

PENAKSIR REGRESI CUM RASIO UNTUK RATA-RATA POPULASI DENGAN MENGGUNAKAN KOEFISIEN KURTOSIS DAN KOEFISIEN SKEWNESS

Yustina Wulan Sari^{1*}, Arisman Adnan², Haposan Sirait²

¹Mahasiswa Program S1 Matematika

²Dosen Jurusan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Riau
Kampus Binawidya Pekanbaru, 28293, Indonesia

*yustina.wulan@unri.ac.id

ABSTRACT

This paper discusses the estimators of regression cum-ratio proposed by Yan and Tian for the population mean. These three estimators use the auxiliary variables of the coefficients of kurtosis and skewness in simple random sampling. All estimators are biased estimators. The most efficient estimator is the estimator that has the smallest mean square error (*MSE*) obtained by comparing their *MSEs*. An example is given at the end of discussion.

Keywords: *regression cum-ratio estimator, coefficient of kurtosis, coefficient of skewness, biased and mean square error*

ABSTRAK

Tulisan ini membahas tiga penaksir regresi cum-rasio yang diajukan oleh Yan dan Tian untuk rata-rata populasi. Ketiga penaksir menggunakan informasi tambahan koefisien kurtosis dan skewness pada sampling acak sederhana. Masing-masing penaksir merupakan penaksir bias. Penaksir yang paling efisien merupakan penaksir yang memiliki *Mean Square Error (MSE)* terkecil yang diperoleh dengan membandingkan *MSE* dari masing-masing penaksir. Contoh numerik diberikan pada akhir pembahasan.

Kata kunci: penaksir regresi cum-rasio, koefisien kurtosis, koefisien skewness, bias dan mean square error

1. PENDAHULUAN

Penaksir regresi cum-rasio merupakan suatu metode yang digunakan untuk meningkatkan ketelitian suatu penaksir, dengan mengambil manfaat hubungan antara y_i dan x_i dimana y_i adalah unit dari populasi berkarakter Y dan x_i adalah unit dari

populasi berkarakter X . Variabel x_i merupakan suatu variabel pendukung yang berkorelasi positif dengan variabel y_i dan variabel x_i telah diteliti sebelumnya sehingga variabel x_i dapat digunakan sebagai variabel bantu untuk menaksir variabel y_i [1].

Bentuk umum penaksir rasio untuk rata-rata populasi \bar{Y} dari variabel yang diteliti Y dirumuskan dengan

$$\hat{\bar{Y}}_R = \frac{\bar{y}}{\bar{x}} \bar{X} = \hat{R} \bar{X}$$

dengan \bar{y} adalah rata-rata sampel dari populasi Y , \bar{x} adalah rata-rata sampel dari populasi X dan \bar{X} adalah rata-rata populasi X .

Dalam artikel ini dibahas tiga modifikasi penaksir regresi cum-rasio untuk rata-rata populasi pada sampling acak sederhana dengan menggunakan koefisien kurtosis dan skewness yang diajukan oleh Yan dan Tian [5], yaitu

$$\bar{Y}_{r,1} = (\bar{y} - b(\bar{x} - \bar{X})) \left(\frac{\bar{X} + \beta_1}{\bar{x} + \beta_1} \right) \quad (1)$$

$$\bar{Y}_{r,2} = (\bar{y} - b(\bar{x} - \bar{X})) \left(\frac{\beta_1 \bar{X} + \beta_2}{\beta_1 \bar{x} + \beta_2} \right) \quad (2)$$

$$\bar{Y}_{r,3} = (\bar{y} - b(\bar{x} - \bar{X})) \left(\frac{\beta_2 \bar{X} + \beta_1}{\beta_2 \bar{x} + \beta_1} \right) \quad (3)$$

dengan β_1 adalah koefisien skewness dan β_2 adalah koefisien kurtosis.

Ketiga modifikasi penaksir regresi cum-rasio untuk rata-rata populasi tersebut merupakan penaksir bias, kemudian ditentukan MSE . Berdasarkan ide dari Yan dan Tian [5], penulis membandingkan MSE dari masing-masing penaksir untuk memperoleh penaksir regresi cum-rasio yang efisien. Penaksir yang memiliki nilai MSE terkecil merupakan penaksir yang efisien.

