

**E - J u r n a l**

**M a t e m a t i k a**

[OPEN JOURNAL SYSTEMS](#)

[Journal Help](#)

USER

Username

Password

Remember me

NOTIFICATIONS

- [View](#)
- [Subscribe / Unsubscribe](#)

JOURNAL CONTENT

Search

- [By Issue](#)
- [By Author](#)
- [By Title](#)
- [Other Journals](#)

- [For Readers](#)
- [For Authors](#)
- [For Librarians](#)

---

[HOME](#)   [ABOUT](#)   [LOG IN](#)   [REGISTER](#)   [SEARCH](#)   [CURRENT](#)   [ARCHIVES](#)

[Home](#) > ***E-Jurnal Matematika***

---

## E-Jurnal Matematika

---

E-Jurnal Matematika merupakan salah satu jurnal elektronik yang ada di Universitas Udayana, sebagai media komunikasi antar peminat di bidang ilmu matematika dan terapan, seperti statistika, matematika finansial, pengajaran matematika dan terapan matematika dibidang ilmu lainnya. Jurnal ini lahir sebagai salah satu bentuk nyata peran serta jurusan Matematika FMIPA UNUD guna mendukung percepatan tercapainya target mutu UNUD, selain itu jurnal ini terbit didorong oleh surat edaran Dirjen DIKTI tentang syarat publikasi karya ilmiah bagi program Sarjana di Jurnal Ilmiah. E-jurnal Matematika juga menerima hasil-hasil penelitian yang tidak secara langsung berkaitan dengan tugas akhir mahasiswa meliputi penelitian atau artikel yang merupakan kajian keilmuan.

### Editorial Team

Ketua : Desak Putu Eka Nilakusumawati, S.Si., M.Si  
Sekretaris : I Made Eka Dwipayana S.Si. M.Si.

Penyunting :

1. Tjokorda Bagus Oka Ph.D.
2. Komang Dharmawan Ph.D.
3. Drs. GK Gandhiadi MT.
4. Ir. I Komang Gde Sukarsa M.Si.
5. Ir. I Putu Eka Nila Kencana MT

ISSN: 2303-1751

---

# E - J u r n a l M a t e m a t i k a

[OPEN JOURNAL SYSTEMS](#)

[Journal Help](#)

USER

Username

Password

Remember me

NOTIFICATIONS

- [View](#)
- [Subscribe](#) / [Unsubscribe](#)

JOURNAL CONTENT

Search

All

Browse

- [By Issue](#)
- [By Author](#)
- [By Title](#)
- [Other Journals](#)

FONT SIZE

INFORMATION

- [For Readers](#)

- [For Authors](#)
- [For Librarians](#)

HOME ABOUT LOG IN REGISTER SEARCH CURRENT ARCHIVES

[Home](#) > [Archives](#) > **Vol 3, No 2 (2014)**

## Vol 3, No 2 (2014)

### Table of Contents

#### Articles

|  |                     |
|--|---------------------|
| <a href="#">ANALISIS REGRESI BAYES LINEAR SEDERHANA DENGAN PRIOR NONINFORMATIF</a>   | <a href="#">PDF</a> |
| <i>ANAK AGUNG ISTRI AGUNG CANDRA ISWARI, I WAYAN SUMARJAYA, I GUSTI AYU MADE SRINADI</i>                                       | 38 - 44             |
| <a href="#">PERBANDINGAN REGRESI ROBUST PENDUGA MM DENGAN METODE RANDOM SAMPLE CONSENSUS DALAM MENANGANI PENCILAN</a>          | <a href="#">PDF</a> |
| <i>NI PUTU NIA IRFAGUTAMI, I GUSTI AYU MADE SRINADI, I WAYAN SUMARJAYA</i>   | 45 - 52             |
| <a href="#">PENGELOMPOKAN BERBAGAI MERK MI INSTAN BERDASARKAN KEMIRIPAN KANDUNGAN GIZI DENGAN MENGGUNAKAN ANALISIS BIPLLOT</a> | <a href="#">PDF</a> |
| <i>AGUSTINUS ANGELAUS ETE, NI LUH PUTU SUCIPTAWATI, DESAK PUTU EKA NILAKUSMAWATI</i>   | 53 - 63             |
| <a href="#">PERHITUNGAN DANA PENSIUN DENGAN METODE PROJECTED UNIT CREDIT DAN INDIVIDUAL LEVEL PREMIUM</a>                      | <a href="#">PDF</a> |
| <i>I GUSTI AYU KOMANG KUSUMA WARDHANI, I NYOMAN WIDANA, NI KETUT TARI TASTRAWATI</i>   | 64 - 74             |
| <a href="#">MODEL REGRESI TOBIT KONSUMSI SUSU CAIR PABRIK (Studi Kasus Rumah Tangga di Provinsi Bali)</a>                      |                     |
| <i>I PUTU JERYANA, I PUTU EKA NILA KENCANA, G. K. GANDHIADI</i>  | 75 - 85             |

ISSN: 2303-1751

## ANALISIS REGRESI BAYES LINEAR SEDERHANA DENGAN *PRIOR* NONINFORMATIF

ANAK AGUNG ISTRI AGUNG CANDRA ISWARI<sup>1</sup>, I WAYAN SUMARJAYA<sup>2</sup>,  
I GUSTI AYU MADE SRINADI<sup>3</sup>

<sup>1,2,3</sup>Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali  
e-mail: <sup>1</sup>iswari.candra@gmail.com, <sup>2</sup>sumarjaya@unud.ac.id, <sup>3</sup>srinadiigustiyumade@yahoo.co.id

### *Abstract*

*The aim of this study is to apply Bayesian simple linear regression using noninformative prior. The data used in this study is 30 observational data with error generated from normal distribution. The noninformative prior was formed using Jeffreys' rule. Computation was done using the Gibbs Sampler algorithm with 10.000 iteration. We obtain the following estimates for the parameters,  $\alpha_0 = 1,698045$  with 95% Bayesian confidence interval (0,775775; 2,626025),  $\beta = 2,999468$  with 95% Bayesian confidence interval (2,948; 3,052), and  $\sigma^2 = 0,697669$  with 95% Bayesian confidence interval (0,375295; 1,114). These values are not very different compared to the actual value of the parameters, which are  $\alpha_0 = 2$  and  $\beta = 3$*

**Keywords:** *Bayesian regression, noninformative prior, Jeffreys' rule, the Gibbs Sampler algorithm*

### 1. Pendahuluan

Analisis regresi linear sederhana adalah salah satu cara yang dapat digunakan untuk mengetahui hubungan antara variabel bebas dan variabel terikat. Pendugaan parameter model regresi linear sederhana dapat dilakukan dengan berbagai cara tergantung dari pandangan peneliti. Dalam ilmu statistika, terdapat dua pandangan yang sering digunakan sebagai dasar dalam metode-metode untuk mengolah data (William M. Bolstad, 2007). Pandangan pertama merupakan pandangan yang umumnya sering digunakan oleh peneliti (*frequentist*) yakni metode yang digunakan untuk mengolah data adalah metode-metode regresi klasik seperti metode kuadrat terkecil (*least square estimation*) dan metode kemungkinan maksimum (*maximum likelihood estimation*). Pandangan kedua merupakan pandangan yang berbeda dengan para *frequentist*. Pandangan ini menggunakan pengetahuan dari peneliti, yang bersifat

subjektif sebagai *prior* yang kemudian diolah bersama data untuk memperoleh parameter regresi yang diinginkan. Pandangan kedua ini disebut pandangan Bayes.

Dalam pandangan Bayes, seseorang dapat memberikan kepercayaan awal (*prior believe*) terhadap suatu parameter karena adanya asumsi bahwa parameter merupakan suatu variabel acak (William M. Bolstad, 2007). Kepercayaan awal ini dapat diperbarui dengan menggunakan Teorema Bayes ketika diperoleh data amatan. Teorema Bayes menyatakan bahwa distribusi peluang posterior untuk  $\theta$  terhadap data  $x$ , proporsional terhadap produk dari distribusi *prior* untuk  $\theta$  terhadap data dan *likelihood* untuk  $\theta$  jika diberikan data  $x$  (George E. P. Box and George C. Tiao, 1973).

Oleh karena itu, analisis regresi Bayes linear sederhana akan dipengaruhi oleh pemilihan *prior* dan *likelihood* data. Distribusi *prior* adalah distribusi awal parameter  $\theta$

<sup>1</sup> Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

<sup>2</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

sebelum diperolehnya data amatan (Andrew Gelman, et al., 2004). Dengan kata lain distribusi *prior* merupakan tingkat kepercayaan peneliti terhadap setiap nilai parameter yang mungkin. Sehingga distribusi *prior* akan selalu bersifat subjektif karena merupakan representasi kepercayaan peneliti.

Pemilihan *prior* secara umum dilakukan berdasarkan diketahui atau tidaknya informasi mengenai parameter. Jika informasi mengenai parameter diketahui, maka *prior* informatif, yaitu *prior* yang memengaruhi hasil distribusi posterior dan bersifat sangat subjektif dapat digunakan (Andrew Gelman, et al., 2004), sedangkan jika informasi mengenai parameter tidak tersedia, maka digunakan *prior* noninformatif yang tidak memberikan pengaruh yang signifikan terhadap distribusi posterior (George E. P. Box and George C. Tiao, 1973) sehingga informasi yang diperoleh dari data amatan bersifat lebih objektif.

Penelitian ini bertujuan untuk menerapkan analisis regresi Bayes linear sederhana dengan menggunakan *prior* noninformatif. Selain menduga parameter regresi, akan dilakukan inferensi dengan menggunakan selang kepercayaan Bayes.

## 2. Metode Penelitian

Model regresi linear sederhana merupakan salah satu model regresi yang sering digunakan dalam analisis regresi. Pada model ini, hanya terdapat satu variabel bebas dengan fungsi regresi linear. Disebut sederhana karena model ini hanya melibatkan satu variabel bebas dan disebut linear karena linear dalam parameter dan linear dalam variabel bebasnya [5]. Model regresi linear sederhana yang digunakan dalam penelitian ini adalah sebagai berikut:

$$y = \alpha_0 + \beta x + \varepsilon.$$

Data yang digunakan dalam penelitian ini adalah data yang dibangkitkan dengan menggunakan program R versi 3.0.2. Data yang dibangkitkan adalah data dengan galat yang berdistribusi normal dengan *mean* nol dan varians satu. Variabel bebas yang

dibangkitkan merupakan bilangan bulat positif dengan nilai 1, 2, ..., 30. Variabel terikat ditentukan oleh hubungan linear antara variabel bebas dan variabel terikat. Adapun nilai parameter yang dipilih sebagai contoh simulasi dalam penelitian ini adalah  $\alpha_0 = 2$  dan  $\beta = 3$ . Sehingga hubungan linear antara variabel bebas dan variabel terikat yang ditentukan adalah sebagai berikut,

$$y = 2 + 3x + \varepsilon$$

dengan  $\varepsilon$  adalah galat berdistribusi normal yang dibangkitkan. Karena data yang dibangkitkan berdistribusi normal, maka *likelihood* data dinyatakan oleh:

$$L(\mu, \sigma^2) \propto \left(\frac{1}{\sigma}\right)^n \exp \sum_{i=1}^n \left\{ -\frac{1}{2\sigma^2} [y_i - \mu]^2 \right\}. \quad (1)$$

Penelitian ini menggunakan *prior* noninformatif yang tidak memberikan pengaruh terhadap distribusi posterior karena tidak tersedianya informasi awal mengenai parameter. *Prior* noninformatif yang digunakan dapat dibentuk dengan menggunakan aturan Jeffreys (Robert E. Kass and Larry Wasserman, 1996). Berdasarkan aturan Jeffreys, dari *likelihood* pada persamaan (1), dibentuk *prior* noninformatif sebagai berikut:

$$\pi(\mu, \sigma^2) = \frac{1}{\sigma^2}. \quad (2)$$

Dari *likelihood* data pada persamaan (1) dan *prior* noninformatif pada persamaan (2) dibentuk distribusi posterior, yaitu:

$$\begin{aligned} \pi(\mu, \sigma^2 | y_i) &\propto \frac{1}{\sigma^2} \left(\frac{1}{\sigma}\right)^n \exp \left[ \sum_{i=1}^n \left\{ -\frac{1}{2\sigma^2} [y_i - \mu]^2 \right\} \right] \\ &= \frac{1}{\sigma^{n+2}} \exp \left[ -\frac{1}{2\sigma^2} \left\{ \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2 \right\} \right] \\ &= \frac{1}{\sigma^{n+2}} \exp \left[ -\frac{1}{2\sigma^2} \{ (n-1)s^2 + n(\bar{y} - \mu)^2 \} \right] \end{aligned}$$

dengan

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2.$$

Untuk memperoleh distribusi posterior marginal untuk  $\sigma^2$ , distribusi posterior  $\pi(\mu, \sigma^2 | y_i)$  diintegrasikan terhadap  $\mu$ , sehingga  $\pi(\sigma^2 | y_i) \propto \int \frac{1}{\sigma^{n+2}} \exp \left[ -\frac{1}{2\sigma^2} \{ (n-1)s^2 + n(\bar{y} - \mu)^2 \} \right] d\mu$

$$\begin{aligned} &\propto \frac{1}{\sigma^{n+2}} \exp \left( -\frac{1}{2\sigma^2} (n-1)s^2 \right) \sqrt{\frac{2\pi\sigma^2}{n}} \\ &\propto (\sigma^2)^{-\frac{n+1}{2}} \exp \left( -\frac{(n-1)s^2}{2\sigma^2} \right) \end{aligned}$$

yang merupakan fungsi densitas untuk invers- $\chi^2$  berskala (*scaled inverse-chi-square*), dengan kata lain  $\sigma^2|y_i \sim \text{Inv} - \chi^2(n-1, s^2)$ . Fungsi densitas untuk distribusi invers- $\chi^2$  berskala memiliki fungsi densitas yang sama dengan distribusi *Inv-gamma* ( $\frac{n-1}{2}, \frac{n-1}{2}s^2$ ) (Andrew Gelman, et al., 2004).

Pendugaan nilai parameter dilakukan dengan menghitung *mean* dari distribusi posterior (Bradley P. Carlin and Thomas A. Louis, 2009). Salah satu metode komputasi yang dapat digunakan untuk menduga parameter adalah metode Markov Chain Monte Carlo. Metode ini membentuk suatu rantai Markov yang digunakan sebagai sampel Monte Carlo atau dapat dinyatakan sebagai:

$$\hat{\mu}_M = \frac{1}{M-B} \sum_{i=B+1}^M g(X_i)$$

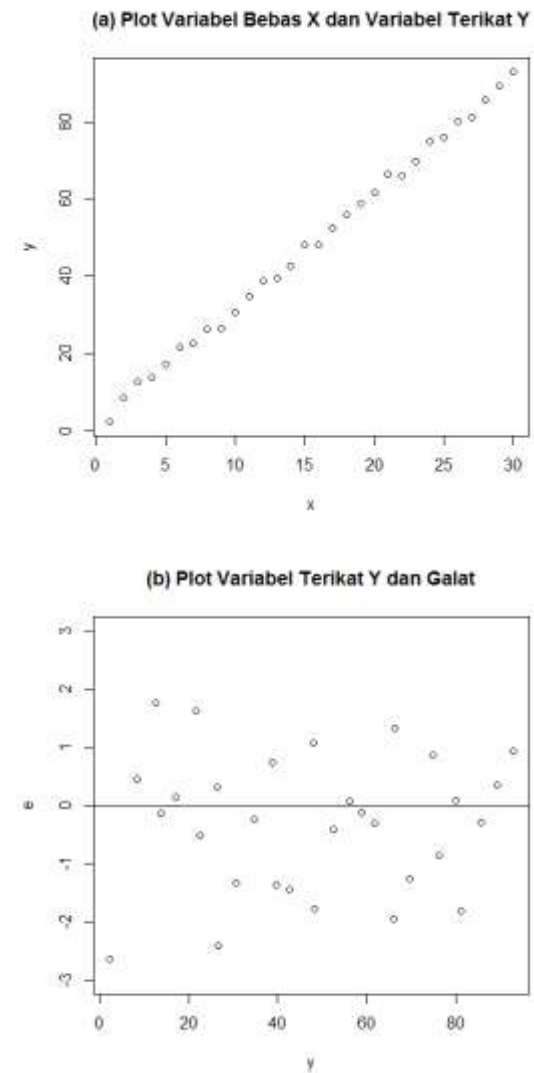
dengan  $M$  merupakan jumlah sampel yang dibangkitkan dan  $B$  merupakan *burn-in* yaitu bilangan bulat non-negatif yang menyatakan jumlah sampel awal yang harus dibuang karena terlalu bias terhadap nilai awal (Radu V. Craiu and Jeffrey S. Rosenthal, 2014).

Dalam penelitian ini digunakan algoritma Gibbs Sampler yang merupakan salah satu algoritma yang termasuk ke dalam kelas algoritma Markov Chain Monte Carlo. Algoritma Gibbs Sampler membangkitkan variabel acak dari suatu distribusi marginal tanpa harus diketahui fungsi densitasnya (Christophe Andrieu, et. al. 2003). Dalam penelitian ini, algoritma Gibbs Sampler dilakukan sebanyak 10.000 kali iterasi dengan *burn-in* sebanyak 1.000 sampel.

### 3. Hasil dan Pembahasan

Adapun data yang dibangkitkan untuk penelitian ini merupakan data dengan galat berdistribusi normal dengan *mean* nol dan varians satu. Plot (a) pada Gambar 1. menunjukkan plot variabel bebas  $X$  dan variabel terikat  $Y$  dan plot (b) menunjukkan plot variabel terikat  $Y$  dengan galat. Plot antara variabel bebas  $X$  dan variabel terikat  $Y$

menunjukkan bahwa data yang dibangkitkan memiliki hubungan yang linear antara variabel terikat dan variabel bebasnya. Sedangkan plot antara variabel terikat  $Y$  dengan galat menunjukkan bahwa galat yang dibangkitkan memiliki varians konstan.



**Gambar 1.** (a) Plot Variabel Bebas  $X$  dan Variabel Terikat  $Y$  dan (b) Plot Variabel Terikat  $Y$  dan Galat

Data variabel bebas  $X$ , galat berdistribusi normal dan variabel terikat  $Y$  yang dibangkitkan ditunjukkan oleh Tabel 1. Data yang telah dibangkitkan memiliki distribusi normal, sehingga *likelihood* data amatan merupakan *likelihood* distribusi normal seperti yang ditunjukkan oleh persamaan (1).

**Tabel 1.** Data Variabel Bebas, Galat, dan Variabel Terikat yang Dibangkitkan

| X  | $\varepsilon$ | $Y = 2 + 3x + \varepsilon$ |
|----|---------------|----------------------------|
| 1  | -2,64753085   | 2,352469                   |
| 2  | 0,45542902    | 8,455429                   |
| 3  | 1,76248521    | 12,762485                  |
| 4  | -0,13762446   | 13,862376                  |
| 5  | 0,13017324    | 17,130173                  |
| 6  | 1,61545617    | 21,615456                  |
| 7  | -0,51863142   | 22,481369                  |
| 8  | 0,30857656    | 26,308577                  |
| 9  | -2,41011001   | 26,589890                  |
| 10 | -1,32809561   | 30,671904                  |
| 11 | -0,23824422   | 34,761756                  |
| 12 | 0,73449361    | 38,734494                  |
| 13 | -1,35865761   | 39,641342                  |
| 14 | -1,44779785   | 42,552202                  |
| 15 | 1,06589440    | 48,065894                  |
| 16 | -1,77573754   | 48,224262                  |
| 17 | -0,42350373   | 52,576496                  |
| 18 | 0,07304432    | 56,073044                  |
| 19 | -0,13094045   | 58,869060                  |
| 20 | -0,30893238   | 61,691068                  |
| 21 | 1,33469871    | 66,334699                  |
| 22 | -1,94607974   | 66,053920                  |
| 23 | -1,25981853   | 69,740181                  |
| 24 | 0,85737874    | 74,857379                  |
| 25 | -0,86449368   | 76,135506                  |
| 26 | 0,07835018    | 80,078350                  |
| 27 | -1,82110904   | 81,178891                  |
| 28 | -0,30040502   | 85,699595                  |
| 29 | 0,33639769    | 89,336398                  |
| 30 | 0,93834574    | 92,938346                  |

Dari data yang dibangkitkan, dilakukan analisis dengan menggunakan metode regresi Bayes linear sederhana. Pendugaan parameter dilakukan dengan menggunakan bantuan program R versi 3.0.2 dan WinBUGS versi 1.4. Luaran dari program tersebut ditunjukkan oleh Tabel 2.

Tabel 2. menunjukkan nilai dugaan untuk masing-masing parameter dengan simpangan baku dan juga kuantil-kuantilnya. Kuantil 2,5% dan 97,5% menunjukkan batas bawah dan batas atas dari selang kepercayaan Bayes

untuk masing-masing parameter.

