

PENAKSIR RASIO REGRESI LINEAR UNTUK RATA-RATA POPULASI MENGGUNAKAN VARIABEL TAMBAHAN, KOEFISIEN VARIASI DAN KOEFISIEN KURTOSIS

Malini Anggraeni^{1*}, Sigit Sugiarto², Rustam Efendi²

¹ Mahasiswa Program Studi S1 Matematika

² Dosen Jurusan Matematika

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

*lini.3011@yahoo.co.id

ABSTRACT

This paper discusses three ratio regression linear estimators for population mean \bar{Y} using two auxiliary variables X and Z , using the coefficient of variation and coefficient of kurtosis on simple random sampling. This paper is a review of the article Singh et. al [*Statistics in Transition* 10(1): 85-100]. The three estimators are biased estimators and their mean square errors are determined. Each estimator compared to the value of the mean square error. Comparison shows that the estimator with the minimum mean square error is an efficient estimator.

Keywords: *ratio estimator, coefficient of variation, coefficient of kurtosis, biased estimator and mean square error.*

ABSTRAK

Pada artikel ini dibahas tiga penaksir rasio regresi linear untuk rata-rata populasi \bar{Y} menggunakan dua variabel tambahan X dan Z , menggunakan koefisien variasi dan koefisien kurtosis pada sampling acak sederhana. Tulisan ini merupakan *review* dari artikel Singh et. al [*Statistics in Transition* 10(1): 85-100]. Ketiga penaksir merupakan penaksir bias dan ditentukan *mean square error*. Setiap penaksir dibandingkan nilai *mean square error*. Perbandingan menunjukkan bahwa penaksir yang memiliki *mean square error* minimum adalah penaksir yang efisien.

Kata kunci: *penaksir rasio, koefisien variasi, koefisien kurtosis, penaksir bias, dan mean square error.*

1. PENDAHULUAN

Terdapat populasi berkarakter Y berukuran N yang akan diteliti, untuk menaksir parameter dari suatu populasi digunakan analisa data sampel yang diperoleh dari populasi tersebut. Metode statistika yang berhubungan dengan pengambilan sebagian dari populasi disebut penarikan sampling. Di sini penarikan sampling dilakukan dengan cara acak karena cara ini dianggap obyektif. Penarikan sampling acak sederhana

(*simple random sampling*) merupakan sebuah metode untuk memilih n unit dari N dalam populasi, dimana setiap unit memiliki kesempatan yang sama untuk diambil sebagai unit sampel. Unit-unit yang terpilih ini sebanyak n merupakan sampel. Banyak kemungkinan sampel yang terpilih dengan peluang anggota populasi terambil sebagai sampel sama adalah C_n^N [2: h. 21].

Sebuah metode dimana bilangan yang telah diambil dari populasi tidak akan terpilih kembali untuk menjadi unit sampel dalam penarikan berikutnya adalah penarikan sampling acak tanpa pengembalian, sebaliknya sebuah metode dimana bilangan yang diambil dari populasi mempunyai kesempatan untuk terpilih kembali menjadi unit sampel dalam penarikan berikutnya disebut penarikan sampling acak dengan pengembalian. Di sini sampling dilakukan tanpa pengembalian agar karakteristik unit-unit lebih akurat.

Setelah sampel y_i berukuran n diperoleh, diharapkan y_i dapat menaksir populasi Y , namun taksiran ini memiliki kesalahan dan variansi yang besar untuk meningkatkan ketelitian taksiran digunakan variabel tambahan X dan selanjutnya dilakukan metode rasio. Dalam metode rasio suatu variabel pendukung x_i yang berhubungan dengan y_i diperoleh untuk setiap unit di dalam sampel. Jumlah populasi X dari x_i harus diketahui. Nilai x_i diperoleh saat dilakukan sensus lengkap. Tujuan metode ini adalah untuk meningkatkan ketelitian dengan mengambil manfaat hubungan antara y_i dan x_i . Metode rasio digunakan saat terdapat korelasi yang tinggi antara variabel pengamatan Y dan variabel tambahan X . Penaksir rasio untuk populasi Y dengan pengambilan sampel secara sampling acak sederhana adalah $\bar{y}_R = \hat{R}\bar{x}$, dengan $\hat{R} = \bar{y}/\bar{x}$ merupakan penaksir untuk rasio $R = \bar{Y}/\bar{X}$ [2: h. 173].

