

PENAKSIR RASIO REGRESI UNTUK RATA-RATA POPULASI MENGUNAKAN MINIMUM DAN MAKSIMUM VARIABEL TAMBAHAN

Iis Novia^{1*}, Firdaus², Haposan Sirait²

¹ Mahasiswa Program S1 Matematika

² Dosen Jurusan Matematika

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

* iis_novia13@yahoo.com

ABSTRACT

This article studies three type of regression ratio estimators for the population mean on simple random sampling using minimum and maximum auxiliary variables. This study is a review from the article of Singh *et. al.* [*Statistics in Transition*, 10 (2009): 85-100]. Then, the mean square error of this three estimators are compared. Estimator with the smallest mean square error is the most efficient estimator. An example is given at the end of the discussion.

Keywords: *Biased estimator, linear regresi, mean square error, ratio estimator, simple random sampling.*

ABSTRAK

Artikel ini membahas tiga tipe penaksir rasio regresi untuk rata-rata populasi pada sampling acak sederhana menggunakan minimum dan maksimum variabel tambahan. Study ini merupakan review dari artikel Singh *et. al.* [*Statistics in Transition*, 10 (2009): 85-100]. Selanjutnya, *mean square error* dari tiga penaksir ini dibandingkan. Penaksir dengan *MSE* terkecil merupakan penaksir yang efisien. Contoh numerik diberikan pada akhir pembahasan.

Kata kunci: *Mean square error*, penaksir bias, penaksir rasio, regresi linier, sampling acak sederhana.

1. PENDAHULUAN

Dalam statistika, data yang telah diperoleh dari penelitian dijadikan sebagai dasar untuk membuat kesimpulan. Cara memperoleh data bisa dengan cara sensus yaitu mengamati semua anggota populasi, tapi untuk populasi berukuran besar cara ini tidak ekonomis, baik dalam hal biaya maupun waktu, dengan demikian penelitian dilakukan berdasarkan sampel. Sampel harus mencerminkan semua unsur dalam populasi secara proposional yang disebut dengan sampel tak bias atau sampel yang representatif [1].

Untuk meningkat ketelitian penaksir, ada beberapa metode yang digunakan diantaranya metode rasio, regresi dan gabungan rasio regresi. Metode rasio yaitu memanfaatkan hubungan rasio antara y_i dan x_i . y_i adalah unit dari populasi Y dan x_i adalah unit dari populasi X . Asumsikan rata-rata populasi X dari x_i

diketahui. Penaksir regresi memanfaatkan hubungan X yang mempunyai regresi dengan Y [1].

Misalkan bahwa informasi pada dua variabel tambahan X dan Z tersedia. 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 berkorelasi dengan Y . Asumsikan bahwa \bar{X} dan \bar{Z} adalah rata-rata populasi dari variabel tambahan dengan X dan Z diketahui. Terdapat korelasi antara Y dengan X , Y dengan Z , tapi X dengan Z bisa terdapat korelasi atau tidak [3].

2. SAMPLING ACAK SEDERHANA

Sampling acak sederhana yaitu suatu metode untuk mengambil n unit dari populasi berukuran N . Setiap unit dari populasi mempunyai kesempatan yang sama untuk terambil menjadi anggota sampel. Pengambilan sampel dapat dilakukan dengan pengembalian atau tanpa pengembalian. Pengambilan sampel dengan pengembalian menyatakan satuan unit yang telah diambil, memiliki kemungkinan terambil kembali menjadi anggota sampel. Pengambilan sampel tanpa pengembalian menyatakan satuan unit yang telah diambil tidak mungkin terambil kembali menjadi anggota sampel, sehingga hasil yang diperoleh lebih akurat [1].

Misalkan sampel berukuran n unit, kesempatan tiap unit terambil menjadi anggota sampel adalah n/N . Probabilitas suatu unit akan terambil menjadi sampel pada pengambilan pertama adalah n/N , pada pengambilan kedua adalah $(n-1)/(N-1)$ dan seterusnya. Probabilitas seluruh n unit tertentu yang terambil dalam n pengambilan adalah $(C_n^N)^{-1}$.