**Tabel 2.** Luaran Pendugaan Nilai Parameter

|            | Mean     |        | Simpangan Baku |         |          |
|------------|----------|--------|----------------|---------|----------|
| $\alpha_0$ | 1,698045 |        | 0,464224       |         |          |
| $\beta$    | 2,999468 |        | 0,026129       |         |          |
| $\sigma^2$ | 0,697669 |        | 0,18857        |         |          |
|            | Kuantil  |        |                |         |          |
|            | 2,5%     | 25%    | 50%            | 75%     | 97,5%    |
| $\alpha_0$ | 0,775775 | 1,392  | 1,699          | 2,001   | 2,626025 |
| $\beta$    | 2,948    | 2,983  | 2,999          | 3,016   | 3,052    |
| $\sigma^2$ | 0,375295 | 0,5631 | 0,6792         | 0,81675 | 1,114    |

Nilai dugaan untuk parameter  $\alpha_0 = 1,698045$  dengan selang kepercayaan Bayes 95% (0,775775; 2,626025). Selang kepercayaan Bayes dapat diinterpretasikan sebagai peluang nilai parameter  $\alpha_0$  berada di antara selang (0,775775; 2,626025) adalah sebesar 95%. Nilai parameter dugaan  $\alpha_0 = 1,698045$  menunjukkan bahwa nilai variabel terikat  $Y$  akan sama dengan 1,698045 jika nilai variabel bebas  $X$  sama dengan nol.

Nilai dugaan untuk parameter  $\beta = 2,999468$  menyatakan bahwa nilai variabel terikat  $Y$  akan mengalami perubahan sebesar 2,999468 jika terjadi perubahan sebesar satu unit satuan pada variabel bebas  $X$ . Selang kepercayaan Bayes 95% (2,948; 3,052) menunjukkan bahwa peluang nilai parameter  $\beta$  berada di antara selang (2,948; 3,052) adalah sebesar 95%.

Nilai dugaan untuk parameter  $\sigma^2 = 0,697669$  dengan selang kepercayaan Bayes 95% (0,375295; 1,114). Dengan kata lain nilai parameter  $\sigma^2$  memiliki peluang sebesar 95% berada di antara selang (0,375295; 1,114).

Masing-masing nilai parameter dugaan yang diperoleh memiliki kesesuaian dengan nilai parameter yang ditentukan. Nilai dugaan dan nilai sesungguhnya dari parameter  $\alpha_0$ , yaitu dua, tidak memiliki perbedaan yang jauh. Selang kepercayaan Bayes meyakinkan bahwa nilai parameter sesungguhnya berada pada selang tersebut.

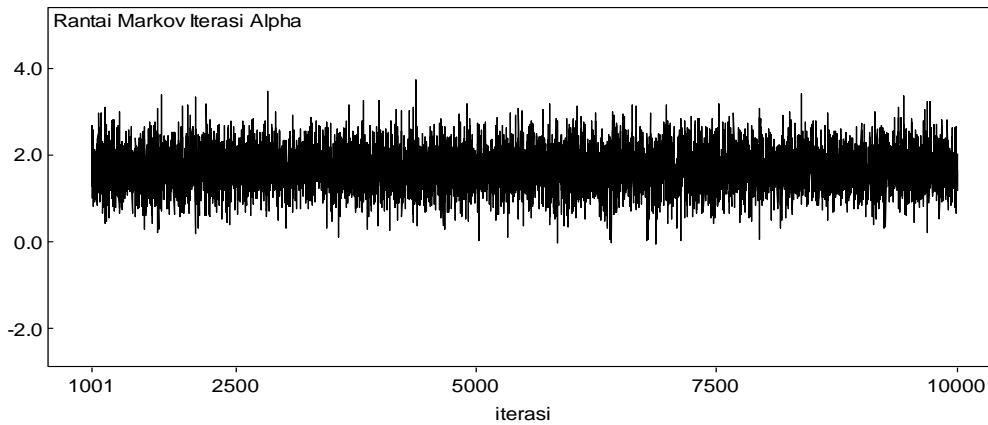
Nilai dugaan untuk parameter  $\beta$  memiliki nilai yang mendekati nilai parameter



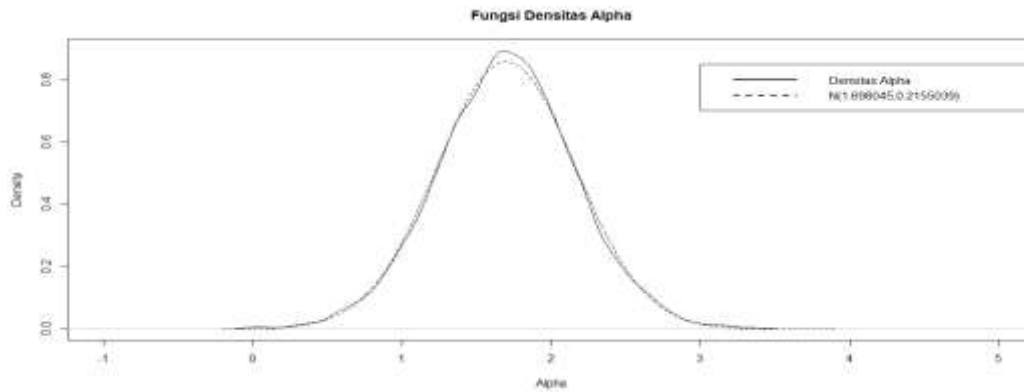
sesungguhnya. Jika dibulatkan, maka nilai parameter dugaan untuk  $\beta$  akan sama dengan nilai parameter  $\beta$  yang sesungguhnya, yaitu tiga. Nilai dugaan yang mendekati nilai parameter yang sesungguhnya ini juga ditunjukkan oleh sempitnya selang kepercayaan Bayes untuk nilai parameter

dugaan  $\beta$ .

Nilai dugaan untuk parameter  $\sigma^2$  juga tidak memiliki perbedaan yang jauh dari nilai parameter yang sesungguhnya. Varians dari galat yang dibangkitkan adalah satu, dan selang kepercayaan Bayes mencakup nilai tersebut.



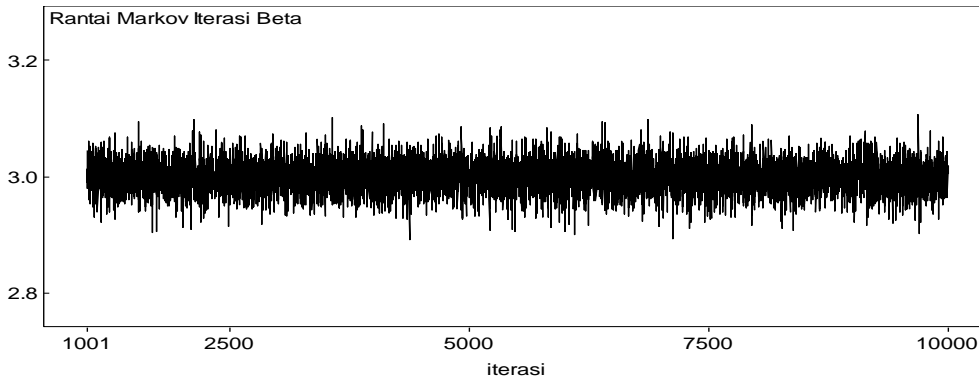
**Gambar 2a.** Rantai Markov untuk Iterasi Parameter  $\alpha_0$



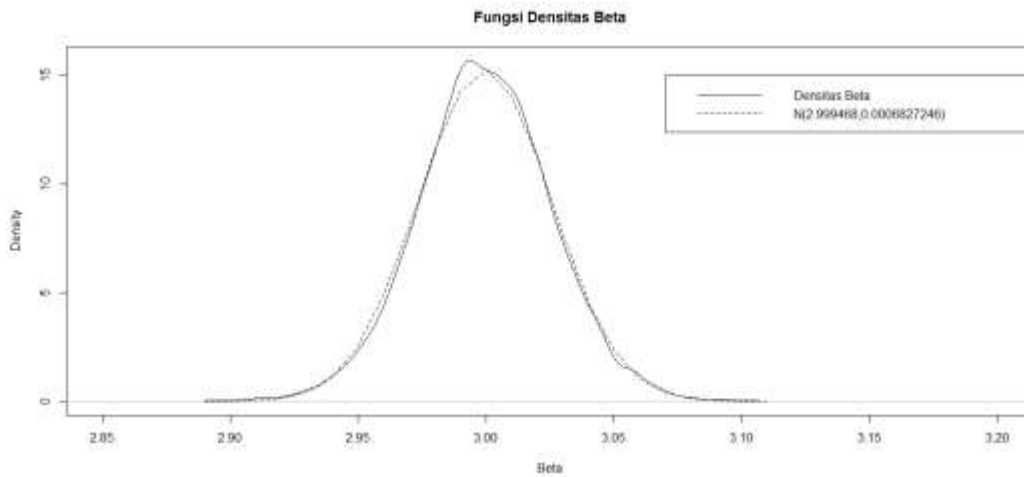
**Gambar 2b.** Plot Fungsi Densitas untuk  $\alpha_0$

Gambar 2a. menunjukkan rantai Markov yang diperoleh dari iterasi Gibbs Sampler untuk parameter  $\alpha_0$ . Dari rantai Markov yang diperoleh, dapat dibentuk suatu plot fungsi

densitas untuk parameter  $\alpha_0$  seperti yang ditunjukkan pada Gambar 2b. Plot fungsi densitas parameter  $\alpha_0$  memiliki bentuk yang menyerupai distribusi normal.



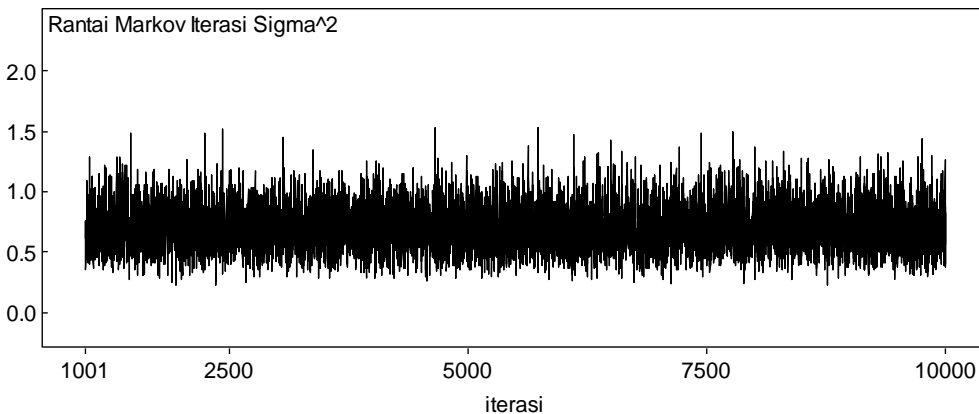
**Gambar 3a.** Rantai Markov untuk Iterasi Parameter  $\beta$



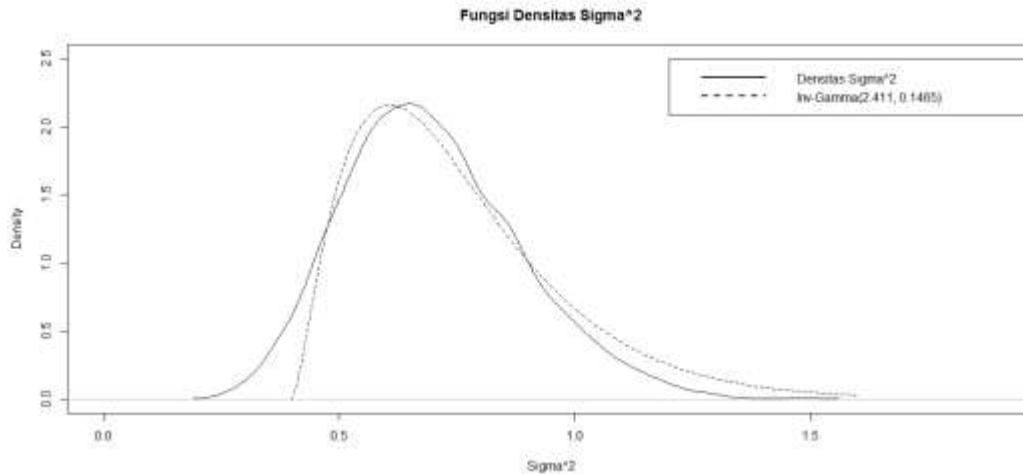
**Gambar 3b.** Plot Fungsi Densitas untuk  $\beta$

Gambar 3a. menunjukkan rantai Markov yang diperoleh dari iterasi Gibbs Sampler untuk parameter  $\beta$ . Dari rantai Markov yang diperoleh, dibentuk suatu plot fungsi densitas

untuk parameter  $\beta$  seperti yang ditunjukkan pada Gambar 3b. Plot fungsi densitas parameter  $\beta$  juga memiliki bentuk yang menyerupai distribusi normal.



**Gambar 4a.**Rantai Markov untuk Iterasi Parameter  $\sigma^2$



**Gambar 4b.** Plot Fungsi Densitas untuk  $\sigma^2$

Gambar 4a. menunjukkan rantai Markov yang diperoleh dari iterasi Gibbs Sampler untuk parameter  $\sigma^2$ . Dari rantai Markov yang diperoleh, dibentuk suatu plot fungsi densitas untuk parameter  $\sigma^2$  seperti yang ditunjukkan pada Gambar 4b. Plot fungsi densitas parameter  $\sigma^2$  memiliki bentuk yang menyerupai distribusi invers-gamma. Hal ini bersesuaian dengan distribusi posterior marginal dari  $\sigma^2$  yang diperoleh.

#### 4. Kesimpulan

Penerapan analisis regresi Bayes linear sederhana dengan menggunakan *prior* noninformatif selain memberikan nilai dugaan untuk parameter, juga memberikan gambaran mengenai kecenderungan distribusi dari parameter-parameter yang diduga. Hal ini dapat digunakan sebagai informasi *prior* jika dilakukan penelitian pada masa mendatang dengan karakteristik data yang sama.

#### Daftar Pustaka

- William M. Bolstad. 2007. *Introduction to Bayesian Statistics*, 2nd ed. New Jersey: Wiley.
- George E. P. Box and George C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Boston: Addison-Wesley Publishing Company, 1973.

- Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2004. *Bayesian Data Analysis*, 2nd ed. New York: Chapman & Hall.

- Andrew Gelman. 2007. *Statistical Modeling, Causal Inference, and Social Science*. [Online]. <http://andrewgelman.com/2007/07/18/informative-and/>

- John Neter, William Wasserman, and Michael H. Kutner. 1983. *Applied Linear Regression*. Illinois: Richard D. Irwin.

- Robert E. Kass and Larry Wasserman. 1996. "The Selection of Prior Distribution by Formal Rules," *Journal of the American Statistical Association*, vol. 91, pp. 1343-1370.

- Bradley P. Carlin and Thomas A. Louis. 2009. *Bayesian Methods for Data Analysis*, 3rd ed. New York: Chapman & Hall.

- Radu V. Craiu and Jeffrey S. Rosenthal. 2014. "Bayesian Computation Via Markov Chain Monte Carlo," *Annual Review of Statistics and Its Application*, vol. I, pp. 179-201.

- Christophe Andrieu, Nando de Freitas, Arnaud Doucet, and Michael I. Jordan. 2003. "An Introduction to MCMC for Machine Learning," *Machine Learning*, vol. 50, pp. 5-43.

## PERBANDINGAN REGRESI *ROBUST* PENDUGA MM DENGAN METODE *RANDOM SAMPLE CONSENSUS* DALAM MENANGANI PENCILAN

NI PUTU NIA IRFAGUTAMI<sup>1</sup>, I GUSTI AYU MADE SRINADI<sup>2</sup>, I WAYAN SUMARJAYA<sup>3</sup>

<sup>1,2,3</sup>Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali  
e-mail: <sup>1</sup>1, <sup>2</sup>srinadiigustiayumade@yahoo.co.id, <sup>3</sup>sumarjaya@unud.ac.id,

### *Abstract*

*The presence of outliers in observation can result in biased in parameter estimation using ordinary least square (OLS). Robust regression MM-estimator is one of the estimations methods that able to obtain a robust estimator against outliers. Random sample consensus (ransac) is another method that can be used to construct a model for observations data and also estimating a robust estimator against outliers. Based on the study, ransac obtained model with less biased estimator than robust regression MM-estimator.*

**Keywords:** *Outlier, Robust Regression, MM-estimate, ransac*

### 1. Pendahuluan

Model regresi dapat menggambarkan hubungan antarpeubah kuantitatif, sehingga satu atau lebih peubah dapat digambarkan oleh peubah lainnya. Model regresi linear sederhana, yakni  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$  merupakan suatu model yang terdiri dari satu peubah bebas dan satu peubah respons. Parameter-parameter yang membentuk model selanjutnya diduga dengan menggunakan metode kuadrat terkecil (MKT).

Metode kuadrat terkecil merupakan metode pendugaan parameter yang dilakukan dengan meminimumkan jumlah simpangan kuadrat  $Q = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2$ . Asumsi kenormalan, kehomogenan ragam, dan tidak terjadi autokorelasi pada komponen sisaan atau galat merupakan syarat dalam penggunaan metode ini. Apabila asumsi ini terpenuhi maka penduga yang dihasilkan akan bersifat takbias. Namun, tidak jarang ditemukan kasus dengan asumsi kenormalan, kehomogenan ragam, dan tidak terjadi autokorelasi tidak terpenuhi. Salah satu penyebabnya adalah adanya pencilan

dalam data amatan yang memengaruhi proses analisis. Salah satu alternatif yang dapat digunakan untuk mengatasi permasalahan ini yakni dengan menerapkan metode regresi *robust*.

Regresi *robust* merupakan suatu metode yang dapat menghasilkan penduga parameter yang kekar terhadap pencilan. Metode ini tetap menggunakan keseluruhan data, termasuk data pencilan, namun dengan memberikan bobot yang kecil pada data pencilan. Metode lain yang dapat diterapkan yakni ransac diperkenalkan pertama kali oleh Fischler dan Bolles pada tahun 1981 sebagai suatu paradigma dalam pembentukan model untuk data eksperimen. Metode ini menggunakan sedikit mungkin himpunan data dan memperluas himpunan ini dengan sekumpulan data yang konsisten (Fischler dan Bolles, 1981). Perbandingan kedua metode yang sama-sama mampu menghasilkan penduga yang kekar terhadap pencilan merupakan fokus dalam penelitian ini. Ketidakbiasan penduga

<sup>1</sup> Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

<sup>2</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

yang dihasilkan oleh masing-masing metode akan menjadi acuan dalam penelitian ini.

## 2. Metode Penelitian

Data yang digunakan pada penelitian ini merupakan data simulasi dengan satu peubah bebas ( $X$ ) dan satu peubah respons ( $Y$ ). Data dibangkitkan dengan bantuan program R versi 2.15.3 Adapun tahapan yang dilakukan dalam penelitian ini, yakni (1) Pembangkitan data. Adapun pembangkitan data dilakukan dengan ketentuan, peubah bebas ( $X$ ) terdiri dari 100 amatan dengan nilai-nilai peubah  $X$  adalah 1, 2, ..., 100. Nilai sisaan yang dibangkitkan merupakan data yang berdistribusi normal dengan rata-rata nol dan ragam satu sebanyak 100. Pembangkitan data pencilan, yang mana pada penelitian pencilan minor merupakan data yang berdistribusi normal dengan rata-rata lima dan standar deviasi 0,1, sedangkan pencilan mayor merupakan data yang berdistribusi normal dengan rata-rata delapan dan standar deviasi 0,1. Adapun pencilan yang dibangkitkan yakni sebesar 5%, 10%, 15%, 20%, 25%, dan 30% dari data, baik pencilan minor maupun mayor. Penempatan pencilan pada bagian bawah, tengah dan bagian atas dari gugus data dilakukan untuk melihat pengaruh posisi pencilan terhadap proses analisis. Setelah terdapat data sisaan yang mengandung pencilan dan peubah bebas, maka dapat dibangkitkan data untuk peubah respon ( $Y$ ). Ditentukan terlebih dahulu hubungan peubah bebas dan peubah respons yaitu  $Y = \beta_0 + \beta_1 X + \varepsilon$ . Pada penelitian ini ditetapkan nilai  $\beta_0 = 2$  dan  $\beta_1 = 3$ . Berdasarkan nilai peubah bebas, koefisien regresi, dan nilai sisaan yang mengandung pencilan, maka akan didapat 36 kelompok data peubah respons yang mengandung pencilan. (2) Melakukan uji *Anderson-Darling* untuk melihat kenormalan data. (3) Melakukan analisis data dengan regresi *robust* penduga MM dan (4) *random sample consensus* untuk setiap data yang dimiliki. (5) Membandingkan penduga parameter yang dihasilkan oleh metode regresi

*robust* penduga MM dan ransac untuk melihat metode manakah yang menghasilkan penduga parameter yang lebih baik.

