterjemahkan oleh Bambang Sumantri. Jurusan Statistika FMIPA IPB.
- [6] Lee, E.T. and Wang Jhon Wenyu. 2003. *Statistical Methods for Survival Data Analysis*. 3rd Edition. John Wiley & Sons. New York.
- [7] Amemiya, T. 1984. Tobit Models : A Survey. *Journal of Econometrics*. Volume 24, Issues 1-2, pp 3-61. Stanford University. Stanford.
- [8] Fair, R. C. 1977. A Note on the Computation of the Tobit Estimator.

Journal Econometrica, Vol. 45, No.7
(Oct., 1977), pp. 1723-1727.

- [9] Jedidi, Kamel. Ramaswamy, Venkaram and Desarbo, Wayne S. 1993. A Maximum Likelihood Method For Latent Class Regression Involving A Censored Dependent Variable. *Journal of Psychometrica*. Vol. 58. pp 375-394.
- [10] Bozdogan, Hamparsum. 2000. Akaike's Information Criterion and Recent Developments in Information Complexity. *Journal of Mathematical Psychology* 44, 62-91.
- [11] Bierens, H. J. (2004). *The Tobit Model*. Retrieved Juni 06, 2013, from <http://econ.la.psu.edu/.. / Tobit. PDF/html>
- [12] Gedam, Kishor. Prasad, Rajendra and V.K. Vijay. 2007. The Study on UHT Processing of Milk : A Versatile Option for Rural Sector. *World Journal of Dairy & Food Science* 2 (2) : 49-53.