Sampling Acak Sederhana

Sampling acak sederhana adalah sebuah metode yang digunakan untuk mengambil n unit sampel dari N unit populasi sehingga setiap unit populasi memiliki kesempatan yang sama untuk dipilih menjadi unit sampel. Dalam hal ini pengambilan sampel dilakukan tanpa pengembalian agar hasil yang diperoleh menjadi representatif [1].

Pada pengambilan sampel tanpa pengembalian probabilitas terpilihnya n dari N populasi terpilih menjadi unit sampel pada pengambilan pertama adalah n/N , probabilitas pada pengambilan kedua adalah $(n-1)/(N-1)$, sampai probabilitas pada

pengambilan ke- n yaitu $1/(N-n+1)$, sehingga peluang seluruh n unit-unit tertentu yang terpilih dalam n pengambilan adalah $({}_N C_n)^{-1}$.

Teorema 2.1 [1 : h. 27] Apabila sampel berukuran n diambil dari populasi berukuran N yang berkarakter Y , dengan sampling acak sederhana tanpa pengembalian maka variansi rata-rata sampel \bar{y} dinotasikan dengan $V(\bar{y})$ yaitu

$$V(\bar{y}) = \frac{S_y^2}{n} \frac{N-n}{N} = \frac{(1-f)}{n} S_y^2$$

dengan $f = n/N$ adalah fraksi penarikan sampel dan $S_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N-1}$ adalah variansi y_i pada populasi berkarakter Y .

Bukti: Bukti dari teorema ini dapat dilihat pada [1].

Teorema 2.2 [1 : h. 29] Jika y_i, x_i adalah sebuah pasangan yang bervariasi dalam unit dalam populasi dan \bar{y}, \bar{x} adalah rata-rata dari sampel acak sederhana berukuran n , maka kovariansi adalah

$$Cov(\bar{y}, \bar{x}) = \frac{1-f}{n} \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{Y})(x_i - \bar{X}) = \frac{1-f}{n} \rho S_y S_x.$$

Bukti: Bukti dari teorema ini dapat dilihat pada [1].

Untuk menentukan MSE dari penaksir dalam bentuk dua variabel digunakan suatu pendekatan dengan menggunakan deret Taylor dua variabel.

Deret Taylor untuk dua variabel [4: h.47] Misalkan $f(x, y)$ adalah suatu fungsi dua variabel dan $f, f', f'', \dots, f^{(n)}$ adalah kontinu pada I dan $f^{(n+1)}$ ada pada I untuk $(x_0, y_0) \in I$. Jika $(x, y) \in I$, maka

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + \left((x-x_0) \frac{\partial}{\partial x} + (y-y_0) \frac{\partial}{\partial y} \right) f(x_0, y_0) + \dots \\ &+ \frac{1}{n!} \left((x-x_0) \frac{\partial}{\partial x} + (y-y_0) \frac{\partial}{\partial y} \right)^{(n)} f(x_0, y_0) \\ &+ \frac{1}{(n+1)!} \left((x-x_0) \frac{\partial}{\partial x} + (y-y_0) \frac{\partial}{\partial y} \right)^{(n+1)} \\ &f(x_0 + \theta(x-x_0), y_0 + \theta(y-y_0)) \end{aligned} \quad (4)$$

dengan $0 < \theta < 1$.

dengan memisalkan $x_0 = \bar{X}$, $x = \bar{X}$, $y_0 = \bar{Y}$, $y = \bar{Y}$, dan mengabaikan pangkat-pangkat yang lebih besar dari satu, maka dari persamaan (4) diperoleh nilai pendekatan unuk mencari *MSE* yaitu

$$f(\bar{x}, \bar{y}) - f(\bar{X}, \bar{Y}) \approx \left((\bar{x} - \bar{X}) \frac{\partial f(\bar{x}, \bar{y})}{\partial x} \Big|_{\bar{x}, \bar{y}} + (\bar{y} - \bar{Y}) \frac{\partial f(\bar{x}, \bar{y})}{\partial y} \Big|_{\bar{x}, \bar{y}} \right) \quad (5)$$

2. PENAKSIR REGRESI UNTUK RATA-RATA POPULASI

Bentuk umum model regresi linear sederhana dalah

$$\bar{Y} = \alpha + \beta X + e, \quad (6)$$

dengan Y adalah variabel tak bebas, X adalah variabel bebas, α dan β adalah parameter (koefisien regresi), e adalah kesalahan pengamatan [3].