## 3. Hasil dan Pembahasan

Setelah dibangkitkan data peubah bebas, data sisaan, dan data peubah respon baik dengan pencilan maupun tanpa pencilan, maka dilakukan pendugaan parameter dengan menggunakan metode kuadrat terkecil. Terdapat dua asumsi yang harus dipenuhi dalam menggunakan metode ini, yakni data yang digunakan menyebar normal dan memiliki ragam yang homogen. Apabila kedua asumsi ini terpenuhi maka penduga yang dihasilkan akan bersifat takbias.

Metode kuadrat terkecil merupakan suatu metode pendugaan parameter dengan meminimumkan jumlah kuadrat galat. Jumlah kuadrat galat dapat dituliskan sebagai berikut:

$$S(\beta_j) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2, \quad (1)$$

Untuk meminimumkan jumlah kuadrat galat maka persamaan (1) akan diturunkan secara kalkulus terhadap parameter  $\beta_0$  dan  $\beta_1$ . Kedua persamaan turunan tersebut selanjutnya disamakan dengan nol. Dengan demikian diperoleh nilai-nilai penduga parameter sebagai berikut,

$$b_0 = \frac{\sum_{i=1}^n Y_i}{n} - b_1 \frac{\sum_{i=1}^n X_i}{n},$$

$$b_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}.$$

Setelah dilakukan analisis pada masing-masing kelompok data yang mengandung pencilan dengan menggunakan metode kuadrat terkecil, maka dihitung nilai sisaan untuk masing-masing model. Nilai sisaan ini selanjutnya digunakan untuk melakukan pengujian asumsi kenormalan dengan menggunakan uji *Anderson-Darling*.

Uji Anderson-Darling melibatkan fungsi distribusi kumulatif dari sebaran data. Misalkan data yang akan diuji diasumsikan berdistribusi normal dengan tingkat signifikan  $\alpha$ , maka statistik uji yang digunakan adalah  $A^2 = -n - S$ , dengan

$$S = \frac{1}{n} \sum_{i=1}^n [2i - 1] [\ln(F(Z_i)) + \ln(1 - F(Z_{n+1-i}))],$$

dengan

$$Z_i = \frac{x_i - \bar{x}}{s}.$$

Adapun hipotesis yang diuji yakni:

$H_0$  : data mengikuti sebaran normal

$H_1$  : data tidak mengikuti sebaran normal,

dengan kaidah keputusannya yakni jika  $A^2$  lebih besar dari nilai kritis uji Anderson-Darling maka tolak  $H_0$  dan jika  $A^2$  lebih kecil dari nilai kritis uji Anderson-Darling maka  $H_0$  gagal ditolak. Apabila nilai  $p$ -value lebih kecil dari  $\alpha$  maka tolak  $H_0$ , apabila sebaliknya, terima  $H_0$ .

Berdasarkan hasil uji *Anderson-Darling* pada data sisaan untuk masing-masing model yang diduga dengan metode kuadrat terkecil, didapat bahwa semua data yang diberikan pencilan, baik pencilan minor maupun mayor dengan persentase 5%, 10%, 15%, 20%, 25%, dan 30% pada bagian bawah, tengah, dan atas, memiliki  $p$ -value yang lebih kecil dari nilai alfa yang ditentukan, yakni  $\alpha = 0,05$ . Hal ini berarti bahwa semua data yang mengandung pencilan tidak memenuhi asumsi kenormalan atau dapat dikatakan data tidak menyebar normal. Sehingga akan diterapkan metode regresi *robust* penduga MM dan metode ransac dalam analisis.

Regresi *robust* merupakan salah satu metode regresi yang mampu menghasilkan penduga parameter yang *robust* (kekar) terhadap pencilan. Metode ini bekerja dengan memberikan bobot pada pencilan yang kemudian membatasi pengaruh pencilan tersebut. Penduga MM merupakan salah satu metode estimasi yang diperkenalkan pertama kali oleh Yohai (1987) sebagai metode yang menggabungkan metode penduga M (*Maximum*

*Likelihood*) dan penduga S (*Scale*). Regresi *robust* dengan penduga S akan menjamin penduga parameter memiliki nilai *breakdown point* yang tinggi dan regresi *robust* dengan penduga M akan menjamin penduga parameter memiliki efisiensi yang tinggi. Hal inilah yang menyebabkan metode regresi *robust* penduga MM memiliki tingkat efisiensi dan *breakdown point* yang tinggi.

Pendugaan parameter dengan metode penduga MM dilakukan dengan dua tahap. Pertama, pendugaan parameter regresi awal yang diawali dengan menghitung nilai penduga skala menggunakan metode penduga S. Kedua, menaksir penduga parameter regresi akhir dengan menggunakan metode penduga M. Adapun langkah-langkah pendugaan parameter dengan menggunakan metode penduga MM, yakni, (1) Menghitung nilai awal penduga parameter  $\tilde{\mathbf{b}}$  dengan menggunakan penduga S (*high breakdown point*), yang mana penduga awal yang dihasilkan tidak harus efisien. Kemudian gunakan penduga ini untuk menghitung residual,

$\varepsilon_{i,0} = \mathbf{y} - \mathbf{x}^T \tilde{\mathbf{b}}$ . (2) Dari nilai residual pada langkah pertama, dihitung penduga skala,  $s_n = s(\varepsilon_1(\beta), \varepsilon_2(\beta), \dots, \varepsilon_n(\beta))$  yang didefinisikan sebagai solusi dari

$$\frac{1}{n} \sum_{i=1}^n \rho_0 \left( \frac{\varepsilon_{i,0}}{s_n} \right) = \delta$$

dengan  $\delta$  merupakan konstanta yang memenuhi  $\delta = E(\phi, \rho(\infty))$ ,  $\phi$  berdistribusi normal standar dan  $\rho(u)$  merupakan fungsi objektif yang menurut (Rousseeuw dan Leroy, 1987) harus memenuhi, (i)  $\rho$  merupakan fungsi simetris dan terdiferensialkan secara kontinu, dan  $\rho(0) = 0$ , (ii) jika terdapat  $a > 0$ , maka  $\rho$  akan meningkat pada  $[0, a]$  dan konstan pada  $[a, \infty)$ , (iii)  $\frac{\rho'(a)}{\rho'(0)} = \frac{1}{2}$ . Fungsi objektif yang digunakan pada langkah ini adalah fungsi objektif *Tukey Bisquare*,

$$\rho(u) = \begin{cases} \frac{c^2}{6} \left( 1 - \left( 1 - \left( \frac{u}{c} \right)^2 \right)^3 \right); & |u| \leq c \\ \frac{c^2}{6} & ; |u| > c \end{cases}$$

dengan  $c = 1,56$  [3].

(3) Setelah mendapatkan nilai sisaan  $\varepsilon_{i,0}$  dan penduga skala  $s_n$ , hitung pembobot awal ( $w_{i,0}$ ) sesuai dengan fungsi pembobot *Tukey Bisquare*,

$$w_{i,0}(\varepsilon_i^*) = \begin{cases} \left( 1 - \left( \frac{\varepsilon_i^*}{c} \right)^2 \right)^2 & ; |u| \leq c \\ 0 & ; |u| > c \end{cases}$$

dengan  $\varepsilon_{i,0}^* = \frac{\varepsilon_{i,0}}{s_n}$  dan  $c = 4,68$  (Yohai, 1987). Kemudian dihitung nilai penduga MM sebagai:

$$b_l = (X^T W_{l-1} X)^{-1} X^T W_{l-1} Y$$

dengan  $W_{l-1}$  merupakan matriks diagonal yang elemen diagonalnya adalah nilai dari fungsi pembobot  $w_{i,l-1}$ , dengan  $l$  menyatakan iterasi. (4) Hitung nilai  $\sum_{i=1}^n |\varepsilon_{i,1}|$ , dengan  $\varepsilon_{i,1} = y - x^T b_l$ . (5) Ulangi langkah kedua hingga empat. Iterasi dihentikan ketika telah mendapatkan nilai  $\sum_{i=1}^n |\varepsilon_{i,1}|$  yang konvergen.

Setelah didapat penduga parameter dengan menggunakan metode regresi *robust* penduga MM, maka dicari selang kepercayaan masing-masing penduganya. Selang kepercayaan digunakan untuk melihat sifat

ketakbiasan penduga yang dihasilkan. Apabila selang kepercayaan mencakup nilai parameter yang sebenarnya, maka penduga parameter yang dihasilkan memiliki sifat takbias. Berdasarkan selang kepercayaan 95% diperoleh bahwa metode regresi *robust* menghasilkan penduga *intercept* ( $\beta_0$ ) yang berbias saat terdapat pencilan minor sebesar 10%, 15%, 20%, 25%, dan 30% pada bagian bawah gugus data, pencilan minor sebesar 30% pada bagian tengah gugus data, pencilan minor sebesar 20%, 25%, dan 30% pada bagian atas gugus data, dan saat terdapat pencilan mayor sebesar 30% pada bagian bawah gugus data.

Metode ini juga menghasilkan penduga *slope* ( $\beta_1$ ) yang berbias saat terdapat pencilan minor sebesar 10%, 15%, 20%, 25%, dan 30% pada bagian bawah gugus data, pencilan minor sebesar 15%, 20%, 25%, dan 30% pada bagian atas gugus data, dan saat terdapat pencilan mayor sebesar 30% pada bagian bawah gugus data. Untuk jenis data pencilan lainnya, metode regresi *robust* penduga MM menghasilkan penduga yang takbias. Hal ini lebih jelasnya dapat dilihat pada Tabel 1.

Tabel 1 Selang Kepercayaan 95%  $\beta_0$  dan  $\beta_1$  dengan Metode Regresi *Robust* Penduga MM

| Pencilan |        | % Pencil-an | $\beta_0$   |            | Keterangan | $\beta_1$   |            | Keterangan |
|----------|--------|-------------|-------------|------------|------------|-------------|------------|------------|
|          |        |             | Batas Bawah | Batas Atas |            | Batas Bawah | Batas Atas |            |
| Minor    | Bawah  | 5%          | 1.6173      | 2.4861     | Tidak Bias | 2.9915      | 3.0066     | Tidak Bias |
|          |        | 10%         | 2.1556      | 3.2633     | Bias       | 2.9800      | 2.9991     | Bias       |
|          |        | 15%         | 3.7913      | 5.3483     | Bias       | 2.9500      | 2.9766     | Bias       |
|          |        | 20%         | 4.8335      | 6.4103     | Bias       | 2.9358      | 2.9628     | Bias       |
|          |        | 25%         | 5.5666      | 6.9994     | Bias       | 2.9289      | 2.9535     | Bias       |
|          |        | 30%         | 6.1440      | 7.5736     | Bias       | 2.9225      | 2.9471     | Bias       |
|          | Tengah | 5%          | 1.7224      | 2.5447     | Tidak Bias | 2.9912      | 3.0054     | Tidak Bias |
|          |        | 10%         | 0.5606      | 3.7267     | Tidak Bias | 2.9846      | 3.0116     | Tidak Bias |
|          |        | 15%         | 1.6920      | 2.7926     | Tidak Bias | 2.9885      | 3.0076     | Tidak Bias |
|          |        | 20%         | 1.7275      | 3.2540     | Tidak Bias | 2.9854      | 3.0116     | Tidak Bias |
|          |        | 25%         | 1.8773      | 3.9812     | Tidak Bias | 2.9795      | 3.0156     | Tidak Bias |
|          |        | 30%         | 2.1315      | 4.5669     | Bias       | 2.9974      | 3.0191     | Tidak Bias |
|          | Atas   | 5%          | 1.7138      | 2.5477     | Tidak Bias | 2.9904      | 3.0047     | Tidak Bias |
|          |        | 10%         | 1.6207      | 2.5320     | Tidak Bias | 2.9917      | 3.0071     | Tidak Bias |
|          |        | 15%         | 0.5354      | 2.0460     | Tidak Bias | 3.0127      | 3.0385     | Bias       |
|          |        | 20%         | -0.1317     | 1.5643     | Bias       | 3.0329      | 3.0619     | Bias       |
|          |        | 25%         | -0.2122     | 1.3202     | Bias       | 3.0428      | 3.0690     | Bias       |
|          |        | 30%         | -0.2361     | 1.1859     | Bias       | 3.0502      | 3.0748     | Bias       |
| Mayor    | Bawah  | 5%          | 1.6137      | 2.4642     | Tidak Bias | 2.9919      | 3.0066     | Tidak Bias |
|          |        | 10%         | 1.6490      | 2.5417     | Tidak Bias | 2.9907      | 3.0062     | Tidak Bias |
|          |        | 15%         | 1.6721      | 2.5997     | Tidak Bias | 2.9900      | 3.0059     | Tidak Bias |
|          |        | 20%         | 1.6524      | 2.5843     | Tidak Bias | 2.9903      | 3.0061     | Tidak Bias |
|          |        | 25%         | 1.7562      | 2.7917     | Tidak Bias | 2.9871      | 3.0050     | Tidak Bias |
|          |        | 30%         | 8.7650      | 11.0408    | Bias       | 2.8752      | 2.9145     | Bias       |
|          | Tengah | 5%          | 1.7244      | 2.5402     | Tidak Bias | 2.9914      | 3.0052     | Tidak Bias |
|          |        | 10%         | 1.7175      | 2.5462     | Tidak Bias | 2.9910      | 3.0052     | Tidak Bias |
|          |        | 15%         | 1.6936      | 2.5612     | Tidak Bias | 2.9905      | 3.0056     | Tidak Bias |
|          |        | 20%         | 1.6856      | 2.5917     | Tidak Bias | 2.9905      | 3.0060     | Tidak Bias |
|          |        | 25%         | 1.6766      | 2.6268     | Tidak Bias | 2.9897      | 3.0060     | Tidak Bias |
|          |        | 30%         | 1.6463      | 2.8850     | Tidak Bias | 2.9873      | 3.0087     | Tidak Bias |
|          | Atas   | 5%          | 1.7153      | 2.5468     | Tidak Bias | 2.9904      | 3.0047     | Tidak Bias |
|          |        | 10%         | 1.6885      | 2.5351     | Tidak Bias | 2.9909      | 3.0056     | Tidak Bias |
|          |        | 15%         | 1.6943      | 2.5675     | Tidak Bias | 2.9902      | 3.0052     | Tidak Bias |
|          |        | 20%         | 1.7552      | 2.6586     | Tidak Bias | 2.9873      | 3.0028     | Tidak Bias |
|          |        | 25%         | 1.7428      | 2.6517     | Tidak Bias | 2.9879      | 3.0034     | Tidak Bias |
|          |        | 30%         | 1.7197      | 2.6484     | Tidak Bias | 2.9984      | 3.0042     | Tidak Bias |



Kemudian metode lain yang dapat digunakan untuk mendapat penduga yang kekar terhadap pencilan yakni metode ransac. Ransac diperkenalkan pertama kali oleh Fischler dan Bolles (1981) sebagai suatu metode pembentukan model bagi data eksperimen. Proses pembentukan model terbaik dengan menggunakan metode ini sangatlah berbeda dengan metode regresi *robust*. Apabila regresi *robust* menggunakan seluruh data dalam pendugaan parameter, maka *random sample consensus* menggunakan semimum mungkin data untuk membentuk model yang mampu menjelaskan keseluruhan data (Fischler dan Bolles, 1981).

Adapun proses pembentukan model dengan menggunakan metode ransac (3), yakni (1) dipilih 2 data secara acak dari sekumpulan 100 data. Selanjutnya kumpulan 2 data ini disebut sebagai himpunan  $S_1$ . (2) Bentuk model dengan menggunakan 2 data tersebut, sebut  $M_1$ . Pembentukan model dapat dilakukan dengan metode kuadrat terkecil (*least square*). (3) Berdasarkan model  $M_1$ , bentuk himpunan baru ( $S_1^*$ ) dari keseluruhan data yang mana anggota himpunannya merupakan data yang memiliki sisaan dalam toleransi *error* ( $t$ ). Himpunan ini selanjutnya disebut sebagai *consensus set*  $S_1$ . (4) Jika ( $S_1^*$ ) lebih besar dari batasan (*threshold*),  $T$ , maka gunakan  $S_1^*$  untuk membentuk model baru, yakni  $M_1^*$ . (5) Jika ( $S_1^*$ ) lebih kecil dari batasan (*threshold*),  $T$ , maka pilih secara acak himpunan baru  $S_2$  dan ulangi langkah (2).

Setelah didapatkan penduga parameter dengan menggunakan metode ransac, maka dicari selang kepercayaan masing-masing penduganya. Berdasarkan selang kepercayaan 95% diketahui bahwa metode ransac menghasilkan penduga parameter  $\beta_0$  yang berbias saat terdapat pencilan minor pada bagian bawah gugus data sebesar 25% dan

30%, dan pada bagian tengah gugus data sebesar 25% dan 30%. Ransac juga menghasilkan penduga parameter  $\beta_1$  yang berbias saat terdapat pencilan minor pada bagian bawah gugus data sebesar 25% dan 30% dan juga saat terdapat pencilan minor sebesar 25% dan 30% pada bagian atas gugus data. Untuk jenis data lainnya, ransac menghasilkan penduga parameter yang bersifat takbias. Hal ini lebih jelasnya dapat dilihat pada Tabel 2.

Tabel 2 Selang Kepercayaan 95%  $\beta_0$  dan  $\beta_1$  dengan Metode Ransac

| Pencilan |        | % Pencilan | $\beta_0$   |            | Keterangan | $\beta_1$   |            | Keterangan |
|----------|--------|------------|-------------|------------|------------|-------------|------------|------------|
|          |        |            | Batas Bawah | Batas Atas |            | Batas Bawah | Batas Atas |            |
| Minor    | Bawah  | 5%         | 1.5705      | 2.4361     | Tidak Bias | 2.9926      | 3.0070     | Tidak Bias |
|          |        | 10%        | 1.5318      | 2.5237     | Tidak Bias | 2.9914      | 3.0075     | Tidak Bias |
|          |        | 15%        | 1.4675      | 2.6009     | Tidak Bias | 2.9904      | 3.0083     | Tidak Bias |
|          |        | 20%        | 1.5143      | 2.7695     | Tidak Bias | 2.9882      | 3.0076     | Tidak Bias |
|          |        | 25%        | 6.3417      | 7.3958     | Bias       | 2.9271      | 2.9446     | Bias       |
|          |        | 30%        | 7.1327      | 8.0673     | Bias       | 2.9200      | 2.9359     | Bias       |
|          | Tengah | 5%         | 1.7240      | 2.4811     | Tidak Bias | 2.9923      | 3.0052     | Tidak Bias |
|          |        | 10%        | 1.7135      | 2.4708     | Tidak Bias | 2.9922      | 3.0050     | Tidak Bias |
|          |        | 15%        | 1.6880      | 2.4685     | Tidak Bias | 2.9922      | 3.0056     | Tidak Bias |
|          |        | 20%        | 1.6812      | 2.4827     | Tidak Bias | 2.9923      | 3.0056     | Tidak Bias |
|          |        | 25%        | 1.7051      | 2.5243     | Tidak Bias | 2.9908      | 3.0044     | Tidak Bias |
|          |        | 30%        | 1.8892      | 2.7211     | Tidak Bias | 2.9893      | 3.0030     | Tidak Bias |
|          | Atas   | 5%         | 1.7112      | 2.4953     | Tidak Bias | 2.9909      | 3.0050     | Tidak Bias |
|          |        | 10%        | 1.6711      | 2.4758     | Tidak Bias | 2.9912      | 3.0065     | Tidak Bias |
|          |        | 15%        | 1.7948      | 2.6465     | Tidak Bias | 2.9878      | 3.0050     | Tidak Bias |
|          |        | 20%        | 1.7147      | 2.5840     | Tidak Bias | 2.9869      | 3.0054     | Tidak Bias |
|          |        | 25%        | 0.3074      | 1.2391     | Bias       | 3.0539      | 3.0700     | Bias       |
|          |        | 30%        | 0.1732      | 1.0939     | Bias       | 3.0596      | 3.0752     | Bias       |
| Mayor    | Bawah  | 5%         | 1.5705      | 2.4361     | Tidak Bias | 2.9926      | 3.0070     | Tidak Bias |
|          |        | 10%        | 1.5318      | 2.5237     | Tidak Bias | 2.9914      | 3.0075     | Tidak Bias |
|          |        | 15%        | 1.4667      | 2.6009     | Tidak Bias | 2.9904      | 3.0083     | Tidak Bias |
|          |        | 20%        | 1.5143      | 2.7695     | Tidak Bias | 2.9882      | 3.0076     | Tidak Bias |
|          |        | 25%        | 1.5215      | 2.9693     | Tidak Bias | 2.9856      | 3.0074     | Tidak Bias |
|          |        | 30%        | 1.0735      | 2.7088     | Tidak Bias | 2.9892      | 3.0131     | Tidak Bias |
|          | Tengah | 5%         | 1.7240      | 2.4811     | Tidak Bias | 2.9923      | 3.0052     | Tidak Bias |
|          |        | 10%        | 1.7135      | 2.4708     | Tidak Bias | 2.9922      | 3.0050     | Tidak Bias |
|          |        | 15%        | 1.6880      | 2.4685     | Tidak Bias | 2.9922      | 3.0052     | Tidak Bias |
|          |        | 20%        | 1.6812      | 2.4827     | Tidak Bias | 2.9923      | 3.0056     | Tidak Bias |
|          |        | 25%        | 1.7051      | 2.5243     | Tidak Bias | 2.9908      | 3.0044     | Tidak Bias |
|          |        | 30%        | 1.7544      | 2.5698     | Tidak Bias | 2.9911      | 3.0045     | Tidak Bias |
|          | Atas   | 5%         | 1.7112      | 2.4953     | Tidak Bias | 2.9909      | 3.0050     | Tidak Bias |
|          |        | 10%        | 1.6711      | 2.4758     | Tidak Bias | 2.9912      | 3.0065     | Tidak Bias |
|          |        | 15%        | 1.6672      | 2.5029     | Tidak Bias | 2.9900      | 3.0068     | Tidak Bias |
|          |        | 20%        | 1.8465      | 2.7319     | Tidak Bias | 2.9845      | 3.0034     | Tidak Bias |
|          |        | 25%        | 1.6924      | 2.5707     | Tidak Bias | 2.9869      | 3.0069     | Tidak Bias |
|          |        | 30%        | 1.6605      | 2.5605     | Tidak Bias | 2.9867      | 3.0086     | Tidak Bias |

#### 4. Kesimpulan

Adanya pencilan pada bagian ujung gugus data memengaruhi proses pendugaan parameter. Hal ini dilihat dari penduga yang dihasilkan oleh metode regresi *robust* penduga MM bersifat bias saat terdapat pencilan minor sebesar 10% hingga 30% pada bagian bawah gugus data, pencilan minor sebesar 15% hingga 30% pada bagian atas gugus data, dan pada saat terdapat pencilan mayor sebesar 30% pada bagian bawah gugus data. Sedangkan, metode ransac menghasilkan penduga yang berbias saat terdapat pencilan minor sebesar 25% hingga 30% baik pada bagian bawah gugus data maupun bagian atas gugus data

Dapat dilihat pula metode ransac menghasilkan lebih sedikit model yang mengandung penduga parameter berbias dibandingkan dengan metode regresi *robust* penduga MM. Dari 36 model yang dihasilkan oleh kedua metode, ransac hanya menghasilkan empat model yang mengandung penduga parameter berbias. Sedangkan dengan metode regresi *robust* penduga MM, terdapat 11 model yang mengandung penduga parameter berbias.

