

PENAKSIR RASIO REGRESI LINEAR YANG EFISIEN UNTUK RATA-RATA POPULASI PADA SAMPLING ACAK SEDERHANA

Ratni Dakhyu^{1*}, Rustam Efendi², Haposan Sirait²

¹ Mahasiswa Program S1 Matematika

² Dosen Jurusan Matematika

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

* ratnii_dachyou@yahoo.co.id

ABSTRACT

In this article we study three linear regression ratio estimators for the population mean on simple random. This study is a review of the article of Singh, et. al [*Statistics in Transition-new series*, 10: 85-100]. Each estimator is a biased estimator and their mean square errors are determined. Estimator with the smallest mean square error is the most efficient one. An example is given at the end of the discussion.

Keywords: *Simple random sampling, ratio regression estimator, bias and mean square error.*

ABSTRAK

Dalam artikel ini dibahas tiga penaksir rasio regresi linear untuk rata-rata populasi pada sampling acak sederhana yang merupakan tinjauan dari artikel Singh, et. al [*Statistics in Transition-new series*, 10: 85-100]. Masing-masing penaksir merupakan penaksir bias dan ditentukan *mean square error*. Penaksir dengan *mean square error* terkecil merupakan penaksir yang paling efisien. Sebuah contoh diberikan pada akhir pembahasan.

Kata kunci: *Sampling acak sederhana, penaksir rasio regresi, bias dan mean square error.*

1. PENDAHULUAN

Untuk memperoleh sampel yang representatif dilakukan teknik pengambilan sampel yang disebut teknik sampling. Metode pengambilan sampling ada dua yaitu sampling acak dan sampling tidak acak. Sampling acak yang terbagi atas sampling acak sederhana, sampling acak sistematis, sampling acak berstrata dan sampling acak cluster. Sampling acak sederhana yaitu metode pengambilan sampel dari anggota populasi secara acak sehingga mempunyai kesempatan yang sama untuk dipilih [1: h. 21].

Metode yang digunakan untuk menaksir rata-rata populasi pada sampling acak sederhana diantaranya adalah metode rasio, regresi dan gabungan rasio regresi.

Metode rasio yaitu memanfaatkan hubungan rasio antara y_i dan x_i , dengan y_i adalah unit dari populasi Y dan x_i adalah unit dari populasi X , dan diasumsikan rata-rata populasi X dari x_i diketahui. Penaksir regresi memanfaatkan hubungan X yang mempunyai regresi dengan Y .

Misalkan bahwa informasi pada dua variabel tambahan X dan Z tersedia, dengan Z dapat berupa nilai saat dilakukan sensus yang memiliki hubungan rasio dengan variabel X dan Y . Variabel X merupakan variabel tambahan lainnya yang dijumpai saat Y diteliti dan sangat berkorelasi dengan Y . Asumsikan bahwa \bar{X} dan \bar{Z} adalah rata-rata populasi dari variabel tambahan dengan X dan Z diketahui dan terdapat korelasi yang tinggi antara Y dan X , Y dan Z , dimana mungkin atau tidak terdapat korelasi apapun antara X dan Z [3].

2. SAMPLING ACAK SEDERHANA

Pengambilan sampel acak sederhana merupakan suatu metode yang digunakan untuk mengambil n unit sampel dari N unit populasi dimana setiap unit memiliki kesempatan yang sama untuk terpilih sebagai unit sampel. Pengambilan sampel dapat dilakukan dengan pengembalian atau tanpa pengembalian. Pengambilan sampel dengan pengembalian menyatakan satuan unit yang telah diambil memiliki kemungkinan akan terpilih kembali menjadi anggota sampel, sedangkan pengambilan sampel tanpa pengembalian menyatakan unit sampel yang telah diambil tidak akan mungkin terpilih pada pengambilan berikutnya. Oleh karena itu, dianjurkan sampling dilakukan tanpa pengembalian agar karakteristik unit-unit lebih akurat.

Pada pengambilan sampel tanpa pengembalian probabilitas terpilih n dari N unit populasi sebagai unit sampel pada pengambilan pertama yaitu $1/N$, probabilitas pada pengambilan kedua yaitu $1/(N-1)$ sampai probabilitas pada pengambilan ke- n yaitu $1/(N-(n-1))$. Maka probabilitas seluruh n unit-unit tertentu yang terpilih dalam n pengambilan adalah $(C_n^N)^{-1}$.

Misalkan suatu populasi berukuran N dengan nilai variabel y_i untuk masing-masing unit, $i = 1, 2, 3, \dots, N$ maka rata-rata populasi \bar{Y} adalah

$$\bar{Y} = \frac{1}{N} \sum_{i=1}^N y_i.$$

Kemudian akan diambil sampel berukuran n unit dengan variabel y_i , $i = 1, 2, 3, \dots, n$ maka rata-rata sampel \bar{y} adalah

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

Untuk menentukan bias dan MSE dari masing-masing penaksir rasio regresi linear digunakanlah teorema mengenai variansi dan kovariansi.