Pada suatu penaksiran terdapat bias yang dapat mempengaruhi kesalahan taksiran, untuk mengetahui seberapa besar kesalahan tersebut dapat dilihat dengan menghitung besar bias dari penaksir. Apabila ekspektasi dari suatu penaksir menghasilkan penaksir itu sendiri penaksir tersebut dinamakan penaksir tak bias, dan apabila ekspektasi dari penaksir tidak sama dengan parameter yang ditaksir maka penaksir tersebut dinamakan penaksir bias. Dari penaksir rasio sederhana tersebut Singh et. al [3] mengkombinasikan antara penaksir rasio dan penaksir regresi menjadi penaksir rasio regresi untuk rata-rata populasi.

2. PENAKSIR RASIO REGRESI UNTUK RATA-RATA POPULASI

Untuk mendapatkan nilai-nilai taksiran dari parameter yang efisien digunakan metode penaksir rasio regresi linear untuk \bar{Y} . Penaksir rasio regresi untuk rata-rata populasi yang akan dibandingkan merupakan penaksir rasio regresi menggunakan dua variabel tambahan X dan Z , koefisien variasi dan koefisien kurtosis, yaitu sebagai berikut

$$\bar{y}_{RR1} = \bar{y} + B_{yx} (\bar{X} - \bar{x}) \left[\frac{\bar{Z} + C_z}{\bar{z} + C_z} \right]$$

$$\bar{y}_{RR1} = \bar{y}_{lr} \left[u + (-b^*) \right]^1, \quad (1)$$

$$\bar{y}_{RR2} = \bar{y} + B_{yx} \left[\bar{x} - \bar{x} \right] \left[\frac{\bar{z} + C_z}{\bar{Z} + C_z} \right]^{-\alpha}$$

$$\bar{y}_{RR2} = \bar{y}_{lr} \left[u + (-b^*) \right]^\alpha, \quad (2)$$

$$\bar{y}_{RR3} = \bar{y} + B_{yx} \left[\bar{x} - \bar{x} \right] \left[\frac{\bar{Z}\beta_2 \bar{c} + C_z}{\bar{z}\beta_2 \bar{c} + C_z} \right]$$

$$\bar{y}_{RR3} = \bar{y}_{lr} \left[u + (-c) \right], \quad (3)$$

dengan

$$b^* = \frac{\bar{Z}}{\bar{Z} + C_z}, c = \frac{\bar{Z}\beta_2 \bar{c}}{\bar{z}\beta_2 \bar{c} + C_z}, u = \frac{\bar{z}}{\bar{Z}}, \text{ dan } y_{lr} = \bar{y} + B_{yx} \left[\bar{x} - \bar{x} \right].$$

$$C_z = \frac{S_z}{\bar{X}} \text{ merupakan koefisien variasi,}$$

$$S_z = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{Z})^2}{N-1}} \text{ merupakan simpangan baku,}$$

$$\beta_2 \bar{c} = \frac{N(N+1) \sum_{i=1}^N (x_i - \bar{Z})^4}{(N-1)(N-2)(N-3) S_z^4} - \frac{3(N-1)^2}{(N-2)(N-3)} \text{ merupakan koefisien kurtosis,}$$

α merupakan sebuah konstanta dengan nilai $0 < \alpha < 1$, dan

B_{yx} merupakan koefisien regresi untuk variabel Y dan X .

Penaksir pada persamaan (1), (2), dan (3) merupakan penaksir untuk parameter Y . Selanjutnya akan ditentukan sifat bias pada penaksir dan *mean square error* (MSE).