#### Daftar Pustaka

- Martin A. Fischler and Robert C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Application to Image Analysis and Automated Cartography," *Graphics and Image Processing*, vol. 24, no. 6, pp. 381-395, June 1981.
- Peter J. Rousseeuw and Annick M. Leroy, *Robust Regression and Outlier Detection*. Canada: John Wiley & Sons, Inc., 1987.
- Victor J. Yohai, "High Breakdown-Point and High Efficiency Robust Estimates for Regression," *The Annals of Statistics*, vol. 15, no. 2, pp. 642-656, Jun 1987.

## PENGELOMPOKAN BERBAGAI MERK MI INSTAN BERDASARKAN KEMIRIPAN KANDUNGAN GIZI DENGAN MENGUNAKAN ANALISIS BILOT

AGUSTINUS ANGELAUS ETE<sup>1</sup>, NI LUH PUTU SUCIPTAWATI<sup>2</sup>, DESAK PUTU EKA  
NILAKUSMAWATI<sup>3</sup>

<sup>1,2,3</sup>Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali  
e-mail:<sup>1</sup>agustinusete@yahoo.com, <sup>2</sup>putusuciptawati@yahoo.co.id, <sup>3</sup>nilakusmawati\_desak@yahoo.com

### *Abstract*

*At this time, almost everyone once to consume instant noodles. The high interest of public on the instant noodles should be balanced with enough knowledge about the noodles and its nutritional content, either on it's instant noodles which have similar nutrient content and nutrient content that become identifier of each this group of noodles. The method can be used to obtain information on several brands of instant noodles that have similar nutrient content and nutrient content type that become identifier of each group of instant noodles is biplot analysis. Biplot analysis can show mie and nutrient content types simultaneously in a two-dimension plot. So that from a plot shows noodles and nutritional content types simultaneously, so that obtain information about the instant noodle that have similar nutrient content and nutrient content types into identifier of each group of instant noodles. This study was used 33 brands of instant noodles as observed objects with the type of nutrient content were observed there were nine. This study aims to find out some instant noodles that have similar nutrient content and nutrient content type that become identifier of each group of instant noodles. From the biplot analysis, obtained six groups of instant noodles with different identifier variables.*

**Keywords:** *Biplot Analysis, Similarity, Type of Nutrition*

### **1. Pendahuluan**

Dewasa ini, hampir semua orang pernah mengkonsumsi mi instan. Sifatnya yang praktis, mudah dihidangkan, rasanya enak, dan harganya yang terjangkau membuat mi instan menjadi daya tarik yang luar biasa bagi semua kalangan, mulai dari anak-anak hingga orang dewasa. Tingginya minat masyarakat terhadap mi instan perlu diimbangi dengan pengetahuan yang cukup tentang mi dan kandungan gizinya, baik itu tentang mi instan-mi instan yang memiliki kemiripan kandungan gizi maupun jenis kandungan gizi yang menjadi penciri dari masing-masing kelompok mi instan. Mengkonsumsi mi instan secara rutin tanpa mengetahui kandungan gizi yang terkandung di dalam mi instan bisa berdampak buruk bagi kesehatan.

Untuk mengetahui kemiripan kandungan gizi antarmi, maka perlu dilakukan pengelompokan mi instan dengan menggunakan analisis biplot. Analisis biplot dapat memberikan gambaran mengenai kedekatan antarmi berdasarkan kemiripan kandungan gizinya dalam bentuk plot dua dimensi sehingga baik digunakan sebagai acuan untuk melakukan pengelompokan. Mi instan yang memiliki kemiripan kandungan gizi akan digambarkan sebagai dua titik dengan posisi yang berdekatan dan mendapatkan sumbangan keragaman yang relatif sama dari vektor jenis kandungan gizi yang dimaksud. Jadi dengan menggunakan analisis biplot, bisa diperoleh tambahan informasi yang lebih baik tentang mi instan dan kandungan gizinya, baik itu tentang beberapa mi instan yang memiliki kemiripan kandungan gizi

<sup>1</sup> Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

<sup>2</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

maupun jenis kandungan gizi yang menjadi penciri dari masing-masing kelompok mi instan.

## 2. Ulasan Pustaka

### 2.1 Analisis Biplot

Analisis biplot merupakan salah satu teknik peubah ganda yang menyajikan plot pengamatan  $n$  dan peubah  $p$  secara bersamaan dalam bidang dua dimensi (Jolliffe, I.T. 2002). Penyajian plot pengamatan  $n$  dan peubah  $p$  secara bersamaan dapat memberikan tambahan informasi yang lebih baik tentang hubungan antara peubah dan pengamatan.

Empat hal penting yang bisa didapatkan dari tampilan biplot adalah (Mattjik, A.A. dan Sumertajaya, I M., 2011):

1. Kedekatan antarobjek yang diamati.  
Kedekatan antarobjek diinterpretasikan sebagai kemiripan sifat dua objek. Semakin dekat letak dua objek maka kemiripan sifat dua objek tersebut semakin tinggi.
2. Keragaman peubah.  
Keragaman peubah bisa dilihat dari panjang vektor peubah. Peubah dengan nilai keragaman kecil akan digambarkan sebagai vektor pendek sedangkan peubah dengan nilai keragaman yang tinggi akan digambarkan sebagai vektor panjang.
3. Korelasi antarpeubah.  
Dua peubah dikatakan memiliki korelasi positif apabila digambarkan sebagai dua buah vektor yang membentuk sudut lancip. Dua peubah dikatakan memiliki korelasi negatif apabila digambarkan sebagai dua buah vektor yang membentuk sudut tumpul. Sedangkan dua peubah dikatakan tidak memiliki korelasi apabila digambarkan sebagai dua buah vektor yang membentuk sudut siku-siku.
4. Nilai peubah pada suatu objek.  
Objek yang terletak searah dengan arah dari vektor peubah, memiliki nilai di atas rata-rata. Sebaliknya, objek yang terletak berlawanan dengan arah vektor peubah, memiliki nilai di bawah rata-rata. Nilai peubah pada suatu objek digunakan untuk melihat peubah penciri dari setiap objek.

Nilai peubah pada suatu objek dapat dilihat dengan melakukan proyeksi ortogonal dari objek ke vektor peubah.

### 2.1.1 Penguraian Nilai Singular

Analisis biplot merupakan gambaran dua dimensi dari suatu matriks data  $X$  yang menampilkan titik untuk masing-masing vektor pengamatan  $n$  (vektor baris matriks  $X$ ) bersama dengan titik untuk masing-masing  $p$  peubah (vektor kolom matriks  $X$ ) (Rencher, Alvin C., 2001). Dengan demikian biplot dibangun dari suatu matriks data  $X$ , dengan masing-masing baris mewakili objek penelitian dan masing-masing kolom mewakili suatu peubah.

$${}_nX_p = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

matriks  $X$  memuat peubah-peubah yang akan diteliti sebanyak  $p$  dan objek penelitian sebanyak  $n$ .

Analisis biplot didasarkan pada penguraian nilai singular. Definisi dari penguraian nilai singular adalah suatu matriks  ${}_nX_p$  yang dapat diuraikan menjadi:

$${}_nX_p = {}_nU_r \quad {}_rL_r \quad {}_pA'_r \quad (2)$$

Keterangan:

1.  $U$  dan  $A$  adalah matriks dengan kolom ortonormal ( $U'U = A'A = I_r$ ).  ${}_pA_r = [a_1, a_2, \dots, a_r]$  merupakan vektor ciri dari  $X'X$  dan  ${}_nU_r = \left[ \frac{Xa_1}{\sqrt{\lambda_1}}, \frac{Xa_2}{\sqrt{\lambda_2}}, \dots, \frac{Xa_r}{\sqrt{\lambda_r}} \right]$ , dengan  $\lambda_1, \lambda_2, \dots, \lambda_r$  merupakan akar ciri tidak nol dari  $X'X$ .
2.  $L$  adalah matriks diagonal berukuran  $r \times r$  dengan unsur-unsur diagonalnya  $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_r}$ . Unsur-unsur diagonal matriks  $L$  ini disebut nilai singular matriks  $X$ .  

$$L = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_r} \end{bmatrix} \quad (3)$$
3. Kolom-kolom matriks  $A$  adalah vektor ciri dari  $X'X$ .
4.  $r$  adalah pangkat matriks  $X$ ,  $r \leq \min(n, p)$ .

Kemudian didefinisikan  $L^\alpha$  untuk  $0 \leq \alpha \leq 1$  merupakan matriks diagonal berukuran  $r \times r$  dengan unsur-unsur diagonalnya  $\lambda_1^{\alpha/2}, \lambda_2^{\alpha/2}, \dots, \lambda_r^{\alpha/2}$ . Definisi  $L^\alpha$  berlaku juga untuk  $L^{1-\alpha}$ , sehingga diperoleh unsur-unsur diagonalnya  $\lambda_1^{1-\alpha/2}, \lambda_2^{1-\alpha/2}, \dots, \lambda_r^{1-\alpha/2}$ . Misalkan  $G = UL^\alpha$  dan  $H' = L^{1-\alpha}A'$  dengan  $0 \leq \alpha \leq 1$ . Persamaan (2) dapat ditulis menjadi (Jolliffe, I.T. 2002):

$$X = UL^\alpha L^{1-\alpha} A' = GH' \quad (4)$$

Matriks  $G$  memuat skor komponen utama yang merupakan titik-titik koordinat dari  $n$  objek dan matriks  $H$  memuat vektor eigen yang merupakan titik-titik koordinat dari  $p$  peubah. Gambaran biplot dari matriks data  $X$  diperoleh dengan mengambil dua kolom pertama dari matriks  $G$  dan dua kolom pertama dari matriks  $H$ .

### 2.1.2 Parameter $\alpha$

Ada dua nilai  $\alpha$  yang digunakan untuk mendefinisikan  $G = UL^\alpha$  dan  $H' = L^{1-\alpha}A'$ , yaitu  $\alpha = 0$  dan  $\alpha = 1$ . Jika  $\alpha = 0$ , maka diperoleh persamaan sebagai berikut:

1.  $G = U$  dan  $H' = LA'$ .
2.  $X'X = HH' = S(n-1)$ .

Hasil kali  $h_j'h_k$  sama dengan  $(n-1)$  kali kovarian  $s_{jk}$  antara peubah ke- $j$  dan peubah ke- $k$ , di mana  $j = 1, 2, \dots, p$ .

3.  $\|h_j\| = \sqrt{n-1}s_j$ .

Panjang vektor  $h_j'h_j$  memberikan gambaran keragaman peubah ke- $j$ .

4.  $r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj}s_{kk}}} = \frac{h_j'h_k}{\|h_j\|\|h_k\|} = \cos \theta$ .

Korelasi antara peubah ke- $j$  dan peubah ke- $k$  sama dengan nilai cosinus sudut antara vektor  $h_j$  dan  $h_k$ .

5.  $\delta_{hi}^2 = (n-1)d_{hi}^2$ .

Jarak Mahalanobis antara 2 pengamatan  $x_h$  dan  $x_i$  sebanding dengan jarak Euclidean antara dua pengamatan  $g_h$  dan  $g_i$ . Dengan  $\delta_{hi}^2 = (x_h - x_i)'S^{-1}(x_h - x_i)$  merupakan jarak Mahalanobis antara 2 pengamatan  $x_h$  dan  $x_i$ . Sedangkan  $d_{hi}^2 = (g_h - g_i)'(g_h -$

$g_i)$  merupakan jarak Euclidean antara dua pengamatan  $g_h$  dan  $g_i$ .

Secara keseluruhan pemilihan  $\alpha = 0$ , akan memberikan kesesuaian untuk data dan keragaman. Tampilan biplot akan lebih mampu menggambarkan keragaman peubah, hubungan antarpeubah sekaligus mendapatkan informasi mengenai pola objek. Jika  $\alpha = 0$  yang digunakan, maka hasil pemfaktoran disebut *GH* atau *CMP biplot (Column Metric Preserving)* (Lipkovich, I. and Smith, E.P., 2002).

Jika  $\alpha = 1$ , maka diperoleh persamaan sebagai berikut:

1.  $G = UL$  dan  $H' = A'$ .
2.  $(x_h - x_i)'(x_h - x_i) = (g_h - g_i)'(g_h - g_i)$ . Jarak Euclidean antara vektor pengaruh baris  $g_h$  dan  $g_i$  sama dengan jarak Mahalanobis antara  $x_h$  dan  $x_i$ .
3.  $x_{ij} = \sum_{k=1}^r z_{ik}a_{jk}$ .

Vektor pengaruh baris ke- $i$  ( $g_i$ ) sama dengan skor komponen utama ke- $k$  dari objek ke- $i$ . Di mana  $z_{ik} = u_{ik}\lambda_k^{1/2}$  adalah skor komponen utama ke- $k$  dari objek ke- $i$  dan  $a_{jk}$  adalah pembobot pada peubah ke- $j$  pada komponen utama ke- $k$ . Untuk  $G = UL$ , maka unsur ke- $k$  dari vektor pengaruh baris  $g_i$  sama dengan  $z_{ik}$  pada analisis komponen utama. Untuk  $H = A$ , maka vektor pengaruh kolom  $h_j$  sama dengan  $a_j$  pada analisis komponen utama.

Dengan menggunakan  $\alpha = 1$ , tampilan biplot akan lebih memberikan gambaran jarak antara pasangan baris sehingga baik digunakan untuk melihat kedekatan objek-objek. Jika  $\alpha = 1$  yang digunakan, maka hasil pemfaktoran disebut *JK* atau *RMP biplot (Row Metric Preserving)* (Lipkovich, I. and Smith, E.P., 2002).

### 2.1.3 Ukuran Kelayakan Biplot

Ukuran kelayakan biplot dua dimensi dirumuskan sebagai berikut (Gabriel, K.R., 1971):

$$\rho_2^{(2)} = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^r \lambda_k} \quad (5)$$

dengan  $\rho_2^{(2)}$  adalah ukuran kelayakan biplot dua dimensi untuk nilai  $\alpha$  bersesuaian,  $\lambda_1$  adalah

nilai eigen terbesar pertama,  $\lambda_2$  adalah nilai eigen terbesar kedua, dan  $\lambda_k$  adalah nilai eigen terbesar ke- $k$  dengan  $k = 1, 2, \dots, r$ . Apabila  $\rho_2^{(2)}$  mendekati nilai satu, maka biplot memberikan penyajian yang semakin baik mengenai informasi data yang sebenarnya.

## 2.2 Mi Instan

Mi instan dibuat dari adonan terigu atau tepung beras atau tepung lainnya sebagai bahan utama dengan atau tanpa penambahan bahan lainnya. Bahan-bahan yang digunakan untuk membuat mi instan dibagi menjadi dua, yaitu bahan utama dan bahan baku lain yang dapat ditambahkan. Bahan utama merupakan bahan dasar yang digunakan untuk membuat mi, seperti terigu, tepung beras, air, atau tepung lainnya. Bahan baku lain yang dapat ditambahkan merupakan bahan yang digunakan untuk membuat bumbu mi, seperti garam, gula, lemak, minyak, vitamin, mineral, bahan penyedap rasa, dan aroma yang diijinkan<sup>[1]</sup>.

Pada mi instan, terkandung berbagai macam zat gizi berupa protein, lemak berupa lemak jenuh, karbohidrat meliputi serat pangan dan gula, dan mineral berupa natrium. Selain keempat zat gizi di atas, pada kemasan mi instan juga dicantumkan informasi mengenai energi baik itu energi total maupun energi dari lemak.

## 3. Metode Penelitian

Jenis data yang digunakan dalam penelitian ini adalah data sekunder. Peneliti menggunakan 33 merk mi instan sebagai objek amatan dengan jenis kandungan gizi yang diamati sebanyak sembilan. Perangkat lunak (*software*) yang digunakan dalam penelitian ini adalah *Macros for Excel*.

Adapun sembilan jenis kandungan gizi yang digunakan sebagai peubah penelitian, yaitu:

1. ET : Energi Total (kkal)
2. EL : Energi dari Lemak (kkal)
3. LT : Lemak Total (%AKG)
4. LJ : Lemak Jenuh (%AKG)
5. P : Protein (%AKG)
6. KT : Karbohidrat Total (%AKG)

7. SP : Serat Pangan (%AKG)
8. G : Gula (mg)
9. N : Natrium (%AKG)

Metode analisis data dalam penelitian ini adalah sebagai berikut:

1. Melakukan Standardisasi data.
2. Membuat biplot data yang telah distandardisasi. Pada tahap ini meliputi penguraian nilai singular dan pembuatan plot.
3. Interpretasi hasil

## 4. Hasil dan Pembahasan

### 4.1 Data Penelitian

Peneliti menggunakan 33 merk mi instan sebagai objek amatan dengan kandungan gizi yang diamati sebanyak sembilan. Informasi mengenai nilai kandungan gizi dari 33 merk mi instan dapat dilihat pada Tabel 1.

Berdasarkan data kandungan gizi pada Tabel 1, terdapat lima macam merk mi instan yang digunakan dalam penelitian ini, yaitu sarimi, indomie, mie sedaap, supermi, dan mi ABC. Masing-masing merk mi instan memiliki rasa yang beraneka ragam. Misalnya sarimi, yang terdiri dari sarimi rasa ayam, sarimi mi goreng rasa ayam kecap, sarimi soto koya jeruk nipis, sarimi rasa kari spesial, dan sarimi rasa soto. Sekilas meskipun mi instan-mi instan tersebut dihasilkan dari produsen mi yang sama, tetapi nilai kandungan gizi yang terkandung di dalam masing-masing mi instan berbeda satu sama lain untuk beberapa jenis kandungan gizi. Perbedaan nilai kandungan gizi ini disebabkan karena penggunaan bahan dasar dan bahan tambahan yang tidak sama baik secara takaran maupun variasi bahan dari masing-masing merk mi instan.

