

## Analisis Regresi Tobit pada Jumlah Kasus Kematian HIV AIDS di Provinsi Jawa Barat

Fiore Rosie Kestana\*, Teti Sofia Yanti

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

\*fiorerosiekestana@gmail.com, tetisofiyanti@unisba.ac.id

**Abstract.** There are several types of regression analysis, one of which is linear regression. In fact, in the field there are many cases where the dependent variable doesn't complete information or is called censored data. Regression takes account censored data on the dependent variable is tobit regression. Parameter estimation in tobit regression uses the Maximum Likelihood Estimation method. This research takes data related to cases of death due to HIV/AIDS. In June 2022, West Java in third place with the highest of HIV/AIDS cases in Indonesia. The number of cases of AIDS death in Indonesia, one of which in West Java, the probability of this occurring is 0 to not death. This data is called censored data. This research aims was to look at the factors that influence the number of AIDS death in West Java with the the variables  $X$  of the number of cases of HIV positive blood donor  $X_1$ , the number of cases of syphilis  $X_2$  and the number of condom  $X_3$  using a significance level of 5%. Based on the results, it is stated that for every people 10 people increasing the number of cases of syphilis, the number of deaths due to AIDS will increase by 5 people and every time the number of condom uses increases by  $10^{1000}$ , the number of deaths due to AIDS will decrease by 1,2 cases. The variables that influence the number of deaths due to AIDS are the number of cases of syphilis  $X_2$  of the number of condom  $\log X_3$ .

**Keywords:** *Censored Data, Tobit Regression, Maximum Likelihood Estimation*

**Abstrak.** Terdapat beberapa macam analisis regresi salah satunya yaitu analisis regresi linier. Faktanya, dilapangan banyak ditemukan kasus dimana variabel terikatnya tidak memberikan informasi secara lengkap atau disebut data tersensor. Regresi yang memperhatikan data tersensor pada variabel terikatnya yaitu analisis regresi tobit. Metode ini digunakan untuk mengatasi ketika variabel terikatnya bersifat terbatas. Pendugaan parameter pada regresi tobit yaitu menggunakan metode *Maksimum Likelihood Estimation*. Penelitian ini mengambil data terkait kasus kematian oleh HIV/AIDS. Provinsi Jawa Barat sampai dengan Juni 2022 terdapat pada urutan ketiga dengan jumlah kasus HIV/AIDS terbanyak di Indonesia. Virus HIV/AIDS merupakan salah satu penyakit dengan peringkat atas dalam menyebabkan kematian. Dalam jumlah kasus kematian AIDS di Indonesia salah satunya di Jawa Barat kemungkinan yang terjadi adalah 0 sampai tak hingga kematian. Struktur data seperti ini disebut data tersensor. Pada penelitian ini bertujuan untuk melihat faktor-faktor yang mempengaruhi pada jumlah kematian AIDS di Jawa Barat dengan variabel  $X$  nya terdiri dari jumlah kasus donor darah positif HIV  $X_1$ , jumlah kasus penyakit *sifilis*  $X_2$  dan jumlah penggunaan kondom  $X_3$  dengan menggunakan taraf nyata 5%. Berdasarkan hasil penelitian menyatakan bahwa setiap penambahan jumlah kasus penyakit *sifilis* 10 orang maka jumlah kematian akibat AIDS akan meningkat sebesar 5 orang dan setiap jumlah penggunaan kondom bertambah  $10^{1000}$  maka jumlah kasus kematian akibat AIDS akan menurun sebesar 1,2 kasus. Selain itu, berdasarkan hasil penelitian ini bahwa variabel yang mempengaruhi jumlah kematian akibat AIDS adalah jumlah kasus penyakit *sifilis*  $X_2$  dan logaritma jumlah penggunaan kondom  $\log X_3$ .