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} dirumuskan sebagai berikut

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

dengan

$$f = \frac{n}{N} \text{ adalah fraksi penarikan sampel}$$

$$S_y^2 = \frac{\sum_{i=1}^N (y_i - \bar{Y})^2}{N - 1} \text{ adalah variansi } y_i \text{ pada populasi berkarakter } Y.$$

Bukti: Dapat dilihat pada [1: h.27]. ■

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

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

Bukti: Dapat dilihat pada [1: h.29]. ■

3. BIAS DAN *MSE* PENAKSIR RASIO REGRESI LINEAR YANG EFISIEN

UNTUK RATA-RATA POPULASI PADA SAMPLING ACAK SEDERHANA

Bentuk dari penaksir rasio regresi linear untuk rata-rata populasi pada sampling acak sederhana yang diajukan oleh Singh, et. al [3] yaitu:

$$\left(\hat{Y}_{RR}^{(1)} \right) = \left(\bar{y} + b_{yx} (\bar{X} - \bar{x}) \left[2 - \left(\frac{\bar{z}}{\bar{Z}} \right)^\alpha \right] \right) \quad (1)$$

$$\left(\hat{Y}_{RR}^{(2)} \right) = \left(\bar{y} + b_{yx} (\bar{X} - \bar{x}) \right) \alpha + \left[1 + (1 - \alpha) \left(\frac{\bar{z}}{\bar{Z}} \right)^{-1} \right] \quad (2)$$

$$\left(\hat{Y}_{RR}^{(3)} \right) = \left(\bar{y} + b_{yx} (\bar{X} - \bar{x}) \right) \left[\alpha + (1 - \alpha) \frac{\bar{z}}{\bar{Z}} \right] \quad (3)$$

dengan α merupakan sebuah konstanta yang tepat dengan nilai $0 < \alpha < 1$.

Masing-masing penaksir yang dibahas merupakan penaksir bias. Akan ditentukan besarnya bias dan *MSE* dari masing-masing penaksir. Suatu penaksir dikatakan efisien apabila memiliki *MSE* minimum. Berikut adalah bias dan *MSE* dari ketiga penaksir:

1. Bias dan *MSE* penaksir rasio regresi linear $\hat{Y}_{RR}^{(1)}$ pada persamaan (1) yaitu

$$B(\hat{Y}_{RR}^{(1)}) \approx \frac{\alpha^2 - \alpha\bar{Y}}{\bar{Z}^2} S_z^2 - \frac{2\alpha}{\bar{Z}} \rho_{yz} + \frac{2\alpha}{\bar{Z}} B_{yx} \rho_{xz}$$

$$MSE(\hat{Y}_{RR}^{(1)}) \approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{xy}^2 C_y^2 + \alpha^2 C_z^2 - 2\alpha \rho_{yz} C_y C_z + 2\alpha \rho_{yx} \rho_{xz} C_y C_z). \quad (4)$$

2. Bias dan *MSE* penaksir rasio regresi linear $\hat{Y}_{RR}^{(2)}$ pada persamaan (2) yaitu

$$B(\hat{Y}_{RR}^{(2)}) \approx \alpha\bar{Y} + 2(1-\alpha)\bar{Y} - \frac{(1-\alpha)}{\bar{Z}} \rho_{lr} S_y S_z + \bar{Y} \cdot \bar{Z}$$

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

3. Bias dan *MSE* penaksir rasio regresi linear $\hat{Y}_{RR}^{(3)}$ pada persamaan (3) yaitu

$$B(\hat{Y}_{RR}^{(3)}) \approx \alpha\bar{Y} + \frac{\alpha}{\bar{Z}} \rho_{lr} S_y S_z + \bar{Y} \cdot \bar{Z}$$

$$MSE(\hat{Y}_{RR}^{(3)}) \approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{xy}^2 C_y^2 + (1-\alpha)^2 C_z^2 + 2(1-\alpha) \rho_{yz} C_y C_z - 2(1-\alpha) \rho_{yx} \rho_{xz} C_y C_z). \quad (6)$$

4. PENAKSIR RASIO REGRESI YANG EFISIEN

Untuk menentukan penaksir yang lebih efisien dari penaksir bias, dapat ditentukan dengan cara membandingkan *MSE* dari masing-masing penaksir. Penaksir yang lebih efisien merupakan penaksir dengan *MSE* terkecil.