Metode yang digunakan untuk mendapatkan parameter α dan β adalah metode kuadrat terkecil. Misalkan x_i dan y_i adalah n pasangan data pengamatan $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, dengan demikian persamaan (6) dapat ditulis

$$y_i = \alpha + \beta x_i + e_i, \quad i = 1, 2, \dots, n. \quad (7)$$

Diasumsikan $E(e_i) = 0$ sehingga $E(y_i) = \alpha + \beta x_i$.

Dari persamaan (7) maka jumlah kuadrat kesalahan pengamatan data terhadap garis regresi ditulis

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2. \quad (8)$$

Dengan mengganti α dan β pada persamaan (8) dengan masing-masing taksirannya adalah a dan b , sehingga diperoleh nilai taksiran sebagai berikut

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2},$$

$$b = \frac{S_{xy}}{S_x^2}$$

dan

$$a = \bar{y} - b\bar{x} \quad (9)$$

Ketika garis regresi linear melalui titik pangkal $a = 0$, maka persamaan (9) menjadi

$$\bar{y} = b\bar{x}, \quad (10)$$

b yang diperoleh dari sampel digunakan juga untuk populasi sehingga rata-rata populasi dinotasikan dengan \hat{Y}

$$\hat{Y} = b\bar{X}. \quad (11)$$

Dari pengurangan persamaan (11) dengan persamaan (10) secara aljabar, diperoleh

$$\hat{Y} = \bar{y} + b(\bar{X} - \bar{x}).$$

\hat{Y} disebut penaksir regresi linier untuk rata-rata populasi yang dinotasikan dengan \hat{Y}_{LR} sehingga

$$\hat{Y}_{LR} = \bar{y} + b(\bar{X} - \bar{x}).$$

3. BIAS DAN *MSE* PENAKSIR RASIO UNTUK RATA-RATA POPULASI

Bias dan *MSE* penaksir regresi cum-rasio yang diajukan oleh Yan dan Tian [5] untuk rata-rata populasi pada sampling acak sederhana dari masing-masing penaksir sebagai berikut

Bias dan *MSE* dari persamaan (1) diperoleh

$$B(\bar{Y}_{r1}) \approx \frac{1-f}{n} R_{r1}^2 \frac{S_x^2}{Y}$$

$$MSE(\bar{Y}_{r1}) \approx \frac{1-f}{n} (R_{r1}^2 S_x^2 + S_y^2 (1 - \rho^2))$$

dengan ρ adalah koefisien korelasi dan $R_{r1} = \frac{\bar{Y}}{\bar{X} + \beta_1}$.

Bias dan *MSE* dari persamaan (2) diperoleh

$$B(\bar{Y}_{r2}) \approx \frac{1-f}{n} R_{r2}^2 \beta_1^2 \frac{S_x^2}{Y}$$

$$MSE(\bar{Y}_{r2}) \approx \frac{1-f}{n} (R_{r2}^2 \beta_1^2 S_x^2 + S_y^2 (1 - \rho^2))$$

dengan $R_{r2} = \frac{\bar{Y}}{\beta_1 \bar{X} + \beta_2}$

Bias dan MSE dari persamaan (3) diperoleh

$$B(\bar{Y}_{r3}) \approx \frac{1-f}{n} R_{r3}^2 \beta_2^2 \frac{S_x^2}{\bar{Y}}$$

$$MSE(\bar{Y}_{r3}) = \frac{1-f}{n} (R_{r3}^2 \beta_2^2 S_x^2 + S_y^2 (1-\rho^2))$$

dengan $R_{r3} = \frac{\bar{Y}}{\beta_2 \bar{X} + \beta_1}$

4. PENAKSIR REGRESI CUM-RASIO YANG EFISIEN

Untuk menentukan penaksir yang efisien dari penaksir yang bias, dapat ditentukan dengan cara membandingkan MSE dari masing-masing penaksir tersebut.