**Kata Kunci:** *Data Tersensor, Regresi Tobit, Estimasi Maksimum Likelihood*

## A. Pendahuluan

Metode statistika dalam perkembangannya memiliki berbagai macam pendekatan, salah satunya yang paling umum digunakan adalah analisis regresi linier. Menurut Priyatno (2013) analisis regresi linier bertujuan untuk mengidentifikasi hubungan antara variabel terikat dengan satu atau lebih variabel bebas dengan menggunakan persamaan regresi linier. Dimana variabel terikatnya merupakan variabel yang berskala kontinu dan nilai pada variabel terikatnya tidak terbatas. Namun, faktanya di lapangan tidak sedikit ditemukan kasus dimana variabel terikatnya yang diteliti tidak memberikan informasi secara lengkap disebabkan adanya batasan waktu penelitian, alat atau batasan nilai. Kondisi variabel terikat yang seperti ini disebut dengan data tersensor (1). Terdapat beberapa macam data tersensor yaitu tersensor kanan, kiri dan interval. Tersensor kanan merupakan penelitian telah berakhir sehingga kemungkinan terdapat subjek yang belum mengalami kejadian dan tersensor kiri terjadi ketika apabila subjek telah mengalami kejadian sebelum periode waktu tertentu atau sebelum penelitian dimulai sedangkan tersensor interval adalah jika informasi mengenai terjadinya suatu kejadian tidak diketahui secara pasti hanya diketahui terjadi pada suatu interval tertentu (2). Kasus seperti ini memerlukan metode atau pendekatan yang tepat untuk analisisnya. Regresi yang memperhatikan data tersensor pada variabel terikatnya adalah regresi tobit.

Regresi tobit pertama kali diusulkan oleh Tobin pada tahun 1958, penelitian tersebut menganalisis antara pengeluaran rumah tangga barang tahan lama dan pendapatan rumah tangga. Diketahui karakteristik dari data variabel terikatnya terdapat beberapa observasi dimana pengeluarannya nol (Amemiya, 1984). Hal ini berpengaruh terhadap hasil analisis regresinya, dimana jika tetap menggunakan analisis regresi linier dengan metode *Ordinary Least Square* (OLS) menyebabkan hasilnya tidak konsisten pada parameter intercept ( $\beta_0$ ) dan parameter koefisien regresi ( $\beta$ ) serta memberikan hasil nilai penduga yang bias (Jeryana dkk, 2014). Oleh karena itu, masalah ini dapat ditangani dengan menggunakan analisis regresi tobit. Metode ini digunakan untuk mengatasi masalah ketika variabel terikatnya bersifat terbatas (3). Tobin (1958) juga mengasumsikan bahwa variabel terikatnya memiliki sejumlah nilai yang dapat dikelompokkan pada suatu nilai pembatas. Umumnya nilai pembatas yang digunakan adalah nol (4). Pendugaan parameter pada analisis regresi tobit dapat diestimasi dengan berbagai metode, salah satunya dengan metode *Maximum Likelihood Estimation* (MLE).

Telah dipaparkan sebelumnya bahwa faktanya di lapangan sering ditemukan kasus data tersensor salah satunya yaitu terkait kasus kematian yang disebabkan oleh Human Immunodeficiency Virus (HIV) atau Acquired Immunodeficiency Syndrome (AIDS). HIV merupakan virus yang menyerang sistem kekebalan tubuh manusia sehingga menyebabkan lemahnya kemampuan tubuh untuk melawan infeksi dan penyakit. Sedangkan AIDS adalah kondisi dimana HIV telah pada tahap akhir infeksi. AIDS juga merupakan salah satu penyakit infeksi dengan peringkat atas dalam menyebabkan kematian (5). Indonesia merupakan salah satu negara di Asia Tenggara dengan penambahan kasus HIV/AIDS tercepat dengan estimasi peningkatan angka kejadian infeksi lebih dari 36% (6). Berdasarkan hasil laporan Perkembangan HIV/AIDS dan Penyakit Infeksi Menular Seksual (PIMS) di Indonesia 2022 (7) menyatakan bahwa dari 503 kabupaten/kota jumlah kumulatif dari tahun 2009 sampai dengan Juni 2022 terkait kasus HIV terdapat 478.784 orang dan kasus AIDS terdapat 139.500 orang. Ada lima provinsi dengan jumlah kasus HIV tertinggi salah satunya Jawa Barat menduduki peringkat ketiga dengan jumlah kasus sebanyak 54.897.

Dalam jumlah kasus kematian AIDS di Indonesia salah satunya kemungkinan yang terjadi adalah 0 sampai tak hingga kematian. Hal tersebut menunjukkan bahwa jumlah kasus kematian HIV/AIDS datanya bersifat tersensor. Pada penelitian ini mengamati jumlah kasus kematian AIDS di Jawa Barat pada tahun 2018 dimana data tersebut termasuk data tersensor kiri karena pada pengambilan datanya sebelum bulan Desember tahun 2018, hal ini dilihat berdasarkan laporan Perkembangan HIV/AIDS dan Penyakit Infeksi Menular Seksual (PIMS) di Indonesia (2018) dimana dilaporkan atau pelaksanaan penelitiannya pada bulan Desember tahun 2018.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Apa saja faktor-faktor yang mempengaruhi pada jumlah kasus

kematian akibat penyakit HIV/AIDS di Provinsi Jawa Barat 2018 dengan menggunakan analisis regresi tobit?". Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

Untuk mengetahui permodelan analisis regresi tobit pada jumlah kasus kematian akibat penyakit HIV/AIDS di Provinsi Jawa Barat tahun 2018.