Ada tiga satuan pengukuran yang digunakan untuk menyatakan nilai dari setiap jenis kandungan gizi yang terkandung dalam mi instan, yaitu kkal, %AKG, dan mg. Jenis kandungan gizi yang menggunakan satuan

pengukuran kkal adalah energi total (ET) dan energi dari lemak (EL). Jenis kandungan gizi yang menggunakan satuan pengukuran %AKG adalah lemak total (LT), lemak jenuh (LJ), protein (P), karbohidrat total (KT), serat pangan (SP), dan natrium (N). Sedangkan jenis kandungan gizi yang menggunakan satuan mg adalah gula (G).

Dalam analisis biplot, perhitungan jarak Euclidean, covarians, dan juga korelasi sangat

rentan terhadap perbedaan satuan pengukuran antarpeubah. Karena satuan pengukuran yang digunakan pada penelitian ini tidak sama, maka data yang digunakan perlu distandardisasi terlebih dahulu. Standardisasi peubah dilakukan dengan cara mengurangkannya terhadap rata-rata kolom dan membaginya dengan standar deviasi kolom.

Tabel 1. Data Kandungan Gizi dari 33 Merk Mi Instan (Sebelum Distandardisasi)

| No | Merk Mi  | Jenis Kandungan Gizi |     |    |    |    |    |    |    |    |
|----|----------|----------------------|-----|----|----|----|----|----|----|----|
|    |          | ET                   | EL  | LT | LJ | P  | KT | SP | G  | N  |
| 1  | SRA      | 310                  | 110 | 19 | 32 | 11 | 15 | 9  | 4  | 71 |
| 2  | SSKJN    | 320                  | 120 | 21 | 33 | 10 | 15 | 9  | 4  | 60 |
| 3  | SMGRAK   | 590                  | 220 | 39 | 65 | 18 | 27 | 16 | 13 | 59 |
| 4  | SRKS     | 540                  | 200 | 37 | 62 | 16 | 24 | 15 | 6  | 77 |
| 5  | SRS      | 520                  | 200 | 36 | 62 | 16 | 24 | 14 | 6  | 77 |
| 6  | IRAS     | 310                  | 110 | 19 | 30 | 12 | 15 | 9  | 4  | 63 |
| 7  | IMGP     | 360                  | 140 | 24 | 44 | 13 | 16 | 8  | 4  | 33 |
| 8  | IMGRAP   | 590                  | 240 | 43 | 96 | 23 | 25 | 10 | 14 | 61 |
| 9  | IRKA     | 320                  | 100 | 19 | 28 | 13 | 16 | 6  | 4  | 58 |
| 10 | IMGRS    | 370                  | 130 | 24 | 35 | 13 | 17 | 7  | 8  | 45 |
| 11 | IMGKP    | 400                  | 140 | 25 | 48 | 14 | 19 | 15 | 9  | 43 |
| 12 | IMGKB    | 430                  | 150 | 27 | 42 | 14 | 20 | 14 | 9  | 51 |
| 13 | IMGRR    | 420                  | 170 | 30 | 49 | 16 | 18 | 16 | 9  | 51 |
| 14 | IRAB     | 320                  | 120 | 22 | 36 | 12 | 14 | 6  | 3  | 64 |
| 15 | IMGRCI   | 410                  | 180 | 32 | 53 | 13 | 17 | 14 | 8  | 41 |
| 16 | IMGC     | 420                  | 150 | 26 | 60 | 16 | 20 | 10 | 5  | 46 |
| 17 | IRSM     | 310                  | 120 | 21 | 33 | 12 | 14 | 10 | 3  | 60 |
| 18 | IMKRAP   | 410                  | 160 | 29 | 45 | 18 | 17 | 12 | 7  | 65 |
| 19 | IMKGS    | 410                  | 160 | 28 | 42 | 18 | 18 | 12 | 9  | 48 |
| 20 | MSRAS    | 320                  | 120 | 21 | 35 | 13 | 14 | 8  | 3  | 41 |
| 21 | MSMKRS   | 360                  | 140 | 25 | 45 | 14 | 15 | 9  | 3  | 45 |
| 22 | MSMKRKA  | 340                  | 130 | 23 | 41 | 14 | 15 | 8  | 3  | 43 |
| 23 | MSRKS    | 360                  | 140 | 26 | 50 | 14 | 15 | 8  | 3  | 48 |
| 24 | MSMKRAB  | 320                  | 120 | 22 | 42 | 13 | 14 | 9  | 3  | 41 |
| 25 | SPRAS    | 340                  | 130 | 23 | 38 | 12 | 16 | 9  | 5  | 63 |
| 26 | SPRS     | 340                  | 130 | 23 | 34 | 12 | 16 | 11 | 5  | 69 |
| 27 | SPMG     | 410                  | 150 | 27 | 56 | 13 | 19 | 11 | 7  | 38 |
| 28 | SPRSA    | 350                  | 140 | 25 | 42 | 11 | 16 | 19 | 5  | 53 |
| 29 | SPRAB    | 320                  | 110 | 20 | 34 | 12 | 15 | 11 | 3  | 63 |
| 30 | MABCRSAP | 330                  | 120 | 22 | 36 | 10 | 15 | 9  | 1  | 95 |
| 31 | MABCRSTP | 330                  | 120 | 22 | 38 | 10 | 16 | 9  | 1  | 90 |
| 32 | MABCRAB  | 310                  | 110 | 19 | 32 | 12 | 15 | 9  | 1  | 83 |
| 33 | MABCRGAP | 310                  | 110 | 21 | 33 | 10 | 16 | 9  | 3  | 90 |

## 4.2 Analisis Biplot

### 4.2.1 Pengelompokan Mi Instan

Untuk memperoleh plot dua dimensi dari masing-masing merk mi instan dan jenis kandungan gizi, maka langkah awal yang perlu dilakukan adalah mencari titik koordinat dari masing-masing merk mi dan jenis kandungan gizi. Koordinat dari masing-masing merk mi instan dapat diperoleh dari matriks  $G$  sedangkan koordinat dari masing-masing jenis kandungan gizi dapat diperoleh dari matriks  $H$ .

Berdasarkan data kandungan gizi yang sudah distandardisasi, dengan menggunakan bantuan *Macros for Excel*, maka diperoleh matriks  $G$  yang memuat skor komponen dari dua komponen utama pertama dari data sebagai berikut.



Tabel 2. Skor Komponen 1 dan Komponen 2 Data yang Distandardisasi

| No | Merk Mi  | Skor Komponen 1 | Skor Komponen 2 |
|----|----------|-----------------|-----------------|
| 1  | SRA      | -0.15896        | 0.090301        |
| 2  | SSKJN    | -0.14235        | 0.007353        |
| 3  | SMGRAK   | 0.427535        | 0.102982        |
| 4  | SRKS     | 0.278853        | 0.313828        |
| 5  | SRS      | 0.26241         | 0.298518        |
| 6  | IRAS     | -0.1521         | 0.00388         |
| 7  | IMGP     | -0.04495        | -0.2636         |
| 8  | IMGRAP   | 0.519936        | -0.01635        |
| 9  | IRKA     | -0.15884        | -0.07069        |
| 10 | IMGRS    | -0.03707        | -0.19054        |
| 11 | IMGKP    | 0.082446        | -0.13442        |
| 12 | IMGKB    | 0.09857         | -0.04236        |
| 13 | IMGRR    | 0.149335        | -0.0613         |
| 14 | IRAB     | -0.14749        | -0.00267        |
| 15 | IMGRCI   | 0.122221        | -0.13335        |
| 16 | IMGC     | 0.090839        | -0.11409        |
| 17 | IRSM     | -0.13873        | -0.00878        |
| 18 | IMKRAP   | 0.095842        | 0.017998        |
| 19 | IMKGS    | 0.112945        | -0.15503        |
| 20 | MSRAS    | -0.1287         | -0.21534        |
| 21 | MSMKRS   | -0.04329        | -0.15094        |
| 22 | MSMKRKA  | -0.07887        | -0.18859        |
| 23 | MSRKS    | -0.03581        | -0.13074        |
| 24 | MSMKRAB  | -0.10628        | -0.20648        |
| 25 | SPRAS    | -0.07879        | 0.021135        |
| 26 | SPRS     | -0.07722        | 0.096355        |
| 27 | SPMG     | 0.074325        | -0.17618        |
| 28 | SPRSA    | -0.00457        | 0.034224        |
| 29 | SPRAB    | -0.13397        | 0.035512        |
| 30 | MABCRSAP | -0.16088        | 0.373114        |
| 31 | MABCRSTP | -0.1484         | 0.335649        |
| 32 | MABCRAB  | -0.17622        | 0.224355        |
| 33 | MABCRGAP | -0.16176        | 0.306256        |

Setiap pasangan skor dari komponen 1 dan komponen 2 merupakan koordinat titik dari setiap merk mi instan pada plot dua dimensi.

Hal yang sama juga dilakukan untuk memperoleh matriks *H*. Untuk matriks *H*, dengan menggunakan bantuan *Macros for Excel* diperoleh matriks *H* yang memuat dua vektor eigen pertama yang merupakan titik-titik koordinat dari masing-masing jenis kandungan gizi. Dua Vektor eigen pertama dari data yang distandardisasi disajikan pada Tabel 3.

Tabel 3. Vektor Eigen 1 dan Vektor Eigen 2 dari Data yang Distandardisasi

| Jenis Kandungan Gizi | Vektor Eigen 1 | Vektor Eigen 2 |
|----------------------|----------------|----------------|
| ET                   | 0.98264846     | 0.107442468    |
| EL                   | 0.980140195    | 0.052882124    |
| LT                   | 0.975680706    | 0.086443833    |
| LJ                   | 0.915279474    | -0.026805601   |
| P                    | 0.863652378    | -0.209719768   |
| KT                   | 0.926975894    | 0.218594094    |
| SP                   | 0.601275551    | 0.173647016    |
| G                    | 0.865762037    | -0.219658275   |
| N                    | -0.159155514   | 0.96108148     |

Selanjutnya bila ingin diketahui sejauh mana biplot mampu memberikan informasi mengenai kemiripan kandungan gizi antarmi dalam plot dua dimensi, maka perlu dicari nilai eigen dari data awal yang distandardisasi dengan menggunakan bantuan *Macros for Excel*. Tabel 4 menyajikan nilai eigen yang diperoleh dari data kandungan gizi yang sudah distandardisasi.

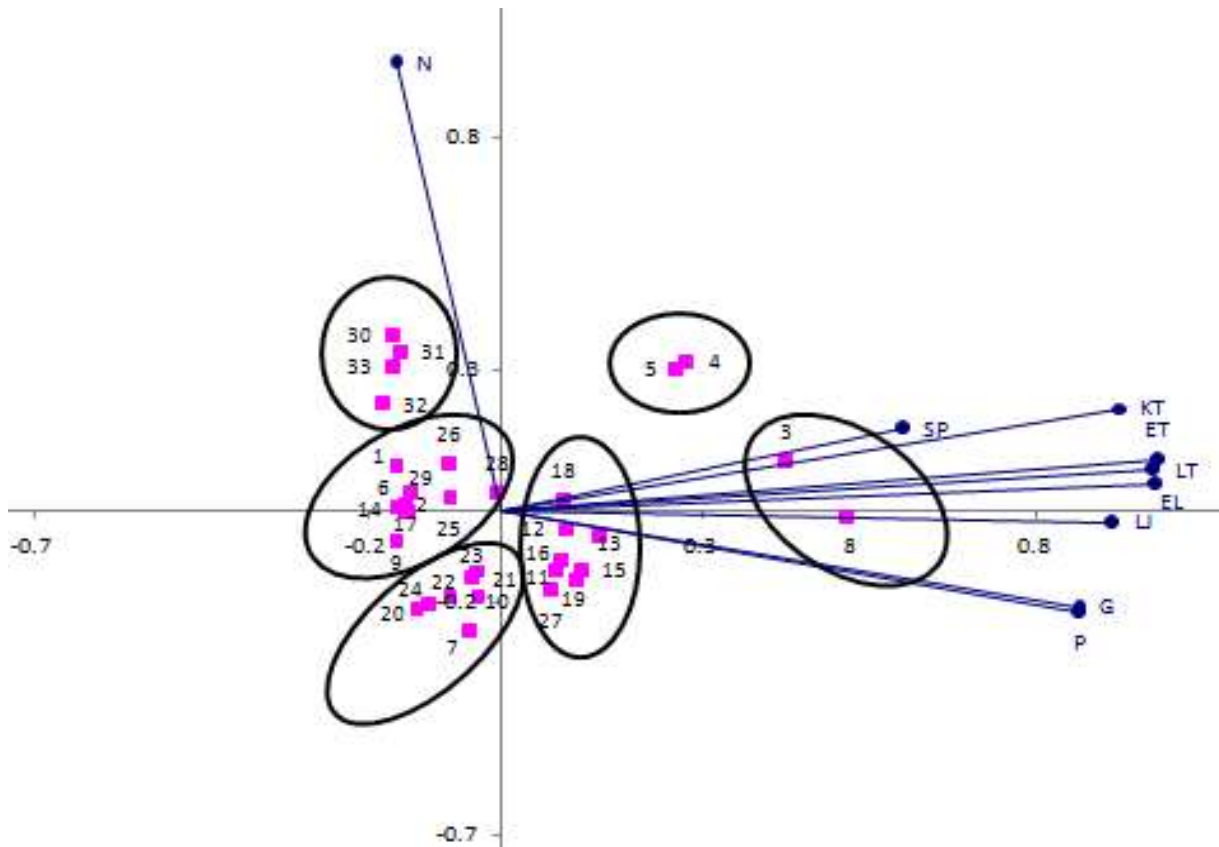
Tabel 4 Nilai Eigen dari Data Awal yang Distandardisasi

| Eigen Values | Cumulative % of Eigenvalues |
|--------------|-----------------------------|
| 6.457549     | 0.717505                    |
| 1.116378     | 0.841547                    |

Berdasarkan data pada Tabel 4, pasangan komponen utama 1 dan komponen utama 2 dapat menjelaskan keragaman data sebesar 84.15%. Dengan total keragaman data yang dapat dijelaskan dari kedua komponen utama pertama sebesar 84.15%, biplot dianggap sudah cukup memberikan informasi mengenai kemiripan kandungan gizi antarmi.

Untuk mengetahui mi instan-mi instan yang memiliki kemiripan kandungan gizi, maka perlu dilihat posisi antarmi pada plot dan juga posisi dari masing-masing mi terhadap setiap vektor jenis kandungan gizi. Mi instan yang memiliki kemiripan kandungan gizi akan digambarkan sebagai dua titik dengan posisi yang berdekatan dan mendapatkan sumbangan keragaman yang relatif sama dari vektor jenis kandungan gizi yang dimaksud. Berikut ini adalah penggambaran biplot yang menampilkan

titik untuk setiap merk mi instan dan vektor jenis kandungan gizi.



Gambar 1. Hasil Pemetaan Biplot untuk Mi dan Jenis Kandungan Gizi

Berdasarkan posisi dari 33 merk mi instan yang ditunjukkan pada Gambar 1, maka dapat dibentuk enam kelompok mi instan. Keenam kelompok mi instan tersebut antara lain sebagai berikut:

1. Kelompok pertama, yaitu sarimi mi goreng rasa ayam kecap (3) dan indomie mi goreng rasa ayam panggang (8).
2. Kelompok kedua, yaitu sarimi rasa kari spesial (4) dan sarimi rasa soto (5).
3. Kelompok ketiga, yaitu mi ABC rasa semur ayam pedas (30), mi ABC rasa sup tomat pedas (31), mi ABC rasa ayam bawang (32), dan mi ABC rasa gulai ayam pedas (33).
4. Kelompok keempat, yaitu indomie mi goreng kriuk pedas (11), indomie mi goreng kriuk bawang (12), indomie mi goreng rasa rendang (13), indomie mi goreng rasa cabe ijo (15), indomie mi goreng cakalang (16), indomie mi keriting rasa ayam panggang (18), indomie mi keriting goreng spesial (19), dan supermi mi goreng (27).
5. Kelompok kelima, yaitu sarimi rasa ayam (1), sarimi soto koya jeruk nipis (2), indomie rasa ayam spesial (6), indomie rasa kari ayam (9), indomie rasa ayam bawang (14), indomie rasa soto mie (17), supermi rasa ayam spesial (25), supermi rasa soto (26), supermi rasa semur ayam (28), dan supermi rasa ayam bawang (29).
6. Kelompok keenam, yaitu indomie mi goreng pedas (7), indomie mie goreng rasa sate (10), mie sedaap rasa ayam spesial (20), mie sedaap mie kuah rasa soto (21), mie sedaap mie kuah rasa kari ayam (22), mie sedaap rasa kari spesial (23), mie sedaap mie kuah rasa ayam bawang (24).

Setelah terbentuknya kelompok-kelompok mi instan, maka akan dilihat vektor jenis kandungan gizi mana yang memberikan

sumbangan keragaman yang relatif besar terhadap masing-masing kelompok mi instan yang selanjutnya disebut sebagai peubah penciri dari masing-masing kelompok mi tersebut. Untuk itu akan dilakukan plot ulang hasil pengelompokan mi tadi, di mana data yang digunakan merupakan data nilai rata-rata dari setiap jenis kandungan gizi yang dimiliki masing-masing kelompok mi instan. Tabel 5 memberikan informasi mengenai nilai rata-rata dari setiap jenis kandungan gizi yang dimiliki masing-masing kelompok mi instan. Perhitungan nilai rata-rata menggunakan data kandungan gizi pada Tabel 1.

Tabel 5 Nilai Rata-Rata Setiap Jenis Kandungan Gizi dari Masing-Masing Kelompok Mi (Sebelum Distandardisasi)

| Kelompok Mi | ET     | EL     | LT    | LJ    | P     | KT    | SP   | G    | N    |
|-------------|--------|--------|-------|-------|-------|-------|------|------|------|
| Kelompok 1  | 590    | 230    | 41    | 80.5  | 20.5  | 26    | 13   | 13.5 | 60   |
| Kelompok 2  | 530    | 200    | 36.5  | 62    | 16    | 24    | 14.5 | 6    | 77   |
| Kelompok 3  | 320    | 115    | 21    | 34.75 | 10.5  | 15.5  | 9    | 1.5  | 89.5 |
| Kelompok 4  | 413.75 | 157.5  | 28    | 49.38 | 15.25 | 18.5  | 13   | 7.88 | 47.9 |
| Kelompok 5  | 324    | 119    | 21.2  | 34    | 11.7  | 15.2  | 9.9  | 4    | 62.4 |
| Kelompok 6  | 347.14 | 131.43 | 23.57 | 41.71 | 13.43 | 15.14 | 8.14 | 3.86 | 42.3 |

Karena satuan pengukuran dari beberapa jenis kandungan gizi pada Tabel 5 tidak sama, maka data yang digunakan perlu distandardisasi terlebih dahulu. Standardisasi peubah dilakukan dengan cara mengurangkannya terhadap rataan kolom dan membaginya dengan standar deviasi kolom.

Untuk memperoleh plot dua dimensi dari masing-masing kelompok mi dan jenis kandungan gizi, maka langkah awal yang perlu dilakukan adalah mencari titik koordinat dari masing-masing kelompok mi dan jenis kandungan gizi. Koordinat dari masing-masing kelompok mi instan dapat diperoleh dari matriks *G*, sedangkan koordinat dari masing-masing jenis kandungan gizi dapat diperoleh dari matriks *H*.

Berdasarkan data kelompok mi yang sudah distandardisasi, dengan menggunakan bantuan *Macros for Excel*, maka diperoleh matriks *G* yang memuat skor komponen dari dua

komponen utama pertama dari data sebagai berikut.

Tabel 6. Skor Komponen 1 dan Komponen 2 Data yang Distandardisasi

| Kelompok Mi | Skor Komponen 1 | Skor Komponen 2 |
|-------------|-----------------|-----------------|
| Kelompok 1  | 0.685567        | -0.13223        |
| Kelompok 2  | 0.348697        | 0.504135        |
| Kelompok 3  | -0.43914        | 0.562315        |
| Kelompok 4  | 0.053534        | -0.32936        |
| Kelompok 5  | -0.3584         | -0.0567         |
| Kelompok 6  | -0.29026        | -0.54815        |

Setiap pasangan skor dari komponen 1 dan komponen 2 merupakan koordinat titik dari setiap kelompok mi instan pada plot dua dimensi.

Hal yang sama juga dilakukan untuk memperoleh matriks *H*. Untuk matriks *H*, dengan menggunakan bantuan *Macros for Excel* diperoleh matriks *H* yang memuat dua vektor eigen pertama yang merupakan titik-titik koordinat dari masing-masing jenis kandungan gizi. Dua Vektor eigen pertama dari data kelompok mi yang sudah distandardisasi disajikan pada Tabel 7.