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

Terdapat bias yang dapat mempengaruhi kesalahan taksiran. Besar bias dari suatu penaksir diperoleh dengan menentukan nilai ekspektasi dari penaksir tersebut. Untuk mengetahui seberapa besar kesalahan taksiran tersebut dapat dilihat dengan menghitung bias dari penaksir yang diperoleh dengan menggunakan deret Taylor untuk k variabel

Definisi 1 (Deret Taylor untuk k variabel) [5: h. 47] Misalkan $f(x_1, x_2, \dots, x_k)$ adalah suatu fungsi k variabel dan $f, f', f'', \dots, f^{(k)}$ adalah kontinu pada domain D dan $f^{(k+1)}$ ada pada D dengan $(x_1, x_2, \dots, x_k) \in D$, maka

$$\begin{aligned}
 f(x_1, x_2, \dots, x_k) &= f(x_{10}, x_{20}, \dots, x_{k0}) + \sum_{r=1}^k \frac{1}{r!} \left((x_1 - x_{10}) \frac{\partial f(x_{10}, x_{20}, \dots, x_{k0})}{\partial x_1} \right. \\
 &+ (x_2 - x_{20}) \frac{\partial f(x_{10}, x_{20}, \dots, x_{k0})}{\partial x_2} + \dots + (x_k - x_{k0}) \frac{\partial f(x_{10}, x_{20}, \dots, x_{k0})}{\partial x_k} \Big)^r \\
 &+ \frac{1}{(k+1)!} \left((x_1 - x_{10}) \frac{\partial f(x_{10}, x_{20}, \dots, x_{k0})}{\partial x_1} + (x_2 - x_{20}) \frac{\partial f(x_{10}, x_{20}, \dots, x_{k0})}{\partial x_2} \right. \\
 &+ \dots + (x_k - x_{k0}) \frac{\partial f(x_{10}, x_{20}, \dots, x_{k0})}{\partial x_k} \Big)^{k+1}
 \end{aligned}$$

Menentukan besar bias digunakan Definisi 1 dimana $k = 3$ dan dengan mengabaikan pangkat yang lebih besar dari dua, maka besar bias setiap penaksir pada persamaan (1), (2), dan (3) adalah sebagai berikut:

$$\begin{aligned}
 B_{RR1} &\approx \frac{\bar{Y}}{\bar{C}_z} S_z^2 - 2 \frac{1}{\bar{C}_z} S_{yz} + 2 \frac{1}{\bar{C}_z} B_{yx} S_{xz}, \\
 B_{RR2} &\approx \frac{1}{2} \frac{\alpha^2 \bar{Y} + \alpha \bar{Y}}{\bar{C}_z} S_z^2 - \frac{\alpha}{\bar{C}_z} S_{yz} + \frac{\alpha}{\bar{C}_z} B_{yx} S_{xz}, \\
 B_{RR3} &\approx \frac{\bar{Y} \beta_2}{\bar{C}_z} S_z^2 - \frac{1}{\bar{C}_z} S_{yz} + \frac{1}{\bar{C}_z} B_{yx} S_{xz},
 \end{aligned}$$

dengan

$$B_{yx} = \frac{S_{yx}}{S_x^2}.$$

Dari hasil yang diperoleh bias dari masing-masing penaksir tidak sama dengan nol, maka penaksir tersebut merupakan penaksir yang bias. Selanjutnya, ketelitian suatu penaksir bias ditinjau dari besar MSE untuk masing-masing penaksir. Definisi yang digunakan untuk menentukan besar MSE untuk setiap penaksir adalah sebagai berikut