Untuk mengidentifikasi faktor-faktor yang mempengaruhi pada jumlah kasus kematian akibat penyakit HIV/AIDS di Provinsi Jawa Barat tahun 2018 dengan menggunakan analisis regresi tobit.

## B. Metodologi Penelitian

Peneliti menggunakan penerapan metode analisis regresi tobit dalam jumlah kasus kematian akibat HIV/AIDS di Provinsi Jawa Barat tahun 2018. Data yang digunakan pada penelitian ini merupakan data sekunder. Data terdiri dari 27 pengamatan berdasarkan Kabupaten/Kota dan data diperoleh dari *website* resmi *Open Data Jabar*. Variabel dalam penelitian ini terdiri dari satu variabel terikat ( $Y$ ) yaitu jumlah angka kematian akibat AIDS di Provinsi Jawa Barat dan dua variabel bebas ( $X$ ) dimana  $X_1$  yaitu jumlah kasus donor darah terindikasi positif HIV di Provinsi Jawa Barat,  $X_2$  yaitu jumlah kasus penyakit *Sifilis* dan  $X_3$  yaitu jumlah penggunaan kondom di Provinsi Jawa Barat.

## C. Hasil Penelitian dan Pembahasan

### Analisis Deskriptif

Berikut adalah penelitian mengenai analisis deskriptif pada data jumlah kasus kematian akibat AIDS Di Provinsi Jawa Barat.

**Tabel 1.** Analisis Statistika Deskriptif

Variabel	Rata-rata	Minimum	Maksimum	Simpangan Baku
Jumlah Angka Kematian akibat AIDS ( $Y$ )	2,3704	0	15	4,1058
Jumlah Kasus Donor Darah Positif HIV ( $X_1$ )	64,5556	0	550	122,2615
Jumlah Kasus Penyakit <i>Sifilis</i> ( $X_2$ )	20,2593	3	36	9,1928
Jumlah Penggunaan Kondom ( $X_3$ )	5639,33	435	25827	6014,45

Berdasarkan pada Tabel 1 dapat dilihat bahwa 27 kabupaten/kota di Jawa Barat pada tahun 2018 dengan memiliki rata-rata jumlah kasus kematian akibat AIDS sebanyak  $2,3704 \approx 2$  kasus kematian dan kasus kematian akibat AIDS yang tertinggi terdapat di Kota Cimahi dengan angka kematiannya mencapai 15 kasus kematian. Selain itu, terdapat juga enam belas kabupaten/kota yang tidak memiliki kasus kematian akibat AIDS diantaranya adalah: Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Ciamis, Kabupaten Cirebon, Kabupaten Sumedang, Kabupaten Subang, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Pangandaran Kota Bogor, Kota Bandung, Kota Depok, Kota Bekasi dan Kota Banjar.

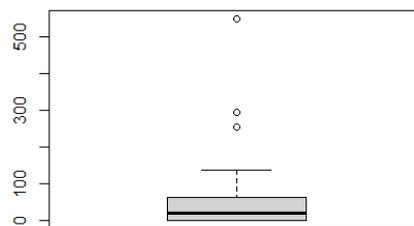
Pada kasus donor darah positif HIV di Jawa Barat pada tahun 2018 memiliki rata-rata jumlah kasus sebanyak  $64,5556 \approx 65$  kasus donor dan kasus donor darah positif HIV tertinggi terdapat di Kabupaten Sukabumi dengan kasus mencapai 550 kasus. Selain itu, terdapat juga sembilan kabupaten/kota yang tidak memiliki kasus donor darah positif HIV diantaranya adalah: Kabupaten Cianjur, Kabupaten Bandung, Kabupaten Garut, Kabupaten Tasikmalaya,

Kabupaten Sumedang, Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bandung Barat, Kota Cimahi, Kota Tasikmalaya dan Kota Banjar.

Pada kasus penyakit Sifilis di Jawa Barat pada tahun 2018 memiliki rata-rata sebesar 20,2593  $\approx$  20 kasus dan kasus penyakit Sifilis tertinggi terdapat di Kabupaten Kuningan dengan jumlah kasus sebanyak 36 kasus sedangkan kasus penyakit Sifilis yang terendah yaitu sebanyak 3 kasus terdapat di Kabupaten Pangandaran.