1. Perbandingan antara penaksir regresi cum-rasio \bar{Y}_{r2} dengan penaksir regresi cum-rasio \bar{Y}_{r1} diperoleh $MSE(\bar{Y}_{r2}) < MSE(\bar{Y}_{r1})$ jika

$$-\beta_1 \bar{X} - \sqrt{\beta_1^2 \bar{X}^2 + \beta_1^4 + 2\bar{X}\beta_1^3} < \beta_2 < -\beta_1 \bar{X} + \sqrt{\beta_1^2 \bar{X}^2 + \beta_1^4 + 2\bar{X}\beta_1^3} .$$

2. Perbandingan antara penaksir regresi cum-rasio \bar{Y}_{r3} dengan penaksir regresi cum-rasio \bar{Y}_{r1} diperoleh $MSE(\bar{Y}_{r3}) < MSE(\bar{Y}_{r1})$ jika

$$\beta_2 > \frac{\beta_1 \bar{X} + \sqrt{\beta_1^2 \bar{X}^2 + \beta_1^4 + 2\bar{X}\beta_1^3}}{\beta_1^2 + 2\bar{X}\beta_1} \text{ atau } \beta_2 < \frac{\beta_1 \bar{X} - \sqrt{\beta_1^2 \bar{X}^2 + \beta_1^4 + 2\bar{X}\beta_1^3}}{\beta_1^2 + 2\bar{X}\beta_1} .$$

3. Perbandingan antara penaksir regresi cum-rasio \bar{Y}_{r2} dengan penaksir regresi cum-rasio \bar{Y}_{r3} diperoleh $MSE(\bar{Y}_{r2}) < MSE(\bar{Y}_{r3})$ jika

- a. $\beta_1 > 0$ dan $\beta_2 < -\beta_1 (\bar{X} + \sqrt{\bar{X}^2 - 1})$ atau $-\beta_1 (\bar{X} - \sqrt{\bar{X}^2 - 1}) < \beta_2 < -\beta_1$ atau $\beta_2 > \beta_1$.

- b. $\beta_1 < 0$ dan $\beta_2 < \beta_1$ atau $-\beta_1 < \beta_2 < -\beta_1 (\bar{X} - \sqrt{\bar{X}^2 - 1})$.

5. CONTOH

Contoh berikut merupakan pendapatan petani karet per hektar per tahun dan biaya tenaga kerja dari Desrita Lestari [2]. Data tersebut diteliti di Air Tiris, Kabupaten Kampar, Kecamatan Kampar, Provinsi Riau pada tahun 2008. Data pendapatan dan biaya tenaga kerja diberikan pada Tabel 1.

Tabel 1. Data pendapatan dan biaya tenaga kerja

No	Pendapatan (Y)	Biaya Tenaga Kerja (X)	No	Pendapatan (Y)	Biaya Tenaga Kerja (X)
1	16.096.000	4.050.000	21	18.768.000	3.400.000
2	16.524.000	4.400.000	22	23.052.000	4.950.000
3	14.280.000	5.100.000	23	15.912.000	4.650.000
4	16.320.000	3.200.000	24	18.971.000	4.050.000
5	16.728.000	5.200.000	25	16.932.000	3.250.000
6	16.524.000	3.650.000	26	21.828.000	3.400.000
7	13.185.600	3.400.000	27	17.748.000	3.250.000
8	15.504.000	3.950.000	28	14.076.000	3.650.000
9	19.152.000	5.000.000	29	13.056.000	5.100.000
10	12.240.000	3.950.000	30	22.032.000	3.950.000
11	17.952.000	4.950.000	31	12.052.800	3.950.000
12	17.952.000	5.150.000	32	22.908.000	4.000.000
13	14.592.000	4.050.000	33	19.788.000	4.400.000
14	15.096.000	4.350.000	34	12.288.000	4.550.000
15	14.892.000	3.850.000	35	16.524.000	4.700.000
16	16.728.000	3.550.000	36	15.096.000	3.750.000
17	13.056.000	4.850.000	37	12.864.000	3.050.000
18	20.800.800	3.900.000	38	15.504.000	3.850.000
19	10.036.800	3.250.000	39	18.564.000	4.600.000
20	11.832.000	3.650.000	40	20.808.000	3.700.000

Sumber: Desrita Lestari [2].