Definisi 2 [1: h.309] Jika $\hat{\theta}^*$ adalah penaksir bias untuk θ , maka $MSE(\hat{\theta}^*)$ didefinisikan

$$MSE(\hat{\theta}^*) = E(\hat{\theta}^* - \theta)^2.$$

Berdasarkan Definisi 2 untuk menentukan MSE digunakan Definisi 1 dimana $k = 3$ dan dengan mengabaikan pangkat yang lebih besar dari satu, maka besar MSE untuk setiap penaksir pada persamaan (1), (2), dan (3) adalah sebagai berikut:

$$MSE(\hat{y}_{RR1}) \approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + b^{*2} C_z^2 + 2b^* \rho_{yx} C_y \rho_{xz} C_z - 2b^* \rho_{yz} C_y C_z) \quad (4)$$

$$MSE(\hat{y}_{RR2}) \approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + \alpha^2 b^{*2} C_z^2 + 2\alpha b^* \rho_{yx} C_y \rho_{xz} C_z - 2\alpha b^* \rho_{yz} C_y C_z) \quad (5)$$

$$MSE(\hat{y}_{RR3}) \approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + c^2 C_z^2 + 2c \rho_{yx} C_y \rho_{xz} C_z - 2c \rho_{yz} C_y C_z) \quad (6)$$

dengan

$$C_y^2 = \frac{S_y^2}{\bar{Y}^2}, \rho_{yx}^2 = \frac{S_{yx}^2}{S_y^2 S_x^2}, \text{ dan } \rho_{yx} = \frac{S_{yx}}{S_y S_x}.$$

ρ_{yx} merupakan koefisien korelasi antara variabel Y dan X .

Setelah besar bias dan MSE diketahui dari masing-masing penaksir, dan menghasilkan besar bias dan MSE yang berbeda. Selanjutnya akan ditentukan penaksir rasio regresi yang efisien untuk rata-rata populasi dari setiap penaksir.

4. PENAKSIR RASIO REGRESI YANG EFISIEN UNTUK RATA-RATA POPULASI

Penaksir pada persamaan (1), (2), dan (3) telah diketahui merupakan penaksir yang bersifat bias, maka untuk mengetahui penaksir yang efisien dapat digunakan dengan menentukan efisiensi relatifnya yang didefinisikan sebagai berikut

Definisi 3 [4: h. 272] Misalkan $\hat{\theta}_1^*$ dan $\hat{\theta}_2^*$ adalah penaksir bias untuk θ , dengan MSE berturut-turut dari penaksir $\hat{\theta}_1^*$ dan $\hat{\theta}_2^*$ adalah $MSE(\hat{\theta}_1^*)$ dan $MSE(\hat{\theta}_2^*)$, maka efisien relatif $\hat{\theta}_1^*$ terhadap $\hat{\theta}_2^*$ dinotasikan dengan $RE(\hat{\theta}_1^*, \hat{\theta}_2^*)$ dan didefinisikan dengan

$$RE(\hat{\theta}_1^*, \hat{\theta}_2^*) = \frac{MSE(\hat{\theta}_2^*)}{MSE(\hat{\theta}_1^*)}$$

Berdasarkan Definisi 3, untuk menentukan penaksir yang efisien adalah dengan cara membandingkan besar nilai MSE yang diperoleh dari persamaan (4), (5), dan (6) sebagai berikut:

a. Perbandingan $MSE(\hat{y}_{RR1})$ dan $MSE(\hat{y}_{RR2})$

Penaksir \bar{y}_{RR2} lebih efisien dari penaksir \bar{y}_{RR1} jika

$$\rho_{yz} > \frac{C_z b^* (\alpha + \alpha^2)}{2C_y} + \rho_{yx} \rho_{xz} \quad (7)$$

b. Perbandingan $MSE(\hat{y}_{RR1})$ dan $MSE(\hat{y}_{RR3})$

Penaksir \bar{y}_{RR3} lebih efisien dari penaksir \bar{y}_{RR1} jika

$$\rho_{yz} > \frac{b^* + c \hat{C}_z}{2C_y} + \rho_{yx}\rho_{xz}. \quad (8)$$

c. Perbandingan $MSE_{\bar{y}_{RR2}}$ dan $MSE_{\bar{y}_{RR3}}$
 Penaksir \bar{y}_{RR2} lebih efisien dari penaksir \bar{y}_{RR3} jika

$$\rho_{yz} < \frac{b^* + c \hat{C}_z}{2C_y} + \rho_{yx}\rho_{xz}. \quad (9)$$

Akan ditunjukkan penaksir yang memenuhi syarat efisiensi dengan menggunakan contoh yang akan dibahas pada bagian selanjutnya.