Tabel 7. Vektor Eigen 1 dan Vektor Eigen 2 Data yang Distandardisasi

| Jenis Kandungan Gizi | Vektor Eigen 1 | Vektor Eigen 2 |
|----------------------|----------------|----------------|
| ET                   | 0.992152287    | 0.094452681    |
| EL                   | 0.995951675    | 0.042046503    |
| LT                   | 0.992004175    | 0.076702121    |
| LJ                   | 0.987933613    | -0.02236233    |
| P                    | 0.970124555    | -0.22391935    |
| KT                   | 0.978907471    | 0.197265237    |
| SP                   | 0.832722781    | 0.233572743    |
| G                    | 0.909426033    | -0.30343471    |
| N                    | -0.085824967   | 0.978837819    |

Setiap pasangan skor dari vektor eigen 1 dan vektor eigen 2 merupakan koordinat titik dari setiap jenis kandungan gizi pada plot dua dimensi.

Selanjutnya bila ingin diketahui sejauh mana biplot mampu memberikan informasi atas data dalam plot dua dimensi, maka perlu dicari nilai eigen dari data kelompok mi yang sudah distandardisasi dengan menggunakan bantuan *Macros for Excel*. Tabel 8 menyajikan nilai

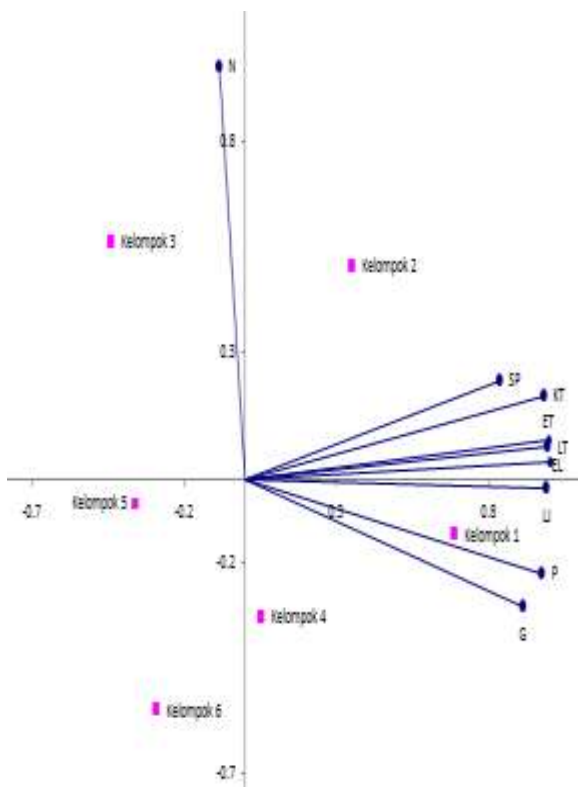
eigen dari data kelompok mi yang distandardisasi.

Tabel 8. Nilai Eigen dari Data Awal yang Distandardisasi

| Eigen Values | Cumulative % of Eigenvalues |
|--------------|-----------------------------|
| 7.363621     | 0.81818                     |
| 1.210878     | 0.952722                    |

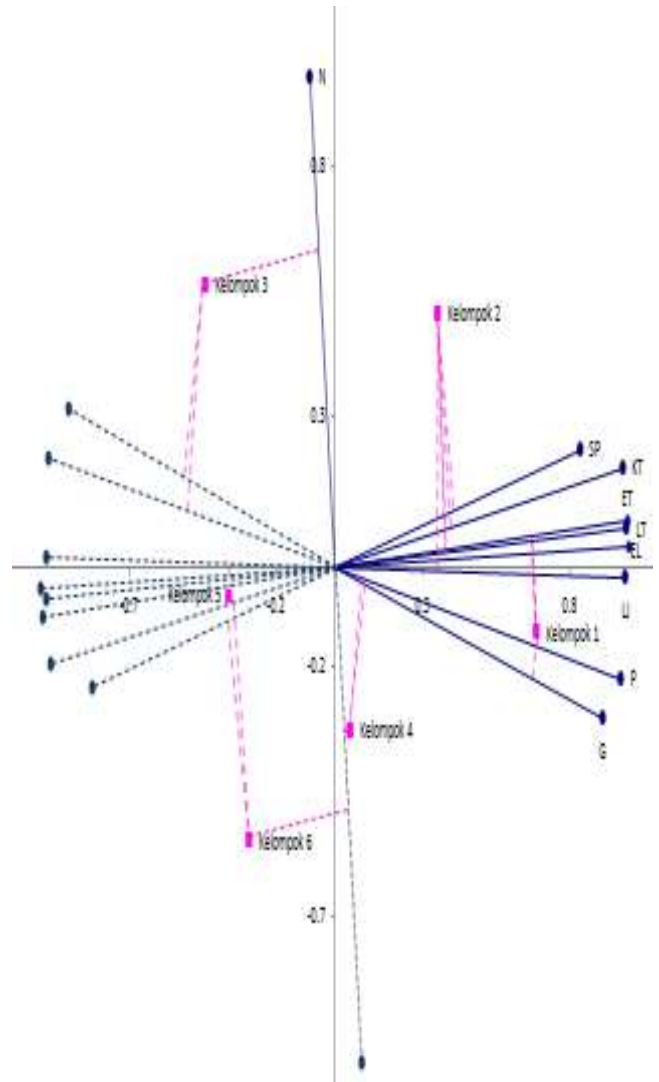
Berdasarkan data pada Tabel 8, pasangan komponen utama 1 dan komponen utama 2 dapat menjelaskan keragaman data sebesar 95.27%. Dengan total keragaman data yang dapat dijelaskan dari kedua komponen utama pertama sebesar 95.27%, maka biplot dianggap sudah cukup memberikan informasi mengenai jenis kandungan gizi yang menjadi penciri dari masing-masing kelompok mi.

Setelah diperoleh koordinat dari masing-masing kelompok mi instan dan jenis kandungan gizi, maka selanjutnya dilakukan plot ulang masing-masing kelompok mi instan dan jenis kandungan gizi. Gambar 2 menunjukkan hasil penggambaran biplot yang menampilkan titik untuk setiap kelompok mi dan vektor jenis kandungan gizi.



Gambar 2. Hasil Pemetaan Biplot untuk Kelompok Mi dan Vektor Jenis Kandungan Gizi

Untuk mengetahui jenis kandungan gizi yang menjadi penciri dari masing-masing kelompok mi instan maka perlu dilakukan proyeksi ortogonal dari masing-masing kelompok mi instan terhadap setiap vektor jenis kandungan gizi. Gambar 3 berikut menampilkan hasil proyeksi ortogonal dari setiap kelompok mi terhadap vektor jenis kandungan gizi yang menjadi pencirinya.



Gambar 3. Hasil Proyeksi Ortogonal dari Setiap Kelompok Mi Terhadap Vektor Jenis Kandungan Gizi yang Menjadi Penciri dari Setiap Kelompok Mi

Semakin jauh dari titik asal sebuah proyeksi dari objek jatuh pada vektor peubah, maka semakin besar nilai objek pada peubah tersebut. Sebaliknya, semakin dekat dengan titik

asal sebuah proyeksi dari objek jatuh pada vektor peubah, maka semakin kecil nilai objek pada peubah tersebut.

Gambar 3 menunjukkan bahwa mi instan kelompok 1 memiliki nilai kandungan gizi yang paling tinggi untuk jenis kandungan gizi protein (P), gula (G), lemak jenuh (LJ), dan energi lemak (EL). Dengan demikian jenis kandungan

gizi yang menjadi penciri dari mi instan pada kelompok 1 adalah protein (P), gula (G), lemak jenuh (LJ), dan energi lemak (EL).

Hasil proyeksi otogonal yang ditunjukkan pada Gambar 3, menunjukkan jenis kandungan gizi yang menjadi penciri dari masing-masing kelompok mi, dan terangkum dalam Tabel 9.

Tabel 9 Jenis Kandungan Gizi yang Menjadi Penciri untuk Setiap Kelompok Mi

| Kelompok Mi | Merk Mi   | Peubah Penciri   |
|-------------|---|--|
| Kelompok 1  | Sarimi Mi Goreng Rasa Ayam Kecap (3)<br>Indomie Mi Goreng Rasa Ayam Panggang (8)  | Protein (P)<br>Gula (G)<br>Lemak Jenuh (LJ)<br>Energi Lemak (EL)                     |
| Kelompok 2  | Sarimi Rasa Kari Spesial (4)<br>Sarimi Rasa Soto (5)  | Serat Pangan (SP)<br>Karbohidrat Total (KT)<br>Energi Total (ET)<br>Lemak Total (LT) |
| Kelompok 3  | Mi ABC Rasa Semur Ayam Pedas (30)<br>Mi ABC rasa Sup Tomat Pedas (31)<br>Mi ABC Rasa Ayam Bawang (32)<br>Mi ABC Rasa Gulai Ayam Pedas (33)  | Natrium (N)<br>Gula (G)<br>Serat Pangan (SP)   |
| Kelompok 4  | Indomie Mi Goreng Kriuk Pedas (11)<br>Indomie Mi Goreng Kriuk Bawang (12)<br>Indomie Mi Goreng Rasa Rendang (13)<br>Indomie Mi Goreng Rasa Cabe Ijo (15)<br>Indomie Mi Goreng Cakalang (16)<br>Indomie Mi Keriting Rasa Ayam Panggang (18)<br>Indomie Mi Keriting Goreng Spesial (19)<br>Supermi Mi Goreng (27)     | Natrium (N)<br>Gula (G)<br>Protein (P)   |
| Kelompok 5  | Sarimi Rasa Ayam (1)<br>Sarimi Soto Koya Jeruk Nipis (2)<br>Indomie Rasa Ayam Spesial (6)<br>Indomie Rasa Kari Ayam (9)<br>Indomie Rasa Ayam Bawang (14)<br>Indomie Rasa Soto Mie (17)<br>Supermi Rasa Ayam Spesial (25)<br>Supermi Rasa Soto (26)<br>Supermi Rasa Semur Ayam (28)<br>Supermi Rasa Ayam Bawang (29) | Energi Total (ET)<br>Lemak Total (LT)<br>Energi Lemak (EL)<br>Lemak Jenuh (LJ)       |
| Kelompok 6  | Indomie Mi Goreng Pedas (7)<br>Indomie Mie Goreng Rasa Sate (10)<br>Mie Sedaap Rasa Ayam Spesial (20)<br>Mie Sedaap Mie Kuah Rasa Soto (21)<br>Mie Sedaap Mie Kuah Rasa Kari Ayam (22)<br>Mie Sedaap Rasa Kari Spesial (23)<br>Mie Sedaap Mie Kuah Rasa Ayam Bawang (24)  | Natrium (N)<br>Serat Pangan (SP)<br>Karbohidrat Total (KT)                           |

Hasil yang ditampilkan pada Tabel 9 menunjukkan bahwa peubah penciri dari masing-masing kelompok mi tidak sama. Dari Tabel 9, diketahui bahwa mi instan-mi instan pada kelompok 3 dihasilkan dari produsen mi yang sama yaitu mi ABC. Hal ini menunjukkan bahwa meskipun mi ABC memiliki banyak aneka rasa tetapi nilai kandungan gizi yang terkandung di dalam setiap mi memiliki kemiripan untuk jenis

kandungan gizi natrium (N), gula (G), dan protein (P).

## 5. Kesimpulan

Dari hasil analisis biplot, diperoleh enam kelompok mi instan dengan peubah penciri yang berbeda.

## Daftar Pustaka

- Badan Standardisasi Nasional. 2000. *Mie Instan*. Tim Penyiapan Konsep Revisi SNI 01-3551-1996.
- Gabriel, K.R. 1971. *The Biplot Graphic Display of Matrices with Application to Principal Componen Analysis*. Jerusalem.
- Jolliffe, I.T. 2002. *Principal Component Analysis*. Springer-verlag. New York.
- Lipkovich, I. and Smith, E.P. 2002. *Biplot and Singular Value Decomposition Macros for Excel*. Blacksburg: Department of Statistics.
- Mattjik, A.A. and Sumertajaya, I M. 2011. *Sidik Peubah Ganda Dengan menggunakan SAS*. Bogor: IPB PRESS.
- Rencher, Alvin C. 2001. *Method of Multivariate Analysis*. United States of America: John Wiley & Sons, Inc.

## **PERHITUNGAN DANA PENSIUN DENGAN METODE *PROJECTED UNIT CREDIT* DAN *INDIVIDUAL LEVEL PREMIUM***

**I GUSTI AYU KOMANG KUSUMA WARDHANI<sup>1</sup>, I NYOMAN WIDANA<sup>2</sup>,  
NI KETUT TARI TASTRAWATI<sup>3</sup>**

<sup>1,2,3</sup>Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali  
e-mail: <sup>1</sup>kusumawardhanie0946@gmail.com, <sup>2</sup>nwidana@yahoo.com, <sup>3</sup>tastrawati@yahoo.com

### ***Abstract***

*The company which provides the pension program needs the actuarial valuation to estimate the fund amount required by the company to pay for pension funding. Actuarial method that used in this research are projected unit credit and individual level premium method. Through this research be obtained the result of valuation pension benefits with career average salary assumption is lower than the other salary assumptions. On the other hand, the result of normal cost final value valuation using individual level premium method is smaller than projected unit credit method that suits for the participants of the pension funding program.*

***Keywords:*** Pension Funding, Normal Cost, Projected Unit Credit, Individual Level Premium.

### **1. Pendahuluan**

Program dana pensiun merupakan salah satu bentuk perencanaan masa depan yang bertujuan untuk menjamin kelangsungan hidup pegawai pada masa pensiun. Merupakan bentuk tanggung jawab perusahaan terhadap pegawai yang telah mendedikasikan dirinya selama masa kerja. Program dana pensiun dapat memberikan rasa aman akan kelangsungan hidup pegawai setelah tidak aktif lagi serta menciptakan ketenangan bagi pegawai karena kesejahteraan pegawai pada hari tua telah terjamin.

Beberapa perusahaan telah menjamin kesejahteraan hari tua para pegawai dengan menyelenggarakan program dana pensiun (Aitken, 1994). Untuk itu, perusahaan membeli asuransi pensiun. Besar premi yang harus dibayarkan tiap pegawai dalam asuransi pensiun disebut dengan iuran normal. Pembayaran iuran normal dilakukan dalam bentuk pemotongan gaji pegawai kemudian diinvestasikan selama masa kerja yang memungkinkan terbentuknya akumulasi dana yang cukup untuk pembayaran manfaat pensiun dalam memelihara kesinam-

bugan penghasilan peserta pada hari tua. Pembayaran manfaat pensiun dilakukan ketika pegawai telah mencapai usia pensiun tertentu berdasarkan ketentuan-ketentuan dalam peraturan dana pensiun.

Beberapa penyebab seorang pegawai pensiun, yaitu kematian yang terjadi pada saat masih aktif bekerja sehingga dibayarkan uang pensiun janda/dudanya, memilih berhenti sebelum usia pensiun sehingga dibayarkan uang pensiun pada saat mencapai usia pensiun dengan memperhatikan masa kerjanya, pegawai yang masih aktif bekerja menjadi cacat sehingga tidak bisa bekerja lagi, atau pegawai yang telah mencapai usia pensiun sehingga uang pensiun segera dibayarkan (Futami, 1993b).

Banyak pegawai yang pensiun pada suatu perusahaan tidak dapat diprediksi, sehingga menyebabkan penurunan pegawai tidak menentu pada perusahaan. Hal ini mengakibatkan perusahaan harus membayarkan uang pensiun kepada pegawainya. Oleh karena itu, apabila perusahaan tidak mempersiapkan dan memperhitungkan pembayaran anuitas bagi pegawai dalam bentuk program dana pensiun

<sup>1</sup> Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

<sup>2</sup> Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

mengakibatkan keuangan perusahaan tidak stabil. Berdasarkan hal tersebut perlu dilakukan perhitungan khusus untuk memproyeksikan dana yang akan dikeluarkan perusahaan dalam membayar uang pensiun pegawainya. Besar manfaat pensiun yang akan diterima serta iuran normal yang harus dibayar oleh pegawai dapat dihitung dengan metode perhitungan aktuarial yang ada.

Metode perhitungan aktuarial yang dapat digunakan adalah metode *projected unit credit* dan *individual level premium*. Metode *projected unit credit* adalah metode perhitungan aktuarial dengan membagi total manfaat pensiun yang kemudian dialokasikan selama masa kerja, sedangkan metode *individual level premium* adalah metode perhitungan aktuarial dengan mengalokasikan total manfaat pensiun secara merata sejak tanggal perhitungan aktuarial. Kedua metode tersebut menggunakan asumsi skala gaji yang akan diestimasi pada masa depan (*future value*) dan diasumsikan bahwa gaji mengalami peningkatan. Terdapat tiga jenis penggunaan skala gaji dalam menghitung besar manfaat pensiun, yaitu rencana gaji terakhir, rencana rata-rata gaji  $n$  tahun terakhir dan rencana rata-rata gaji selama bekerja. Besar manfaat pensiun akan dibayarkan sama setiap tahun selama masa pensiun.

Beberapa penelitian tentang perhitungan dana pensiun telah dilakukan, antara lain oleh Pratiwi (2008) dan Hapsari (2012). Penelitian Pratiwi (2008) memperlihatkan bahwa besar iuran normal dan manfaat pensiun berdasarkan rata-rata gaji selama bekerja lebih kecil dibandingkan dengan perumusan berdasarkan manfaat penghasilan tetap, sedangkan berdasarkan rata-rata gaji terakhir, besar iuran normal dan manfaat pensiun lebih besar dibandingkan dengan manfaat penghasilan tetap. Hapsari (2012) menunjukkan bahwa besar iuran normal dengan menggunakan metode *projected unit credit* terus meningkat seiring dengan pertambahan gaji yang diterima, sedangkan apabila menggunakan metode *entry age normal* besarnya sama untuk tiap tahunnya pada seorang pegawai.

Berdasarkan hal tersebut, penulis tertarik untuk menghitung iuran normal yang dibayarkan tiap tahun oleh peserta program pensiun dan besar manfaat yang diperoleh berdasarkan tiga asumsi skala gaji seperti yang telah disebutkan di bagian terdahulu dengan metode *projected unit credit* dan metode *individual level premium*. Selanjutnya, juga akan dilakukan perbandingan hasil perhitungan antara kedua metode tersebut.

## 2. Ulasan Pustaka

### 2.1 Program Dana Pensiun

Program dana pensiun merupakan bentuk balas jasa pemerintah terhadap pegawai negeri yang telah bertahun-tahun mengabdikan dirinya kepada negara (Taspen, 2013).

Pada sistem dana pensiun, terdapat beberapa manfaat yang diberikan sebagai manfaat tambahan karena kematian, pensiun dipercepat (keluar), pensiun karena tidak bisa bekerja (cacat), dan pensiun pada saat usia pensiun. Adapun manfaat tambahannya adalah sebagai berikut (Futami, 1993b) :

1. Manfaat pensiun janda/duda yang dibayarkan karena meninggal.
2. Manfaat pensiun dipercepat yang dibayarkan bagi peserta yang berhenti bekerja atau keluar.
3. Manfaat pensiun cacat yang dibayarkan bagi peserta yang tidak bisa bekerja karena cacat.
4. Manfaat pensiun yang dibayarkan bagi peserta yang telah mencapai usia pensiun.