Dengan menggunakan data pada Tabel 1 akan ditentukan penaksir regresi cum-rasio yang efisien untuk menaksir rata-rata pendapatan petani karet per hektar per tahun dengan menggunakan syarat penaksir lebih efisien yang diperoleh sebelumnya. Hal ini secara umum dapat ditunjukkan dengan menghitung *MSE* dari masing-masing penaksir. Sebagai informasi tambahan untuk menaksir rata-rata pendapatan petani akret per hektar per tahun digunakan biaya tenaga kerja. Untuk menghitung *MSE* dari masing-masing penaksir terlebih dahulu ditentukan nilai yang dibutuhkan. Informasi yang diperoleh dari pendapatan dan biaya tenaga kerja dengan menggunakan Microsoft Excel, yaitu

$$\begin{array}{lll}
 N = 40 & \rho = 0,1108 & R_{r_3} = -3,5160 \\
 n = 10 & S_x = 630,41 & \beta_1 = 0.2784 \\
 f = 0,25 & S_y = 3249,22 & \beta_2 = -1.0585 \\
 \bar{X} = 4091.25 & R_{r_1} = 4,2947 & \\
 \bar{Y} = 16431.6 & R_{r_2} = 13,3691 &
 \end{array}$$

Dengan menggunakan informasi sebelumnya, diperoleh bahwa

- (i) $MSE(\bar{Y}_{r_2}) < MSE(\bar{Y}_{r_1})$ jika $-2278.08 < \beta_2 < 0.775$.
- (ii) $MSE(\bar{Y}_{r_3}) < MSE(\bar{Y}_{r_2})$ jika $\beta_2 > 1139.51$ atau $\beta_2 < 1138.50$.
- (iii) $MSE(\bar{Y}_{r_2}) < MSE(\bar{Y}_{r_3})$ jika $\beta_2 < -2278.079$ atau $-2278.081 < \beta_2 < -0.2784$
atau $\beta_2 > 0.2784$.

Selanjutnya nilai MSE dari masing-masing penaksir diberikan pada Tabel 2.

Tabel 2. Nilai MSE dari masing-masing penaksir

No	Penaksir	MSE
1	\bar{Y}_{r_1}	17.313,92
2	\bar{Y}_{r_2}	13.779,87
3	\bar{Y}_{r_3}	15.534,02

Berdasarkan Tabel 2, dapat dilihat bahwa penaksir rasio \bar{Y}_{r_2} memiliki nilai MSE yang terkecil dengan syarat bahwa kondisi lebih efisien dapat terpenuhi.

6. KESIMPULAN

Dari pembahasan di atas, dapat disimpulkan bahwa penaksir regresi cum-rasio \bar{Y}_{r_2} merupakan penaksir yang paling efisien dari penaksir regresi cum-rasio \bar{Y}_{r_1} dan penaksir regresi cum-rasio \bar{Y}_{r_3} jika syarat efisien terpenuhi.

DAFTAR PUSTAKA

- [1] Cochran, W. G. 1991. *Teknik Penarikan Sampel, Edisi Ketiga*. Terj. Dari *Sampling Techniques*, oleh Rudiandiyah & E. R Osman. UI Press, Jakarta.
- [2] Lestari, Desrita. 2008. Analisis faktor-faktor yang mempengaruhi Pendapatan Petani Karet di Kelurahan Air Tiris Kecamatan Kampar. Skripsi Jurusan Sosial Ekonomi Pertanian (Agrobisnis) Fakultas Pertanian Universitas Riau, Pekanbaru.
- [3] Myers, R. H. 1986. *Classical and Modern Regression with Applications*. Wadsworth, Inc, America.
- [4] Phillips, G. M. & P. J. Taylor. 1972. *Theory and Applications of Numerical Analysis. Second Edition*. Academic Press, New York.
- [5] Yan Z. & B. Tian. 2010. Ratio Method to the Mean Estimation Using Coefficient of Skewness of Auxiliary Variable. *ICICA, Part II, CCIS*. **106**: 103-110.