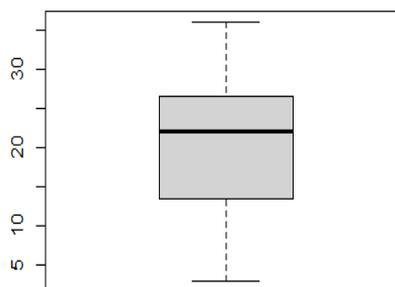
Selain itu, berdasarkan Tabel 1 dapat dilihat melalui nilai simpangan bakunya sebesar 5639,33 bahwa keragaman data yang terbesar itu terdapat pada data jumlah penggunaan kondom ( $x_3$ ) dibandingkan data lainnya.

### Mengecek *Outlier*



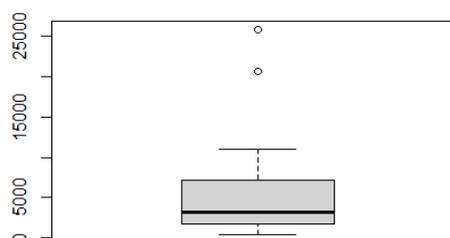
**Gambar 1.** *Boxplot* Variabel Jumlah kasus donor darah Positif HIV

Dilihat pada Gambar 1 menunjukkan adanya *outlier* pada variable jumlah kasus donor darah positif HIV yang terdapat pada data ke-2, 20 dan 22.



**Gambar 2.** *Boxplot* Variabel Jumlah Kasus Penyakit Sifilis

Dilihat pada Gambar 2 di atas menunjukkan tidak terdapat *outlier* pada variabel jumlah kasus penyakit sifilis



**Gambar 3.** *Boxplot* Variabel Jumlah Penggunaan Kondom

Dilihat pada Gambar 3 di atas menunjukkan adanya *outlier* pada variabel jumlah penggunaan kondom yang terdapat pada data ke-6 dan 16. Berdasarkan hasil di atas bahwa terdapat *outlier* pada variabel jumlah kasus darah donor darah positif HIV ( $X_1$ ) dan jumlah penggunaan kondom ( $X_3$ ) sehingga perlu dilakukan transformasi agar hasil analisis menjadi lebih akurat dengan menggunakan transformasi logaritma terhadap variabel  $X_1$  dan  $X_3$ .

**Model Regresi Tobit**

Model regresi tobit diperoleh dengan menggunakan estimasi parameter *maksimum Likelihood* berdasarkan faktor-faktor yang mempengaruhinya yaitu jumlah kasus donor darah positif HIV ( $\log X_1$ ) yang telah ditransformasi, jumlah kasus penyakit *Sifilis* ( $X_2$ ) dan jumlah penggunaan kondom ( $\log X_3$ ) yang telah ditransformasi dengan jumlah pengamatan sebanyak 27 kabupaten/kota di Provinsi Jawa Barat. Penelitian ini menggunakan *software R-Studio* dan memberikan hasil taksiran parameter model regresi tobit sebagai berikut:

$$\widehat{y}_i^* = 24,3851 - 0,5605 \log X_1 + 0,4618 X_2 - 4,1238 \log X_3$$

**Pengujian Sifat MLE (*Maximum Likelihood Estimation*) Tobit dengan Asumsi**

Pengujian sifat MLE tobit dengan asumsi yaitu pengujian ini sama seperti uji asumsi klasik dimana residualnya berdistribusi normal, ragam residualnya bersifat homogen dan tidak terdapat multikolinearitas diantara variabel bebas.

**Uji Normalitas**

Hipotesis yang digunakan:

$H_0$  : Residual mengikuti distribusi normal

$H_1$  : Residual tidak mengikuti distribusi normal

Statistika Uji

$$A = -N - S \tag{1}$$

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(Y_i) + \ln(1 - F(Y_{N+1-i}))]$$

Dengan:

$F$  = fungsi distribusi kumulatif dari distribusi yang ditentukan.

$Y_i$  = data yang diurutkan

**Tabel 2.** Hasil Uji Normalitas

<i>Anderson-Darling (A)</i>	<i>P<sub>value</sub></i>
0,3227	0,5105

Kriteria Uji

Tolak  $H_0$  jika  $P_{value} < \alpha$  dengan  $P_{value} = 0,5105$  dan  $\alpha = 0,05$   
 $0,5105 > 0,05$  maka  $H_0$  diterima.

Kesimpulan: Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa tidak cukup bukti untuk menolak  $H_0$  artinya residual mengikuti distribusi normal.