- a. Perbandingan *MSE* $\hat{Y}_{RR}^{(1)}$ dengan *MSE* $\hat{Y}_{RR}^{(2)}$

Penaksir rasio regresi $\hat{Y}_{RR}^{(2)}$ lebih efisien dari penaksir rasio regresi $\hat{Y}_{RR}^{(1)}$, jika

$$\rho_{yz} > \frac{2(1-2\alpha)C_y \rho_{yx} \rho_{xz} - (1+2\alpha)C_z}{2(1-2\alpha)C_y}.$$

- b. Perbandingan *MSE* $\hat{Y}_{RR}^{(1)}$ dengan *MSE* $\hat{Y}_{RR}^{(3)}$

Penaksir rasio regresi $\hat{Y}_{RR}^{(3)}$ lebih efisien dari penaksir rasio regresi $\hat{Y}_{RR}^{(1)}$, jika

$$\rho_{yz} < \frac{2C_y \rho_{yx} \rho_{xz} + (1+2\alpha)C_z}{2C_y}.$$

- c. Perbandingan *MSE* $\hat{Y}_{RR}^{(2)}$ dengan *MSE* $\hat{Y}_{RR}^{(3)}$

Penaksir rasio regresi $\hat{Y}_{RR}^{(2)}$ lebih efisien dari penaksir rasio regresi $\hat{Y}_{RR}^{(3)}$, jika

$$\rho_{yz} > \rho_{yx} \rho_{xz}.$$

Contoh Diketahui data pada Tabel 1 mengenai frekuensi permintaan responden tanaman hias (Y) di Kecamatan Marpoyan Damai Kota Pekanbaru, Riau. Data ini merupakan sampel. Tapi, pada artikel ini data tersebut dianggap populasi dengan $N = 100$, kemudian dari populasi ini diambil sampel dengan $n = 40$, dan digunakan variabel tambahan yaitu harga tanaman hias (X) dan tingkat pendapatan responden (Z). Permintaan responden tanaman hias ini dipengaruhi oleh harga tanaman hias dan pendapatan responden. Dengan menggunakan data tersebut akan ditentukan penaksir rasio regresi linear yang paling efisien untuk menaksir rata-rata permintaan tanaman hias responden yakni dengan menggunakan syarat penaksir lebih efisien yang diperoleh sebelumnya dan secara umum dapat ditunjukkan dengan menghitung MSE dari masing-masing penaksir yang diberikan.

Tabel 1: Data jumlah permintaan tanaman hias
di Kecamatan Marpoyan Damai
Kota Pekanbaru, Riau.

No	Y Permintaan Tanaman Hias	X Harga Tanaman Hias (Rp)	Z Pendapatan responden/tahun (Rp)
1	24	2.400.000	35.400.000
2	24	1.920.000	30.600.000
3	48	6.000.000	36.000.000
4	24	1.800.000	40.200.000
5	36	4.800.000	36.000.000
6	12	1.200.000	18.600.000
7	24	2.130.000	26.400.000
8	36	3.750.000	37.800.000
9	12	1.260.000	15.600.000
10	12	420.000	18.000.000
11	12	306.000	24.600.000
12	12	600.000	28.200.000
13	24	1.800.000	26.400.000
14	12	600.000	21.000.000
15	36	1.500.000	33.600.000
16	36	2.700.000	48.000.000
17	24	600.000	19.800.000
18	12	420.000	22.800.000
19	12	1.020.000	29.400.000
20	12	540.000	19.200.000
21	12	720.000	18.600.000
22	24	1.920.000	30.000.000

23	36	4.500.000	49.800.000
24	36	4.200.000	36.600.000
25	24	1.200.000	33.000.000
26	12	420.000	30.000.000
27	12	540.000	20.400.000
28	12	360.000	20.400.000
29	12	1.020.000	28.800.000
30	12	1.200.000	25.200.000
31	36	3.600.000	46.800.000
32	12	360.000	19.800.000
33	24	1.980.000	31.200.000
34	12	480.000	29.340.000
35	12	960.000	21.600.000
36	12	600.000	33.360.000
37	12	420.000	23.400.000
38	24	1.800.000	32.400.000
39	12	900.000	21.600.000
40	12	900.000	22.500.000
41	12	810.000	15.600.000
42	36	2.400.000	51.600.000
43	12	480.000	24.000.000
44	24	1.200.000	25.800.000
45	24	1.200.000	28.800.000
46	12	1.200.000	16.800.000
47	24	2.040.000	25.200.000
48	24	2.760.000	29.400.000
49	12	360.000	19.800.000
50	24	3.060.000	33.600.000
51	24	3.600.000	26.400.000
52	36	4.284.000	29.400.000
53	24	1.080.000	21.600.000
54	12	450.000	21.600.000
55	36	3.600.000	49.380.000
56	12	420.000	22.800.000
57	36	2.700.000	32.400.000
58	48	5.700.000	50.700.000
59	12	1.200.000	14.400.000
60	24	1.500.000	31.200.000
61	12	1.200.000	25.800.000
62	24	1.200.000	28.200.000
63	12	300.000	19.800.000