Contoh Sebagai contoh dari pembahasan sebelumnya, diberikan data skunder mengenai analisis faktor-faktor yang mempengaruhi pendapatan pengusaha warung internet di kota Pekanbaru. Data diperoleh dengan penelitian yang dilakukan oleh peneliti [6] di wilayah Kecamatan Bukit Raya, Kecamatan Tampan, Kecamatan Rumbai, Kecamatan Marpoyan Damai dan Kecamatan Sukajadi. Populasi penelitian ini adalah para pengusaha warung internet sebanyak 50 orang/ 50 warnet yang terdapat di Kota Pekanbaru. Penelitian ini dilakukan untuk menaksir rata-rata pendapatan pengusaha yang diterima dengan sampling acak sederhana dan memanfaatkan informasi tambahan berupa biaya operasional, dimana biaya operasional ini adalah biaya yang dikeluarkan oleh pengusaha untuk menunjang berjalannya usaha warnet, seperti upah/gaji karyawan, dan jumlah pelanggan warung internet di kota Pekanbaru.

Tabel 1: Data Pendapatan Pengusaha Warung Internet di Kota Pekanbaru

No.	Y Pendapatan Pengusaha (Rp)	X Biaya Operasional (Rp)	Z Jumlah Pelanggan
1	18.324.000	8.108.500	4.150
2	21.756.000	10.037.500	5.350
3	10.732.000	6.996.000	2.950
4	12.380.000	6.389.500	3.520
5	9.135.000	3.573.250	2.400
6	11.668.000	3.831.500	3.300
7	12.652.000	6.707.500	2.850
8	12.150.000	5.038.500	3.080
9	10.733.000	5.581.500	2.300
10	12.221.000	6.254.750	3.200
11	26.282.000	7.896.750	5.550
12	8.585.000	6.147.500	2.100
13	6.658.000	4.188.250	1.600
14	24.528.000	12.821.500	5.700
15	27.462.000	11.316.000	5.970

16	15.211.000	6.883.250	3.900
17	19.843.000	9.285.000	5.100
18	6.792.000	4.317.750	1.230
19	5.640.000	2.895.000	1.080
20	17.201.000	7.588.750	4.200
21	9.113.000	7.111.000	2.500
22	10.568.000	4.410.250	2.800
23	8.950.000	3.495.250	2.180
24	11.495.000	4.285.000	2.760
25	9.104.000	5.067.250	2.400
26	8.487.000	4.025.000	2.200
27	7.513.000	2.736.750	1.950
28	7.324.000	3.604.250	1.840
29	5.812.000	3.146.250	1.300
30	16.081.000	4.986.000	3.800
31	12.930.000	5.928.250	3.500
32	12.133.000	4.077.000	3.800
33	8.412.000	4.481.000	2.150
34	8.527.000	3.063.750	2.300
35	10.336.000	5.800.000	2.500
36	20.658.000	4.451.500	4.700
37	15.343.000	4.007.500	3.210
38	12.822.000	3.685.750	2.900
39	13.797.000	6.739.250	3.760
40	11.642.000	5.512.500	3.150
41	7.263.000	4.527.500	1.800
42	5.780.000	3.452.500	950
43	5.694.000	3.623.500	1.100
44	7.812.000	4.138.000	2.000
45	11.230.000	5.526.250	3.680
46	6.179.000	3.157.500	1.250
47	6.887.000	3.803.750	1.620
48	7.364.000	4.102.000	1.730
49	9.137.000	4.950.000	2.350
50	5.416.000	3.334.000	1.300
	$\bar{Y} = 11.675.240$	$\bar{X} = 5.341.735$	$\bar{Z} = 2.860,2$

Sumber: Reliabilita (2014).