### 2.2 Tabel Penyusutan Jamak

Tabel penyusutan jamak (*multiple decrement table*) berisi peluang pegawai pensiun yang disebabkan oleh pensiun dipercepat (keluar), kematian, dan pensiun karena tidak bisa bekerja (cacat). Peluang pegawai pensiun di antara usia  $x$  dan  $x + 1$  tahun pada perusahaan yang disebabkan oleh pensiun dipercepat (keluar), kematian, dan pensiun karena tidak bisa bekerja (cacat), secara berturut-turut adalah (Futami, 1993b):



$$q_x^{(w)} = \frac{w_x}{l_x^{(T)}}, q_x^{(d)} = \frac{d_x^T}{l_x^{(T)}}, \text{ dan } q_x^{(i)} = \frac{i_x}{l_x^{(T)}}.$$

Oleh karena itu, peluang pegawai yang masih aktif bekerja di antara usia  $x$  dan  $x + 1$  tahun yang dinotasikan dengan  $p_x^{(T)}$  adalah

$$p_x^{(T)} = 1 - q_x^{(d)} - q_x^{(w)} - q_x^{(i)}. \quad (1.4)$$

### 2.3 Fungsi Manfaat

Fungsi manfaat digunakan untuk menentukan jumlah manfaat yang dibayarkan pada saat pensiun dipercepat (keluar), kematian, pensiun karena tidak bisa bekerja (cacat), atau pensiun pada saat usia pensiun. Jika  $B_r$  adalah besar total manfaat selama peserta aktif bekerja dari umur  $e$  tahun sampai  $r$  tahun, sedangkan besar manfaat yang diterima setiap tahunnya pada peserta yang berusia  $x$  tahun sebesar  $b_x$  (Aitken, 1994):

$$b_x = \frac{B_r}{r-e}. \quad (1.5)$$

Manfaat yang diperoleh peserta program pensiun merupakan proporsi gaji sebesar  $k$  persen yang diakumulasikan selama masa kerja  $(x - e)$  tahun berdasarkan tiga skala gaji, yaitu:

1. Asumsi gaji terakhir

Gaji terakhir pada usia  $r - 1$  tahun yang diharapkan yang dinotasikan dengan  $s_{r-1}$  dirumuskan dengan

$$s_{r-1} = (1 + s)^{r-1-x} s_x. \quad (1.6)$$

Dengan demikian, manfaat pensiun yang akan dibayar sampai usia  $x$  tahun yang dinotasikan dengan  $B_x$  dirumuskan dengan

$$B_x = k s_{r-1} (x - e). \quad (1.7)$$

2. Asumsi rata-rata gaji selama  $n$  tahun terakhir

Rata-rata gaji yang diharapkan  $n$  tahun terakhir (*Final Average Salary*) adalah

$$\begin{aligned} FAS &= \frac{1}{n} (1 + s)^{r-x-1} [(1 + s)^{1-n} + \dots + 1] s_x \\ &= \frac{1}{n} (1 + s)^{r-x-1} \ddot{a}_{n|s} s_x. \end{aligned} \quad (1.8)$$

Oleh karena itu, manfaat pensiun yang masih harus dibayar sampai berusia  $x$  tahun adalah

$$B_x = k FAS (x - e). \quad (1.9)$$

3. Asumsi rata-rata gaji selama bekerja  
Rata-rata gaji yang diharapkan selama bekerja adalah:

$$\begin{aligned} &\frac{1}{r-e} [s_e + \dots + s_x + s_{x+1} + \dots + s_{r-1}]. \\ &= \frac{1}{r-e} \left[ \frac{s_x}{(1+s)^{x-e}} + \dots + s_x + (1+s)s_x + \dots + (1+s)^{r-x-1} s_x \right] \\ &= \frac{1}{r-e} s_x (1+s)^{e-x} [1 + \dots + (1+s)^{x-e} + (1+s)^{x-e+1} + \dots + (1+s)^{r-e-1}]. \end{aligned} \quad (1.10)$$

Nilai akhir anuitas yang dilakukan selama  $r - e$  tahun dengan peningkatan sebesar  $s$  dinotasikan dengan  $s_{r-e|s}$ , sehingga persamaan (1.10) menjadi:

$$\frac{1}{r-e} s_x (1+s)^{e-x} s_{r-e|s}.$$

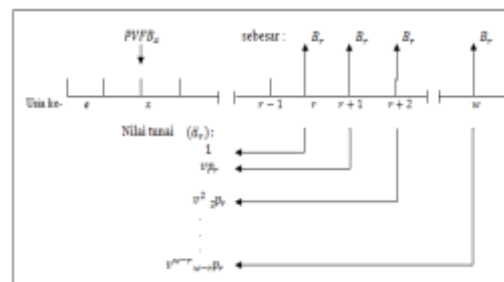
Oleh karena itu, manfaat yang diperoleh pegawai sampai berusia  $x$  tahun adalah:

$$B_x = \frac{k}{r-e} [s_x (1+s)^{e-x} s_{r-e|s}] (x - e).$$

*Present value of future benefit* (PVFB) adalah nilai sekarang dari manfaat pensiun yang akan diterima peserta program dana pensiun saat peserta memasuki usia pensiun yaitu saat peserta berusia  $r$  tahun. Pembayaran manfaat pensiun dilakukan tiap tahun sampai peserta meninggal. Winklevoss (1993) dalam Oktiani (2013), menyatakan  ${}^r(PVFB)_x$  dirumuskan sebagai berikut :

$${}^r(PVFB)_x = B_r v^{r-x} \ddot{a}_{r-r-x} p_x.$$

Rumus tersebut dapat dijelaskan dengan skema pembayaran seperti disajikan pada Gambar 1:



Gambar 1. Skema Pembayaran Manfaat

Berdasarkan skema pembayaran pada Gambar 1, maka nilai keseluruhan pembayaran di tahun ke- $r$  adalah:

$$B_r [1 + vp_r + v^2 {}_2p_r + \dots + v^{w-r} {}_{w-r}p_r] = B_r \ddot{a}_r.$$

Selanjutnya, anuitas manfaat akan diproyeksikan pada tahun sekarang (pada saat peserta berusia  $x$  tahun), sehingga nilai sekarang manfaat pensiun saat peserta berusia  $x$  tahun yang dinotasikan dengan  ${}^r(PVFB)_x$  adalah:

$${}^r(PVFB)_x = B_r \ddot{a}_r v^{r-x} {}_{r-x}p_x^{(T)}. \quad (1.12)$$

## 2.4 Metode Perhitungan Aktuaria

Metode perhitungan aktuaria yang digunakan pada penelitian ini, adalah metode *projected unit credit* dan metode *individual level premium*

### 2.4.1 Metode Projected Unit Credit

Metode *projected unit credit* adalah membagi total manfaat pensiun pada usia pensiun normal dengan total masa kerja menjadi satuan unit manfaat pensiun yang kemudian dialokasikan ke setiap tahun selama masa kerja (Bowers, et al. 1997).

Iuran normal dengan metode *projected unit credit*  ${}^{PUC}(NC)_x$  dirumuskan dengan persamaan (1.13) (Aitken, 1994):

$${}^{PUC}(NC)_x = b_x \frac{D_r^{(T)}}{D_x^{(T)}} \ddot{a}_r. \quad (1.13)$$

Rumus tersebut dapat dijelaskan sebagai berikut:

Berdasarkan definisi iuran normal dengan metode *projected unit credit*, maka diperoleh:

$$\begin{aligned} {}^{PUC}(NC)_x &= \frac{1}{(r-e)} {}^r(PVFB)_x \\ &= \frac{1}{(r-e)} B_r \ddot{a}_r v^{r-x} {}_{r-x}p_x^{(T)}. \end{aligned}$$

(Karena  ${}_{r-x}p_x^{(T)} = \frac{l_r^{(T)}}{l_x^{(T)}}$ , )

$$= \frac{1}{(r-e)} B_r \ddot{a}_r \frac{v^r l_r^{(T)}}{v^x l_x^{(T)}}.$$

(Berdasarkan fungsi komutasi)

$$= \frac{1}{(r-e)} B_r \ddot{a}_r \frac{D_r^{(T)}}{D_x^{(T)}}. \quad (1.14)$$

### 2.4.2 Metode Individual Level Premium

Metode *individual level premium* merupakan metode yang termasuk dalam kelompok metode *projected benefit cost*. Metode *individual level premium* adalah nilai sekarang pada tanggal perhitungan aktuaria dari total manfaat pensiun dialokasikan secara merata pada setiap tahun masa kerja, yaitu sejak tanggal perhitungan aktuaria hingga usia pensiun normal. Metode tersebut menggunakan asumsi kenaikan gaji dengan besar anuitas yang telah ditentukan berdasarkan masa kerja yang telah lewat dan masa kerja di waktu yang akan datang (Jonatan, 2006).

Besar iuran normal dengan menggunakan metode *individual level premium* dirumuskan sebagai berikut (Aitken, 1994):

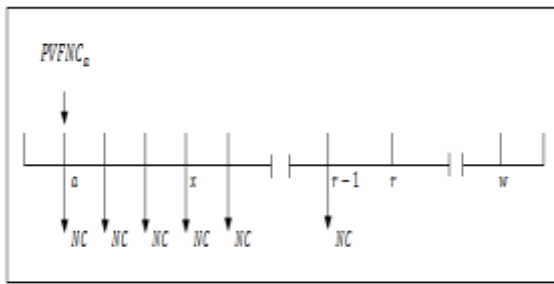
$$NC \left( \frac{N_e^{(T)} - N_r^{(T)}}{D_e^{(T)}} \right) = B_r \ddot{a}_r \frac{D_r^{(T)}}{D_e^{(T)}}. \quad (1.15)$$

Pada dasarnya, iuran normal yang dibayarkan peserta secara berkala (PVFNC) pada saat peserta berusia  $e$  tahun sampai usia  $r$  tahun digunakan untuk membayarkan manfaat (PVFB) yang akan diterima peserta pada saat pensiun. Sehingga, nilai sekarang dari iuran normal saat peserta berusia  $e$  tahun  ${}^r(PVFN C)_e$  nilainya akan sama dengan nilai sekarang dari manfaat pensiun saat peserta berusia  $e$  tahun  ${}^r(PVFB)_e$ . Oleh karena itu, diperoleh persamaan:

$${}^r(PVFB)_e = {}^r(PVFN C)_e. \quad (1.16)$$

*Present value of future normal cost* (PVFNC) adalah nilai sekarang dari iuran normal yang dibayarkan secara berkala oleh peserta dimulai dari peserta berusia  $e$  tahun sampai memasuki usia pensiun berusia  $r - 1$  tahun, yang dinotasikan dengan  ${}^r(PVFN C)_e$ . Besar pembayaran berkala iuran normal yang dilakukan setiap awal tahun sebesar  $NC$  dimulai dari peserta masuk program pensiun (usia  $a$  tahun) sampai memasuki usia pensiun (usia  $r -$

1 tahun), dapat dijelaskan dengan skema pembayaran seperti tampak pada Gambar 2:



Gambar 2 Skema Pembayaran Iuran Normal Selama Masa Kerja

Berdasarkan skema pembayaran pada Gambar 2, pembayaran berkala iuran normal selama masa kerja pada peserta dari usia  $a$  tahun sampai berusia  $r - 1$  tahun adalah:

$$1 + vp_a + v^2 {}_2p_a + \dots + v^{r-1-a} {}_{r-1-a}p_a = \ddot{a}_{a:r-a}$$

Oleh karena itu, nilai sekarang iuran normal saat peserta berusia  $a$  tahun yang dinotasikan dengan  ${}^r(PVFNC)_a$  adalah

$${}^r(PVFNC)_a = NC(\ddot{a}_{a:r-a})$$

Diasumsikan bahwa usia peserta saat masuk program pensiun sama dengan usia peserta saat masuk kerja ( $a = e$ ). Maka dari itu, diperoleh persamaan:

$${}^r(PVFNC)_e = NC(\ddot{a}_{e:r-e}) = NC\left(\frac{N_e^{(T)} - N_r^{(T)}}{D_e^{(T)}}\right)$$

Berdasarkan hal tersebut, maka persamaan (1.16) menjadi:

$$NC\left(\frac{N_e^{(T)} - N_r^{(T)}}{D_e^{(T)}}\right) = B_r \ddot{a}_r \frac{D_r^{(T)}}{D_e^{(T)}} \implies N = B_r \ddot{a}_r \frac{D_r^{(T)}}{N_e^{(T)} - N_r^{(T)}}$$

Oleh karena itu, perhitungan iuran normal dengan metode *individual level premium* saat peserta berusia  $x$  tahun  ${}^{ILP}(NC)_x$  dapat dirumuskan dengan:

$${}^{ILP}(NC)_x = B_r \ddot{a}_r \frac{D_r^{(T)}}{N_e^{(T)} - N_r^{(T)}} \tag{1.17}$$

### 2.5 Nilai Akhir Pembiayaan Iuran Normal

Pada program dana pensiun, nilai akhir pembiayaan iuran normal digunakan untuk mengetahui total pembiayaan iuran normal yang dikeluarkan peserta selama mengikuti program dana pensiun sampai memasuki usia pensiun. Jika seorang peserta masuk program dana pensiun pada usia  $e$  tahun dan masih hidup saat memasuki usia pensiun (berusia  $r$  tahun), maka nilai akhir total iuran normal yang dibayar peserta saat berusia  $r$  tahun yang dinotasikan dengan  ${}^r(NA)_e$  adalah (Oktiani, 2013):

$${}^r(NA)_e = \sum_{x=e}^{r-1} (NC)_x (1+i)^{r-x}$$

### 3. Metode Penelitian

Jenis data yang digunakan dalam penelitian ini adalah data kuantitatif yaitu gaji pegawai Badan Kepegawaian Daerah Provinsi Bali tahun 2012, berupa data peserta program pensiun yang terdiri dari gaji pegawai dan proporsi dari gaji yang dipersiapkan untuk manfaat pensiun. Pada penelitian ini digunakan data pegawai yang diambil secara acak untuk perhitungan dana pensiun.

Pengolahan data dilakukan dengan menggunakan *software Microsoft Excel 2007* sebagai alat bantu perhitungan. Langkah-langkah analisis yang dilakukan untuk memperoleh hasil perhitungan dana pensiun adalah:

1. Penyusunan Tabel Penyusutan Jamak (*Multiple Decrement Table*).
2. Menyusun tabel perhitungan, berdasarkan tabel penyusutan jamak dan Tabel Mortalitas Indonesia 1993 dengan asumsi tingkat suku bunga sebesar 10% per tahun.
3. Menghitung besar manfaat pensiun masing-masing peserta berdasarkan skala gaji.
4. Menghitung nilai sekarang manfaat pensiun  ${}^r(PVFB)_x$ , kemudian membandingkan hasil perhitungan berdasarkan asumsi tiga skala gaji dari sudut pandang perusahaan penyelenggara program dana pensiun.
5. Menghitung besar iuran normal dana pensiun yang dibayar peserta program dana pensiun dengan menggunakan metode

*projected unit credit* dan *individual level premium*, kemudian membandingkan hasil perhitungan antara kedua metode.

6. Menghitung besar nilai akhir pembiayaan iuran normal berdasarkan metode *projected unit credit* dan *individual level premium*.
7. Membandingkan hasil perhitungan nilai akhir pembiayaan iuran normal dana pensiun setiap tahun antara metode *projected unit credit* dan *individual level premium* dari sudut pandang peserta program dana pensiun.

## 4. Hasil dan Pembahasan

### 4.1 Pembuatan Tabel *Multiple Decrement*

Tabel penyusutan jamak (*multiple decrement table*) menunjukkan peluang situasi penurunan pegawai pada perusahaan yang disebabkan oleh pensiun dipercepat, kematian, pensiun karena tidak bisa bekerja, atau pensiun pada saat usia pensiun.

Misalkan banyaknya pegawai yang masih aktif bekerja saat berusia 30 tahun sebesar  $l_{30}^{(T)}$ , sedangkan banyaknya pegawai yang keluar, meninggal atau cacat di antara usia 30 dan 31 tahun, secara berturut-turut sebesar  $w_{30}$ ,  $d_{30}^T$ , dan  $i_{30}$ , maka peluang pegawai pensiun di antara usia 30 dan 31 tahun pada perusahaan yang disebabkan oleh pensiun dipercepat (keluar), kematian, atau pensiun karena tidak bisa bekerja (cacat), secara berturut-turut dinotasikan dengan  $q_{30}^{(w)}$ ,  $q_{30}^{(d)}$ , atau  $q_{30}^{(i)}$ . Berdasarkan data pada *service table* (Futami, 1993b), diperoleh

$$q_{30}^{(w)} = \frac{w_{30}}{l_{30}^{(T)}} = \frac{1.510}{38.728} = 0,038990,$$

$$q_{30}^{(d)} = \frac{d_{30}^{(T)}}{l_{30}^{(T)}} = \frac{29}{38.728} = 0,000749, \text{ atau}$$

$$q_{30}^{(i)} = \frac{i_{30}}{l_{30}^{(T)}} = \frac{9}{38.728} = 0,000232.$$

Dari sini diperoleh, peluang pegawai yang masih aktif bekerja di antara usia 30 dan 31 tahun adalah

$$\begin{aligned} p_{30}^{(T)} &= 1 - q_{30}^{(d)} - q_{30}^{(w)} - q_{30}^{(i)} \\ &= 1 - 0,000749 - 0,038990 - \\ &\quad 0,000232 \end{aligned}$$

$$= 0,960029.$$

Dengan cara yang sama,  $q_x^{(w)}$ ,  $q_x^{(d)}$ ,  $q_x^{(i)}$ , dan  $p_x^{(T)}$  dapat dihitung untuk  $x = 18$  sampai  $x = 60$ , yang semuanya bisa disusun dalam bentuk tabel.

### 4.2 Pembuatan Tabel Perhitungan dengan $i = 10\%$

Tabel perhitungan dibuat untuk memudahkan dalam perhitungan dana pensiun. Tabel perhitungan disusun berdasarkan tabel *multiple decrement* dengan asumsi tingkat bunga sebesar 10% per tahun.

Misalkan pada saat pegawai berusia 30 tahun, nilai komutasi untuk  $v^x$ ,  $v^{r-x}$ , dan  $(1+i)^{r-x}$  secara berturut-turut adalah,

$$v^{30} = \frac{1}{(1+10\%)^{30}} = 0,057309$$

$$v^{56-30} = \frac{1}{(1+10\%)^{26}} = 0,083905$$

$$(1+10\%)^{56-30} = (1,1)^{26} = 11,918177.$$

Perhitungan untuk peluang pegawai berusia 30 tahun yang masih aktif bekerja sampai berusia 31 tahun adalah

$$p_{30}^{(T)} = \frac{l_{31}^{(T)}}{l_{30}^{(T)}} = \frac{37.180}{38.728} = 0.960029.$$

Perhitungan untuk peluang pegawai berusia 30 tahun yang masih aktif bekerja sampai berusia  $r$  tahun adalah:

$${}_{26}p_{30}^{(T)} = \frac{l_{26+30}^{(T)}}{l_{30}^{(T)}} = \frac{26.371}{38.728} = 0.680929.$$

Perhitungan untuk nilai komutasinya adalah:

$$\begin{aligned} D_{30}^{(T)} &= 38.728 (1+0,1)^{-30} \\ &= 2.219,446. \end{aligned}$$

Dari sini diperoleh:

$$\begin{aligned} N_{30}^{(T)} &= D_{30}^{(T)} + \dots + D_{56}^{(T)} \\ &= l_{30}^{(T)} v^{30} + \dots + l_{56}^{(T)} v^{56} \\ &= 38.728 (1,1)^{-30} + \dots + 26.371(1,1)^{-56} \\ &= 2.219,446 + \dots + 126.805 \end{aligned}$$

$$= 19.034,619.$$

Dengan cara yang sama,  $v^x, v^{r-x}, (1+i)^{r-x}, p_x^{(T)}, r-x p_x^{(T)}, D_x^{(T)}$ , dan  $N_x^{(T)}$  dapat dihitung untuk  $x = 25$  sampai  $x = 56$ , yang semuanya bisa disusun dalam bentuk tabel.