**Uji Homogenitas**

Hipotesis yang digunakan:

$H_0$  : Ragam residual homogen

$H_1$  : Ragam residual tidak homogen

Statistika Uji

$$T^3 = \frac{1}{D} [\sum_{i=1}^k a_i (X_{n-i+1} - X_i)]^2 \tag{2}$$

$$i = 1, 2, 3, \dots n$$

$$D = \sum_{i=1}^n a_i (X_i - \bar{X})^2 \tag{3}$$

Dengan:

$n$  = jumlah data

$a$  = koefisien uji *Shapiro-Wilk*

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

**Tabel 3.** Hasil Uji Homogenitas

<i>White Test</i>	<i>P<sub>value</sub></i>
2,3874	0,3031

Kriteria Uji

Tolak  $H_0$  jika  $P_{value} < \alpha$  dengan  $P_{value} = 0,3031$  dan  $\alpha = 0,05$   
 $0,3031 > 0,05$  maka  $H_0$  diterima.

Kesimpulan: Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa tidak cukup bukti untuk menolak  $H_0$  artinya Ragam residual homogen.

### Uji Multikolinearitas

Hipotesis yang digunakan:

$H_0$  : VIF < 10 (tidak terdapat multikolinieritas)

$H_1$  : VIF > 10 (terdapat multikolinieritas)

Statistika Uji

$$VIF = \frac{1}{1-R_i^2} \quad \dots(4)$$

Dengan:

$R_i^2$  = koefisien determinasi ganda antara variabel bebas  $x_i$  dengan variabel bebas lainnya.

**Tabel 4.** Hasil Uji Multikolinearitas

<i>VIF</i>	<i>log x<sub>1</sub></i>	<i>x<sub>2</sub></i>	<i>log x<sub>3</sub></i>
	1,1317	1,2080	1,1440

Kriteria Uji

Tolak  $H_0$  jika nilai VIF > 10 dengan Nilai VIF untuk  $\log x_1 = 1,1317$ ,  $x_2 = 1,2080$  dan  $\log x_3 = 1,1440$  maka  $H_0$  diterima. Maka dapat disimpulkan tidak terdapat multikolinieritas.

### Pengujian Parameter Model Regresi Tobit

Pengujian signifikansi parameter bertujuan untuk mengetahui apakah ada atau tidaknya hubungan parameter didalam model regresi. Pengujian signifikan parameter dapat dilakukan secara serentak atau simultan dan secara individual atau parsial.

#### Uji Simultan

Hipotesis yang digunakan:

$H_0$  :  $\beta_1 = \beta_2 = \beta_3 = 0$

$H_1$  : minimal ada salah satu  $\beta_j \neq 0$

Statistika Uji

$$G = -2 \ln \left( \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) \quad \dots(5)$$

Dengan:

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

$L(\hat{\Omega}) = \prod_{i=1}^n f(y_i; \beta_1, \beta_2, \dots, \beta_k)$  dengan  $\Omega = \{\beta_1, \beta_2, \dots, \beta_k\}$

$L(\hat{\omega})$  = Likelihood tanpa variabel bebas tertentu

$L(\hat{\Omega})$  = Likelihood dengan variabel bebas

**Tabel 5.** Hasil Uji Glejser

<i>Chisquare</i>	<i>P<sub>value</sub></i>
16,453	0,0009

Berdasarkan Tabel 5 dengan metode *likelihood ratio* atau uji Glejser maka diperoleh nilai statistika ujinya sebagai berikut:

$G = 16,453$

Kriteria Uji

Tolak  $H_0$  jika nilai uji  $G > \chi^2_{(\alpha,k)}$  atau  $p\text{-value} < \alpha$  dengan  $\alpha = 0,05$  dan  $\chi^2_{(0,05,2)} = 5,591$  berdasarkan tabel distribusi *Chii-Square*.

$16,453 > 5,591$  atau  $0,0009 < 0,05$  maka  $H_0$  ditolak.

Kesimpulan:

Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa cukup bukti untuk menolak  $H_0$  artinya variabel bebas berpengaruh secara simultan terhadap variabel terikat.

**Uji Parsial**

Hipotesis yang digunakan:

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

$H_1 : \beta_j \neq 0$  (koefisien  $\beta_j$  signifikan secara statistik)

Dengan:  $j = 1, 2, 3$

Statistika Uji

$$W^2 = \frac{\widehat{\beta}_j^2}{(SE_{\beta_j^2})^2} \dots(6)$$

Dengan:

$SE_{\widehat{\beta}_j}$  = standar *error* koefisien parameter

$\beta_j^2$  = nilai koefisien dugaan variabel bebas

**Tabel 6.** Hasil Uji Wald

Variabel	<i>Chiisquare</i>	<i>Pvalue</i>
Log $X_1$	0,9599	0,3272
$X_2$	10,3657	0,0012*
Log $X_3$	7,6652	0,0056*

\*) Signifikan pada  $\alpha = 0,05$

Kriteria Uji

Tolak  $H_0$  jika nilai uji  $W^2 > \chi^2_{(\alpha,k)}$  atau  $p\text{-value} < \alpha$  dengan  $\alpha = 0,05$  dan  $\chi^2_{(0,05,2)} = 5,591$ .