Dari data pada Tabel 1 dan dengan bantuan Microsoft Excel diperoleh nilai-nilai sebagaimana yang tertera pada Tabel 2.

Tabel 2: Nilai-nilai yang diperlukan untuk membandingkan MSE dari kedua penaksir, dengan $\alpha = 0,5$

N	50	S_{yx}	9.666.590.000.000
\bar{Y}	11.675.240	S_{yz}	6.755.609.135
\bar{X}	5.341.735	S_{xz}	2.270.336.840
\bar{Z}	2.860,2	ρ_{yx}	0,811254664
S_x	2.185.030	ρ_{yz}	0,966423844
S_y	5.453.288	ρ_{xz}	0,810577174
S_z	1.281.854	-	-

dengan mensubstitusikan nilai-nilai yang diperoleh dari pendapatan pengusaha warung internet di kota Pekanbaru, biaya operasional, dan jumlah pelanggan ke persamaan (7), (8), (9) maka diperoleh

1. $MSE_{\hat{Y}_{RR2}} < MSE_{\hat{Y}_{RR1}}$ jika $\rho_{yz} > 0,89742444$.
2. $MSE_{\hat{Y}_{RR3}} < MSE_{\hat{Y}_{RR1}}$ jika $\rho_{yz} < 1,618210182$
3. $MSE_{\hat{Y}_{RR2}} < MSE_{\hat{Y}_{RR3}}$ jika $\rho_{yz} < 1,378370285$

Tabel 3: Nilai *Mean Square Error* dan nilai taksiran masing-masing penaksir

Penaksir Rasio	MSE	Nilai Taksiran
\bar{y}_{RR1}	$3,98292 \times 10^{11}$	12.449.334,46
\bar{y}_{RR2}	$1,6396 \times 10^{11}$	12.423.631,55
\bar{y}_{RR3}	$4,00255 \times 10^{11}$	12.449.470,54

Dari Tabel 3 dapat dilihat bahwa \bar{y}_{RR2} memiliki nilai MSE minimum dan memiliki nilai taksiran yang mendekati nilai rata-rata populasi.

5. KESIMPULAN

Setelah diperoleh nilai MSE dari penaksir rasio regresi menggunakan dua variabel tambahan untuk rata-rata populasi pada sampling acak sederhana, kemudian membandingkan MSE dari masing-masing penaksir, sehingga dapat disimpulkan bahwa berdasarkan hasil yang diperoleh pada Tabel 3 penaksir rasio regresi \hat{Y}_{RR2} lebih efisien dibandingkan dengan penaksir rasio regresi \hat{Y}_{RR1} dan penaksir rasio regresi \hat{Y}_{RR3} .

DAFTAR PUSTAKA

- [1] Bain, L. J. & M, Engelhardt. 1991. *Introduction to Probability and Mathematical Statistics, Second Edition*. Duxbury Press, California.
- [2] Cochran, W. G. 1991. *Teknik Penarikan Sampel, Edisi Ketiga*. Terj. Dari *Sampling Techniques*, oleh Rudiansyah & E. R Osman. Penerbit Universitas Indonesia, Jakarta.
- [3] Singh, H. P., Upadhyaya, L. N. & Premchandra, 2009. *An Improved Version of Regression Ratio Estimator With Two Auxiliary Variables in Sample Surveys*, *Statistics in Transition*, 10(1): 85-100.
- [4] Montgomery, D. C. & G. C. Runger. 1999. *Applied Statistics and Probability for Engineers, Second Edition*. John Willey & Sons, New York.
- [5] Phillips, G. M. & P. J. Taylor. 1972. *Theory and Applications of Numerical Analysis, Second Edition*. Academic Press, New York.
- [6] Reliabilita. 2014. Analisis Faktor-faktor yang Mempengaruhi Pendapatan Pengusaha Warung Internet di Kota Pekanbaru. Skripsi Fakultas Ekonomi Universitas Riau, Pekanbaru.