64	36	1.800.000	38.400.000
65	24	3.000.000	25.200.000
66	24	3.000.000	25.800.000
67	48	3.000.000	67.200.000
68	12	1.260.000	22.800.000
69	36	3.600.000	51.000.000
70	24	1.560.000	24.840.000
71	12	1.680.000	27.000.000
72	36	1.800.000	51600.000
73	36	2.640.000	48.000.000
74	12	780.000	20.400.000
75	24	840.000	19.560.000
76	36	4.560.000	32.100.000
77	24	2.040.000	29.100.000
78	24	1.800.000	23.400.000
79	12	960.000	20.400.000
80	36	2.400.000	42.600.000
81	12	600.000	31.800.000
82	12	540.000	18.600.000
83	12	600.000	15.600.000
84	24	2.580.000	26.700.000
85	24	1.800.000	34.200.000
86	12	1.200.000	26.400.000
87	24	3.000.000	30.000.000
88	12	600.000	17.400.000
89	24	1800.000	30.000.000
90	36	4.500.000	36.000.000
91	24	1.800.000	35.400.000
92	12	600000	18.000.000
93	36	3.600.000	37.200.000
94	12	540.000	22.200.000
95	12	540.000	35.700.000
96	12	480.000	18480.000
97	24	1.800.000	32.400.000
98	24	2.220.000	30.600.000
99	12	600.000	24.000.000
100	48	6.600.000	46.200.000
Jumlah	2184	179.310.000	2.922.960.000

Sumber [2].

Untuk menghitung MSE dari masing-masing penaksir rasio regresi linear terlebih dahulu ditentukan nilai yang dibutuhkan. Dengan bantuan Microsoft Excel diperoleh nilai-nilai sebagaimana yang tertera pada Tabel 2.

Tabel 2: Nilai-nilai yang diperlukan untuk membandingkan MSE dari ketiga penaksir dengan $\alpha = 0,6$.

N	100	S_{yx}	12.747.127
\bar{Y}	21,84	S_{yz}	24.835.549
\bar{X}	1.793.100	S_{xz}	$1,97257 \times 10^{12}$
\bar{Z}	29.229.600	ρ_{yx}	0,86347
S_x	1.397.122	ρ_{yz}	0,23542
S_y	10,5665	ρ_{xz}	0,14087
S_z	10.022.132	-	-

Dengan mensubstitusikan nilai-nilai yang diperoleh pada Tabel 2 ke persamaan (4), (5) dan (6), maka diperoleh MSE dari masing-masing penaksir yang dimuat pada Tabel 3.

Tabel 3: Nilai *Mean Square Error* dari data jumlah permintaan tanaman hias di Kecamatan Marpoyan Damai Kota Pekanbaru, Riau.

Penaksir Rasio	MSE
$\hat{Y}_{RR}^{(1)}$	0,566804
$\hat{Y}_{RR}^{(2)}$	0,452613
$\hat{Y}_{RR}^{(3)}$	0,668685

Dari Tabel 3 dapat dilihat bahwa MSE untuk penaksir $\hat{Y}_{RR}^{(2)}$ merupakan MSE yang minimum, sehingga dapat dikatakan bahwa $\hat{Y}_{RR}^{(2)}$ paling efisien dari penaksir $\hat{Y}_{RR}^{(1)}$ dan penaksir $\hat{Y}_{RR}^{(3)}$.

5. KESIMPULAN

Berdasarkan pembahasan efisiensi yang telah dikemukakan maka dapat disimpulkan bahwa penaksir rasio regresi linear $\hat{Y}_{RR}^{(2)}$ relatif efisien dari penaksir $\hat{Y}_{RR}^{(1)}$,

penaksir rasio regresi linear $\hat{Y}_{RR}^{(1)}$ relatif efisien dari penaksir $\hat{Y}_{RR}^{(3)}$, dan penaksir rasio regresi linear $\hat{Y}_{RR}^{(2)}$ relatif efisien dari penaksir $\hat{Y}_{RR}^{(3)}$.

DAFTAR PUSTAKA

- [1] Cochran, W.G. 1977. *Sampling Techniques*, Third Edition. John Wiley & Sons, New York.
- [2] Sinaga, Ruth. 2011. Analisis Permintaan Masyarakat Terhadap Tanaman Hias di Kecamatan Marpoyan Damai Kota Pekanbaru. Skripsi Fakultas Ekonomi Universitas Riau, Pekanbaru
- [3] Singh, H.P., L.N. Upadhyaya, & Premchandra. 2009. An Improved Version of Regression Ratio Estimator With Two Auxiliary Variables in Sample Surveys, *Statistics In Transition*, 10: 85-100.