Terlihat pada Tabel 6 variabel yang signifikan pada model yaitu variabel  $X_2$  dan  $X_3$

Kesimpulan

Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa cukup bukti untuk menolak  $H_0$  artinya koefisien  $\widehat{\beta}_2$  dan  $\widehat{\beta}_3$  signifikan secara statistik atau jumlah kasus penyakit *sifilis* dan jumlah penggunaan kondom memiliki pengaruh secara signifikan terhadap jumlah kematian akibat AIDS. Sedangkan, koefisien  $\widehat{\beta}_1$  tidak signifikan secara statistik atau jumlah kasus donor darah positif HIV tidak memiliki pengaruh secara signifikan terhadap jumlah kematian akibat AIDS.

**Model Regresi Tobit tanpa Variabel tanpa log  $x_1$**

Pada pengujian parsial sebelumnya, dapat disimpulkan bahwa pada variable log  $X_1$  tidak signifikan secara statistik sehingga perlu dibuat kembali model baru regresi tobit dengan variabel yang diikutsertakan dalam model hanya jumlah kasus penyakit *sifilis* ( $X_2$ ) dan jumlah penggunaan kondom ( $\log X_3$ ) dengan menggunakan estimasi parameter *maksimum likelihood*. Diperoleh hasil taksiran parameter model regresi tobit sebagai berikut:

**Tabel 7.** Taksiran Parameter Model Regresi Tobit tanpa log  $x_1$

Parameter	Taksiran	Standard Error	Z	<i>Pvalue</i>
$\beta_0$	-5,1170	3,6436	-1,404	0,1602

Parameter	Taksiran	Standard Error	Z	P <sub>value</sub>
$\beta_2$	0,4505	0,1440	3,128	0,0018
$\beta_3$	-0,0012	0,0005	-2,389	0,0169

Berdasarkan Tabel 7 diperoleh taksiran model tobit adalah sebagai berikut

$$\widehat{y}_i^* = -5,1170 + 0,4505X_2 - 0,0012 \log X_3$$

Pada model regresi tobit di atas menyatakan bahwa  $b_2 = 0,4505$  artinya bahwa setiap penambahan jumlah kasus penyakit *sifilis* 10 orang maka jumlah kematian akibat AIDS akan meningkat sebesar  $0,4505 \times 10 = 4,5 \approx 5$  orang.  $b_3 = -0,0012$  artinya bahwa setiap jumlah penggunaan kondom bertambah  $10^{1000}$  maka jumlah kasus kematian akibat AIDS akan menurun sebesar  $0,0012 \times 1000 = 1,2$  kasus.

#### Pengujian Sifat MLE (Maximum Likelihood Estimation) Tobit dengan Asumsi

Dilakukan kembali uji asumsi pada model baru regresi tobit terbaru tanpa  $\log x_1$  dimana residualnya berdistribusi normal, ragam residualnya bersifat homogen dan tidak terdapat multikolinearitas diantara variabel bebas.

#### Uji Normalitas

Dilakukan uji normalitas kembali dengan hasil statistik uji pada tabel di bawah ini

**Tabel 8.** Hasil Uji Normalitas

<i>Anderson-Darling (A)</i>	<i>P<sub>value</sub></i>
0,9318	0,0765

Tolak  $H_0$  jika  $P_{value} < \alpha$  dengan maka  $H_0$  diterima. Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa tidak cukup bukti untuk menolak  $H_0$  artinya residual mengikuti distribusi normal.

#### Uji Homogenitas

Dilakukan uji homogenitas kembali dengan hasil statistik uji pada tabel di bawah ini

**Tabel 9.** Hasil Uji Homogenitas

<i>White Test</i>	<i>P<sub>value</sub></i>
2,6461	0,2663

Tolak  $H_0$  jika  $P_{value} < \alpha$  maka  $H_0$  diterima. Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa tidak cukup bukti untuk menolak  $H_0$  artinya Ragam residual homogen.