### 4.3 Contoh Kasus Penerapan

Berikut ini, diberikan suatu contoh kasus penerapan. Misalkan, seorang pegawai negeri dengan golongan II/b, berjenis kelamin laki-laki, mulai menjadi peserta pada usia 25 tahun ( $e = 25$ ), dan mulai terhitung pensiun pada tanggal 1 Januari 2035 dengan usia 56 tahun ( $r = 56$ ). Gaji pokok pada tahun pertama ( $s_e$ ) diterima dalam setahun sebesar Rp17.008.800,

Perhitungan pembiayaan pensiun pada saat peserta berusia 34 tahun ( $x = 34$ ) adalah:

#### 1. Perhitungan manfaat pensiun

Perhitungan manfaat pensiun berdasarkan asumsi 3 skala gaji adalah

##### a) Asumsi gaji terakhir

$$\begin{aligned} s_{55} &= (1 + 5\%)^{56-1-25} s_{25} \\ &= (1,05)^{30} 17.008.800 \\ &= 73.511.053. \end{aligned}$$

Manfaat pensiun yang akan diterima peserta pada saat pensiun adalah

$$\begin{aligned} B_{56}^{(1)} &= 2,25\% (56 - 25) s_{55} \\ &= (0,6975) 73.511.053 \\ &= 51.273.959. \end{aligned}$$

Jadi besar total manfaat pensiun yang akan diterima peserta adalah Rp51.273.959,-

##### b) Asumsi rata-rata gaji selama 5 tahun terakhir

Akan dilakukan perhitungan untuk  $\ddot{a}_{5|0,05}$  terlebih dahulu yaitu:

$$\begin{aligned} \ddot{a}_{5|0,05} &= (1 + 0,05)^{-4} + \dots + (1 + 0,05)^0 \\ &= 0,822702 + \dots + 1 \\ &= 4,545951. \end{aligned}$$

Perhitungan skala gaji dengan asumsi rata-rata gaji selama 5 tahun terakhir adalah:

$$FAS = \frac{1}{5} (1 + 5\%)^{56-25-1} \ddot{a}_{5|0,05} s_{25}$$

$$\begin{aligned} &= \frac{1}{5} (1,05)^{30} (4,545951) 17.008.800 \\ &= 66.835.522. \end{aligned}$$

Manfaat pensiun yang akan diterima peserta pada saat pensiun adalah

$$\begin{aligned} B_{56}^{(2)} &= 2,25\% 66.835.522 (56 - 25) \\ &= 46.617.777. \end{aligned}$$

Jadi besar total manfaat pensiun yang akan diterima peserta adalah Rp46.617.777,-

##### c) Asumsi rata-rata gaji selama bekerja:

Pada kasus ini diketahui gaji pada tahun pertama yaitu  $s_e$ , maka:

$$\begin{aligned} &\frac{1}{r-e} [s_e + \dots + s_{r-1}] \\ &= \frac{1}{r-e} [s_e + \dots + s_e (1+s)^{r-1-e}] \\ &= \frac{1}{r-e} s_e [1 + \dots + (1+s)^{r-1-e}] \\ &= \frac{1}{r-e} s_e s_{r-e|s}. \end{aligned}$$

Perhitungan untuk  $s_{56-25|0,05}$  adalah:

$$\begin{aligned} s_{31|0,05} &= 1 + \dots + (1+s)^{56-1-25} \\ &= 1 + \dots + 4,321942 \\ &= 70,760790. \end{aligned}$$

Perhitungan skala gaji dengan asumsi rata-rata gaji selama bekerja adalah:

$$\begin{aligned} &\frac{1}{56-25} s_{25} s_{31|0,05} \\ &= \frac{1}{31} 17.008.800 (70,760790) \\ &= 38.824.391. \end{aligned}$$

Manfaat pensiun yang akan diterima peserta pada saat pensiun adalah

$$\begin{aligned} B_{56}^{(3)} &= 2,25\% 38.824.391 (56 - 25) \\ &= 27.080.013. \end{aligned}$$

Jadi besar total manfaat pensiun yang akan diterima peserta adalah Rp27.080.013,-

2. Perhitungan nilai sekarang manfaat pensiun  
(*Present Value of Future Benefit*)

a) Asumsi gaji terakhir

$$\begin{aligned} {}^{56}(PVFB)_{34}^1 &= B_{56}^{(1)} \ddot{a}_{56} v^{56-34} {}_{56-34}p_{34}^{(T)} \\ &= B_{56}^{(1)} \frac{N_{56}}{D_{56}} v^{56-34} {}_{56-34}p_{34}^{(T)} \\ &= 40.652.747. \end{aligned}$$

Jadi nilai sekarang total manfaat pensiun pada usia 34 tahun sebesar Rp40.652.747,-

b) Asumsi rata-rata gaji selama 5 tahun terakhir

$$\begin{aligned} {}^{56}(PVFB)_{34}^2 &= B_{56}^{(2)} \ddot{a}_{56} v^{56-34} {}_{56-34}p_{34}^{(T)} \\ &= B_{56}^{(2)} \frac{N_{56}}{D_{56}} v^{56-34} {}_{56-34}p_{34}^{(T)} \\ &= 36.961.076. \end{aligned}$$

Jadi nilai sekarang total manfaat pensiun pada usia 34 tahun sebesar Rp36.961.076,-

c) Asumsi rata-rata gaji selama bekerja

$$\begin{aligned} {}^{56}(PVFB)_{34}^3 &= B_{56}^{(3)} \ddot{a}_{56} v^{56-34} {}_{56-34}p_{34}^{(T)} \\ &= B_{56}^{(3)} \frac{N_{56}}{D_{56}} v^{56-34} {}_{56-34}p_{34}^{(T)} \\ &= 21.470.488. \end{aligned}$$

Jadi nilai sekarang total manfaat pensiun pada usia 34 tahun sebesar Rp21.470.488,-

3. Perhitungan iuran normal dengan metode aktuarial.

a) Metode *projected unit credit*

Perhitungan iuran normal dengan metode *projected unit credit* adalah:

$$\begin{aligned} {}^{PUC}NC_{34} &= \frac{B_{56}^{(3)}}{(r-e)} \frac{D_{56}^{(T)}}{D_{34}^{(T)}} \ddot{a}_{56} \\ &= \frac{B_{56}^{(3)}}{(r-e)} \frac{l_{56}^{(T)} v^{56} N_{56}}{l_{34}^{(T)} v^{34} D_{56}} \\ &= \frac{27.080.013}{(56-25)} \frac{126,805}{1.323,722} (8,276628) \\ &= 692.596. \end{aligned}$$

Jadi iuran normal yang harus dibayar peserta saat berusia 34 tahun adalah sebesar Rp692.596,-

b) Metode *individual level premium*

Perhitungan iuran normal dengan metode *individual level premium* adalah:

$$\begin{aligned} {}^{ILP}NC_{34} &= \frac{B_{56}^{(3)} D_{56}^{(T)}}{N_{25}^{(T)} - N_{56}^{(T)}} \ddot{a}_{56} \\ &= \frac{B_{56}^{(3)} (l_{56}^{(T)} v^{56})}{D_{25}^{(T)} + D_{26}^{(T)} + \dots + D_{55}^{(T)}} \frac{N_{56}}{D_{56}} \\ &= \frac{27.080.013 (126,805)}{36.752,72962} (8,276628) \\ &= 773.302. \end{aligned}$$

Jadi pembayaran iuran normal yang dilakukan tiap tahun oleh peserta saat berusia 34 tahun adalah sebesar Rp773.302,-

Dengan cara yang sama, perhitungan nilai sekarang total manfaat pensiun, pembiayaan iuran normal setiap tahun, dan nilai akhir pembiayaan iuran normal dapat dihitung untuk  $x = 25$  sampai  $x = 56$ .

4. Perhitungan nilai akhir pembiayaan iuran normal dengan metode aktuarial adalah:

a) Metode *projected unit credit*

Perhitungan nilai akhir pembiayaan iuran normal dengan metode *projected unit credit* adalah:

$$\begin{aligned} {}^{PUC}NA &= \sum_{x=25}^{55} {}^{PUC}(NC)_x (1+i)^{56-x} \\ &= {}^{PUC}(NC)_{25} (1,1)^{31} + \dots \\ &\quad + {}^{PUC}(NC)_{56} (1,1)^1 \\ &= (3.659.706) + \dots + (7.953.042) \\ &= 185.155.205. \end{aligned}$$

Jadi, total nilai akhir pembiayaan iuran normal dengan metode *projected unit credit* adalah sebesar Rp192.385.244,-

b) Metode *individual level premium*

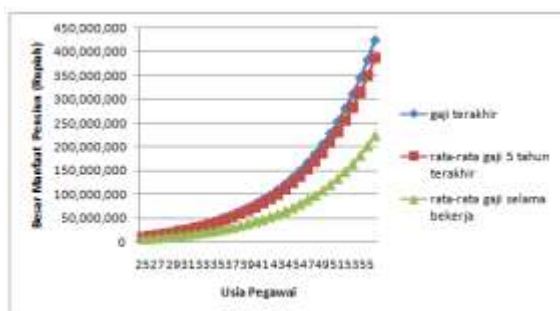
Perhitungan nilai akhir pembiayaan iuran normal dengan metode *individual level premium* adalah:

$$\begin{aligned} ILP NA &= \sum_{x=25}^{55} ILP(NC)_x(1+i)^{56-x} \\ &= ILP(NC)_{25}(1+0,1)^{56-25} + \dots \\ &\quad + ILP(NC)_{56}(1+0,1)^{56-55} \\ &= (773.302)(19,194342) + \dots \\ &\quad + (773.302)(1,1) \\ &= (14.843.022) + \dots + (850.632) \\ &= 154.766.918. \end{aligned}$$

Jadi, total nilai akhir pembiayaan iuran normal dengan metode *individual level premium* adalah sebesar Rp154.766.918,-

#### 4.4 Perbandingan Hasil Perhitungan

Penggunaan asumsi skala gaji untuk menghitung nilai sekarang manfaat pensiun tiap tahunnya disajikan dengan grafik garis seperti pada Gambar 3.



Gambar 3. Grafik nilai sekarang dari manfaat pensiun (*Present Value of Future Benefit*) dengan asumsi 3 skala gaji.

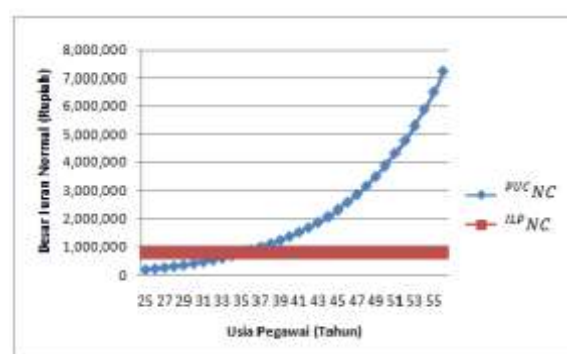
Gambar 3 menunjukkan bahwa penggunaan asumsi gaji terakhir menghasilkan manfaat pensiun yang paling tinggi. Hal ini berarti bahwa penggunaan skala gaji terakhir akan menimbulkan biaya yang paling tinggi dibandingkan dengan penggunaan asumsi gaji yang lain, mengingat gaji pegawai selalu meningkat tiap tahun dan gaji terakhir merupakan gaji tertinggi yang diperoleh pegawai. Selain itu, penggunaan asumsi gaji terakhir dapat merugikan bagi perusahaan karena harus membayar kekurangan

pembiayaan yang terjadi pada awal masa kepesertaan bagi peserta yang memperoleh peningkatan penghasilan tiap tahunnya.

Berdasarkan Gambar 3 juga diperoleh bahwa penggunaan asumsi rata-rata gaji selama 5 tahun terakhir akan menimbulkan biaya dan manfaat pensiun yang lebih kecil dibandingkan dengan penggunaan asumsi gaji terakhir. Hal ini berarti besar manfaat yang diterima peserta sesuai dengan gaji yang diperoleh menjelang masa pensiun. Selain itu, Gambar 3 juga menunjukkan bahwa penggunaan asumsi rata-rata gaji selama bekerja menghasilkan manfaat pensiun paling rendah dibandingkan ketiga asumsi skala gaji yang lain. Akibatnya penggunaan asumsi skala gaji ini menghasilkan manfaat pensiun yang relatif stabil tiap tahun, sehingga perusahaan dapat terhindar dari kewajiban pembiayaan iuran normal akibat peningkatan gaji pada tahun-tahun tertentu.

Berdasarkan hal tersebut, maka asumsi skala gaji yang sesuai bagi perusahaan penyelenggara program pensiun untuk terhindar dari kerugian adalah asumsi rata-rata gaji selama bekerja.

Selanjutnya akan dilakukan perbandingan perhitungan pembiayaan iuran normal dengan metode *projected unit credit* dan *individual level premium* dengan asumsi rata-rata gaji selama bekerja.



Gambar 4. Grafik pembiayaan iuran normal menggunakan metode *projected unit credit* (PUC) dan *individual level premium* (ILP) berdasarkan rata-rata gaji selama bekerja

Grafik garis yang ditunjukkan pada Gambar 4 menunjukkan bahwa pembiayaan iuran normal dengan metode *projected unit*

*credit* terus meningkat setiap tahunnya yang ditunjukkan dengan grafik garis yang berwarna biru sedangkan grafik dengan garis berwarna merah menunjukkan pembiayaan iuran normal dengan metode *individual level premium* yang cenderung tetap selama bekerja sampai memasuki usia pensiun, hal ini disebabkan karena perhitungan pembiayaan iuran normal dengan metode *individual level premium* tidak dipengaruhi oleh usia peserta saat tahun perhitungan aktuarial (saat peserta berusia  $x$  tahun), namun hanya dipengaruhi oleh usia peserta saat memasuki program pensiun (saat peserta berusia  $e$  tahun). Pada awal tahun masuk program pensiun sampai peserta berusia 35 tahun, perhitungan pembiayaan iuran normal dengan metode *individual level premium* lebih tinggi dibandingkan dengan metode *projected unit credit*. Sedangkan perhitungan pembiayaan iuran normal saat peserta berusia 36 tahun sampai memasuki usia pensiun dengan metode *projected unit credit* jauh lebih tinggi dibandingkan dengan perhitungan pembiayaan iuran normal dengan metode *individual level premium* dengan besar manfaat pensiun yang sama untuk kedua metode tersebut.

Berdasarkan hasil perbandingan pembiayaan iuran normal setiap tahun, belum bisa ditentukan metode yang bisa disarankan bagi peserta program dana pensiun. Berikut ini dilakukan perbandingan hasil nilai akhir pembiayaan iuran normal setiap tahun, yang diasumsikan bahwa setiap peserta yang masuk program dana pensiun pada usia  $e$  tahun dan masih hidup saat memasuki usia pensiun (berusia  $r$  tahun). Berdasarkan perhitungan nilai akhir pembiayaan iuran normal, diperoleh total akhir pembiayaan iuran normal dengan metode *individual level premium* sebesar Rp154.766.918,-, sedangkan total nilai akhir pembiayaan iuran normal dengan metode *projected unit credit* sebesar Rp185.155.205,-. Penggunaan metode aktuarial dengan *projected unit credit* menghasilkan nilai akhir pembiayaan iuran normal yang lebih tinggi dibandingkan dengan nilai akhir pembiayaan iuran normal dengan metode *individual level premium*. Oleh

karena itu, perhitungan pembiayaan iuran normal dari sudut pandang peserta asuransi disarankan untuk menggunakan metode *projected unit credit* pada program dana pensiun. Dengan demikian, peserta dapat memaksimalkan manfaat yang diperoleh dalam menginvestasikan dananya dalam bentuk program dana pensiun.

## 5. Kesimpulan

Berdasarkan pembahasan, kesimpulan yang bisa diambil adalah:

1. Penggunaan asumsi rata-rata gaji selama bekerja menghasilkan besar manfaat pensiun yang relatif stabil tiap tahun, sehingga perusahaan dapat terhindar dari kewajiban pembiayaan iuran normal akibat peningkatan gaji pada tahun-tahun tertentu.
2. Perhitungan nilai akhir pembiayaan iuran normal dengan metode *individual level premium* memperoleh hasil bahwa penggunaan metode *individual level premium* lebih baik digunakan dari sudut pandang peserta program dana pensiun dibandingkan dengan metode *projected unit credit*.

## Daftar Pustaka

- Aitken, W. H. 1994. *A Problem Solving Approach to Pension Funding and Valuation*. 2<sup>nd</sup> edition. Winsted : Actex Publications.
- Bowers, Newton. L. *et al.* 1997. *Actuarial Mathematics*. 2<sup>nd</sup> edition. IPC Publishing.
- Futami, T. 1993a. *Matematika Asuransi Jiwa Bagian I*. Herliyanto, Gatot, penerjemah. Tokyo: Oriental Life Insurance Cultural Development Center. Terjemahan dari : *Seimei Hoken Sugaku, Jokan* ("92 Revision).
- \_\_\_\_\_. 1993b. *Matematika Asuransi Jiwa Bagian II*. Herliyanto, Gatot, penerjemah. Tokyo: Oriental Life Insurance Cultural Development Center. Terjemahan dari : *Seimei Hoken Sugaku, Jokan* ("92 Revision).



- Hapsari, Ayu. 2012. Penggunaan Metode Projected Unit Credit dan Entry Age Normal dalam Pembiayaan Pensiun. *Jurnal Gaussian*, Volume 1, Nomor 1, Halaman 47-54.
- Jonatan, Ponso. 2006. Alternatif Pendanaan untuk Imbal Pasca Kerja Berdasarkan UU No 13 Tahun 2003 tentang Ketenagakerjaan. *Tesis*. Jurusan Manajemen FE UI. Jakarta: Universitas Indonesia.
- Oktiani, Irma. 2013. Perhitungan Aktuarial untuk Manfaat Pensiun Normal Menggunakan Metode Projected Unit Credit dan Entry Age Normal. *Skripsi*. Bogor: Institut Pertanian Bogor
- Pratiwi, Anggi Noor. 2008. Penentuan Dana Pensiun dan Perhitungan Premi dengan Metode Accrued Benefit Cost pada Asuransi Dana Pensiun. *Skripsi*. Yogyakarta : Universitas Negeri Yogyakarta.
- Taspen. 2013. *Program Pensiun*. diakses melalui [web: http://www.taspen.com/produk/pensiun#navigasi](http://www.taspen.com/produk/pensiun#navigasi). pada tanggal 2 Maret 2013

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

I PUTU JERYANA<sup>1</sup>, I PUTU EKA NILA KENCANA<sup>2</sup>, G.K. GANDHIADI<sup>3</sup>

<sup>1,2,3</sup>Jurusan Matematika FMIPA Universitas Udayana, Bukit Jimbaran-Bali  
e-mail: <sup>1</sup>jerry\_crock30@yahoo.com, <sup>2</sup>i.putu.enk@gmail.com, <sup>3</sup>gandhiadigk@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].

<sup>1</sup> Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

<sup>2</sup> 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- $t$ .

$\varepsilon_t$  saling bebas dan menyebar normal dengan mean 0 dan variansi  $\sigma^2$

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  $\sigma^2$ .

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  $y_i^*$

$\beta$  : 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  $\sigma^2$

$i$  : 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  $\theta$  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  $\beta$  dan  $\sigma^2$  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  $\sigma^2$  dan  $\beta$

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

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

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

### 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  $\beta_j$  tidak signifikan secara statistik)

$H_1 : \beta_j \neq 0$  (koefisien  $\beta_j$  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  $\beta_j$

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

$H_0$  ditolak jika nilai statistik uji  $|W| < -Z_{\alpha/2}$  atau jika  $p$ -value  $< \alpha$  yang berarti  $\beta_j$  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<sup>2</sup>

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<sup>2</sup>* yang dapat melihat kemampuan model regresi Tobit dalam menerangkan variasi perubahan variabel terikat. Dalam program *E-views* nilai *Pseudo-R<sup>2</sup>* berbentuk McFadden  $R^2$ .

Berikut persamaan *Pseudo-R<sup>2</sup>* 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

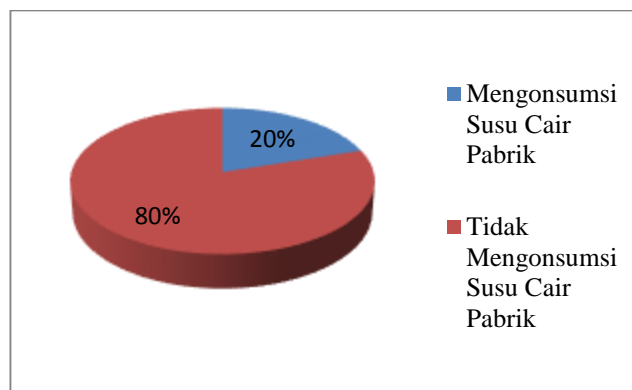
Pada penelitian ini, pemodelan akan dilakukan dengan menggunakan regresi Tobit. Proses analisis data pada penelitian ini menggunakan *software SPSS 20* dan *E-views 5.1*. Langkah-langkah analisis untuk memperoleh model regresi tobit adalah:

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-R<sup>2</sup>* 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.

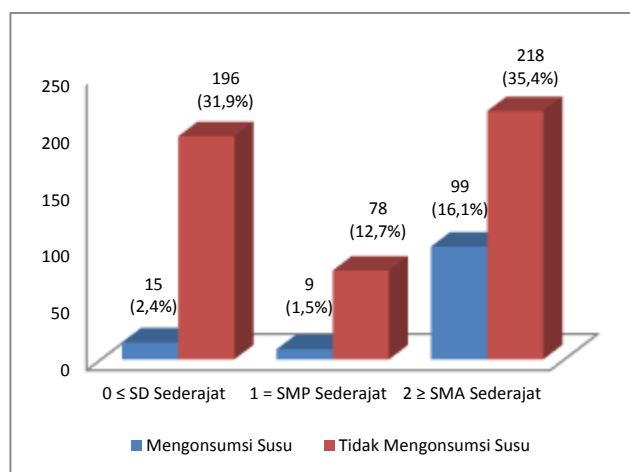
Tabel 3.1 Deskripsi Variabel Bebas yang Memengaruhi Rumah Tangga Mengonsumsi Susu Cair Pabrik (Data Kontinu)

| 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  $\beta_i$  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  $\beta_i$  yang berpengaruh signifikan terhadap variabel tak bebas.

Tabel 3.3 menunjukkan nilai  $\chi_{hitung}^2$  (*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  $\beta_j$  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  $\alpha$  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  $\chi_{hitung}^2$  (*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 <i>Log Likelihood</i>     | 105,109  | Prob. <i>Chi-Square</i> (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  $\beta_1$  dan  $\beta_2$ , 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 | <b>0,167965</b> | <b>6,751572</b> |
|    | $x_1$     | 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*. 3<sup>rd</sup> Edition. McGraw-Hill International Edition, Economic Series.
- [2] Greene, W. H. 2003. *Econometrics Analysis*, 5<sup>th</sup> 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*. 3<sup>rd</sup> 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.