#### Uji Multikolinearitas

Dilakukan uji multikolinearitas kembali dengan hasil statistik uji pada tabel di bawah ini

**Tabel 10.** Hasil Uji Multikolinearitas

<i>VIF</i>	$x_2$	$\log x_3$
	1,0890	1,1346

Tolak  $H_0$  jika nilai  $VIF > 10$  dengan Nilai  $VIF$  maka  $H_0$  diterima. Maka dapat disimpulkan tidak terdapat multikolinieritas.

#### Pengujian Parameter Model Regresi Tobit

Dilakukan kembali pengujian signifikansi secara serentak atau simultan dan secara individual

atau parsial pada model regresi tobit yang terbaru tanpa log  $x_1$ .

**Uji Simultan**

Hipotesis:

$$H_0 : \beta_2 = \beta_3 = 0$$

$$H_1 : \text{minimal ada salah satu } \beta_j \neq 0$$

Statistika Uji

**Tabel 11.** Hasil Uji Glejser

<i>Chisquare</i>	<i>P<sub>value</sub></i>
16,075	0,0003

Berdasarkan Tabel 11 dengan metode *likelihood ratio* atau uji Glejser maka diperoleh nilai statistika ujinya sebagai berikut:  $G = 16,075$

Kriteria Uji

Tolak  $H_0$  jika nilai uji  $G > \chi^2_{(\alpha,k)}$  atau  $p\text{-value} < \alpha$  dengan  $\alpha = 0,05$  dan  $\chi^2_{(0,05,2)} = 5,991$  berdasarkan tabel distribusi *Chi-Square*.  $16,075 > 5,991$  atau  $0,0003 < 0,05$  maka  $H_0$  ditolak.

Kesimpulan

Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa cukup bukti untuk menolak  $H_0$  artinya variabel bebas berpengaruh secara simultan terhadap variabel terikat.

**Uji Parsial**

Hipotesis yang digunakan:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

**Tabel 12.** Hasil Uji Wald

Variabel	<i>Chisquare</i>	<i>P<sub>value</sub></i>
$X_2$	9,7820	0,0018*
Log $X_3$	5,7076	0,0169*

\*) Signifikan pada  $\alpha = 0,05$

Kriteria Uji

Tolak  $H_0$  jika nilai uji  $W^2 > \chi^2_{(\alpha,k)}$  atau  $p\text{-value} < \alpha$  dengan  $\alpha = 0,05$  dan  $\chi^2_{(0,05,2)} = 5,991$ .

Terlihat pada Tabel 12 variabel yang signifikan pada model yaitu variabel  $X_2$  dan log  $X_3$

Kesimpulan

Dengan tingkat kepercayaan sebesar 95% dapat disimpulkan bahwa cukup bukti untuk menolak  $H_0$  artinya koefisien  $\widehat{\beta}_2$  dan  $\widehat{\beta}_3$  signifikan secara statistik atau jumlah kasus penyakit *sifilis* dan jumlah penggunaan kondom memiliki pengaruh secara signifikan terhadap jumlah kematian akibat AIDS.

**Penentuan Kesesuaian Model**

Menurut Sembiring (2019) menentukan kesesuaian model dalam analisis regresi tobit dapat menggunakan metode Akaike Information Criterion (AIC) dan koefisien determinasi ( $R^2$ ).

**Metode Akaike Information Criterion (AIC)**

Menghitung nilai AIC menggunakan persamaan sebagai berikut:

$$AIC = -2 \ln(L) + 2k \tag{7}$$

maka didapat nilai AIC berdasarkan persamaan di atas sebesar 88,2028 dan diperoleh model regresi tobit terbaik untuk faktor yang mempengaruhi jumlah kasus kematian akibat AIDS adalah sebagai berikut:

$$\widehat{y}_i^* = -5,1170 + 0,4505X_2 - 0,0012 \log X_3$$

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

Persamaan  $R^2$  yang digunakan dalam regresi Tobit adalah sebagai berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^n \hat{e}_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad \dots(8)$$

Keterangan:

$Y_i$  = nilai variabel terikat pada observasi ke-i

$\bar{Y}$  = nilai rata-rata variabel terikat

$e_i$  = residual pada model

Maka diperoleh nilai Koefisien Determinasi ( $R^2$ ) adalah sebagai berikut:

$$R^2 = 0,1670$$

Berdasarkan hasil di atas diperoleh nilai koefisien determinasi sebesar 0,1670 artinya sebesar 16,70% variasi dari jumlah kasus kematian akibat AIDS mampu dijelaskan oleh jumlah kasus penyakit *sifilis* dan jumlah penggunaan kondom, sisanya sebesar 83,30% dijelaskan oleh variabel lain yang tidak dianalisis.

### D. Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Analisis regresi tobit dalam estimasi parameter menggunakan metode *maksimum likelihood*, maka diperoleh taksiran model regresi tobit adalah sebagai berikut:

$$\hat{y}_i^* = -5,1170 + 0,4505X_2 - 0,0012 \log X_3$$

Pada model regresi tobit di atas menyatakan bahwa  $b_2 = 0,4505$  artinya bahwa setiap penambahan jumlah kasus penyakit *sifilis* 10 orang maka jumlah kematian akibat AIDS akan meningkat sebesar  $0,4505 \times 10 = 4,5 \approx 5$  orang.  $b_3 = -0,0012$  artinya bahwa setiap jumlah penggunaan kondom bertambah  $10^{1000}$  maka jumlah kasus kematian akibat AIDS akan menurun sebesar  $0,0012 \times 1000 = 1,2$  kasus.

2. Pada pengujian parsial untuk model taksiran regresi tobit yang telah dipaparkan pada poin satu dapat diketahui bahwa jumlah kasus penyakit *sifilis*  $X_2$  dan logaritma jumlah penggunaan kondom  $\log X_3$  berpengaruh secara signifikan terhadap jumlah kasus kematian akibat AIDS.

### Acknowledge

Peneliti mengucapkan terima kasih kepada semua pihak yang terlibat dalam penelitian ini dan kepada orang tua, keluarga, bapak Dr. Apt. Suwendar, S.Si., M.Si selaku Dekan F-MIPA Unisba, Bapak Abdul Kudus, S.Si., M.Si., Ph.D selaku Ketua Program Studi Statistika Unisba, Ibu Teti Sofia Yanti, Dra., M.Si selaku Dosen Pembimbing, dosen-dosen Program Studi Statistika, Bapak Dr. Aceng Komarudin Mutaqin, MT., M.Si. selaku dosen wali dan rekan-rekan yang tidak bisa disebutkan namanya satu persatu.

### Daftar Pustaka

- [1] Nurhanda, D. A. (2018) "*Studi Kasus Konsumsi Rokok Rumah Tangga Provinsi Daerah Istimewa Yogyakarta Tahun 2016*" Yogyakarta: Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga.
- [2] Nirwana, B.M, S. (2018). Metode buckley-james untuk estimasi model regresi linier pada data tersensor kanan. *Statistika*, 6(1), 1-7.
- [3] Amemiya, T. (1984). Tobit model: a survey. *Jurnal of Econometrics*, 24, 3-61.
- [4] Jeryana, P., Kencana, E. N., & Gandhiadi, G. K. (2014). Model regresi tobit konsumsi susu cair pabrik (Studi kasus rumah tangga di provinsi bali). *E-jurnal Matematika*, 3(2), 75-85.
- [5] Sinaga, R. L., Saputra, W., & Qurniawan, H. (2021). Pengelompokan jumlah kasus penyakit aids berdasarkan provinsi menggunakan metode k-means. *Jurnal penerapan*

- sistem informasi (komputer & manajemen)*, **2**(2), 99-107.
- [6] Marlinda, Y., & Azinar, M. (2017). Perilaku pencegahan penularan hiv/aids. *Jurnal of health education*, **2**(1), 192-200.
- [7] Rondonuwu, M. R. (2022). *Laporan Eksekutif Perkembangan HIV AIDS dan Penyakit Infeksi Menular Seksual (PIMS) Triwulan II Tahun 2022*. Jakarta: Kementerian Kesehatan Republik Indonesia Direktorat Jenderal Pencegahan dan Pengendalian Penyakit.
- [8] Ajeng Mega Pratiwi and A. K. Mutaqin, “Penerapan Algoritma Naïve Bayes Classifier dalam Memprediksi Status Keberlanjutan Polis Nasabah Asuransi PT.X,” *Jurnal Riset Statistika*, vol. 1, no. 2, pp. 117–126, Dec. 2021, doi: 10.29313/jrs.v1i2.435.
- [9] Muhammad Rizq Nafisyah Alam and Aceng Komarudin Mutaqin, “Pemodelan Distribusi Poisson-Sujatha pada Data Frekuensi Klaim Asuransi Kendaraan Bermotor di Indonesia,” *Jurnal Riset Statistika*, pp. 71–78, Jul. 2023, doi: 10.29313/jrs.v3i1.1944.
- [10] Aisha Kusuma Putri and Suliadi, “Rekomendasi Destinasi Wisata di Indonesia Menggunakan Metode Item2Vec,” *Jurnal Riset Statistika*, pp. 11–18, Jul. 2023, doi: 10.29313/jrs.v3i1.1770.