Untuk menentukan bias dan MSE suatu penaksir, diperlukan beberapa definisi dan teorema sebagai berikut

Teorema 2.1 [1: h.27] Variansi dari rata-rata sampel \bar{y} untuk sampling acak sederhana tanpa pengembalian adalah

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

dengan

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

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

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

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

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

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

3. BIAS DAN *MSE* PENAKSIR RASIO REGRESI UNTUK RATA-RATA POPULASI MENGGUNAKAN MINIMUM DAN MAKSIMUM VARIABEL TAMBAHAN

1. Bentuk dari penaksir rasio regresi untuk rata-rata populasi menggunakan minimum variabel tambahan adalah

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

Bias dan *MSE* penaksir pada persamaan (1) yaitu

$$B\left(\hat{Y}_{RR}^{(1)}\right) \approx \frac{1-f}{n} 2\bar{Y} \left(d^{*2} C_z^2 + d^* \rho_{yx} \rho_{yz} C_y C_z - d^* \rho_{yz} C_y C_z\right),$$

$$MSE\left(\hat{Y}_{RR}^{(1)}\right) \approx \frac{1-f}{n} \bar{Y}^2 \left(C_y^2 - \rho_{xy}^2 C_y^2 + d^{*2} C_z^2 + 2d^* \rho_{yz} C_y C_z - 2d^* \rho_{yz} \rho_{xz} C_y C_z\right), \quad (2)$$

dengan

$$d^* = \frac{\bar{Z}}{(\bar{Z} + Z_m)}.$$

2. Bentuk dari penaksir rasio regresi untuk rata-rata populasi menggunakan minimum variabel tambahan dengan pangkat $-\alpha$ adalah

$$\left(\hat{Y}_{RR}^{(2)}\right) = (\bar{y} + b_{yx}(\bar{X} - \bar{x})) \left(\frac{\bar{z} + Z_m}{\bar{Z} + Z_m}\right)^{-\alpha}. \quad (3)$$

Bias dan *MSE* penaksir pada persamaan (3) yaitu

$$B\left(\hat{Y}_{RR}^{(2)}\right) \approx \frac{1-f}{n} \bar{Y} \left((\alpha^2 + \alpha) d^{*2} C_z^2 + 2\alpha d^* \rho_{yx} \rho_{yz} C_y C_z - 2\alpha d^* \rho_{yz} C_y C_z\right),$$

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

3. Bentuk dari penaksir rasio regresi untuk rata-rata populasi menggunakan maksimum variabel tambahan adalah

$$\left(\hat{Y}_{RR}^{(3)}\right) = (\bar{y} + b_{yx}(\bar{X} - \bar{x})) \left(\frac{\bar{Z} + Z_M}{\bar{z} + Z_M}\right). \quad (5)$$

Bias dan *MSE* penaksir pada persamaan (5) yaitu

$$B\left(\hat{Y}_{RR}^{(3)}\right) \approx \frac{1-f}{n} 2\bar{Y} \left(d_1^2 C_z^2 + d_1 \rho_{yx} \rho_{yz} C_y C_z - d_1 \rho_{yz} C_y C_z\right),$$

$$MSE\left(\hat{Y}_{RR}^{(3)}\right) \approx \frac{1-f}{n} \bar{Y}^2 \left(C_y^2 - \rho_{xy}^2 C_y^2 + \alpha^2 d_1^2 C_z^2 + 2\alpha d_1 \rho_{yz} C_y C_z - 2\alpha d_1 \rho_{yz} \rho_{xz} C_y C_z\right), \quad (6)$$

dengan

$$d_1 = \frac{\bar{Z}}{(\bar{Z} + Z_M)}.$$

4. PENAKSIR RASIO REGRESI YANG EFISIEN

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

1. 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

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

2. 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

$$\rho_{yz} > \frac{2\rho_{xy}\rho_{xz}C_y - (d^* + d_1)C_z}{2C_y} \quad (8)$$

3. 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

$$\rho_{yz} > \frac{2(\alpha d^* - d_1)\rho_{xy}\rho_{xz}C_y - (\alpha d^* - d_1)(\alpha d^* + d_1)C_z}{2(\alpha d^* - d_1)C_y} \text{ ketika } -1 < (\alpha d^* - d_1) < 0,$$

atau

$$\rho_{yz} < \frac{2(\alpha d^* - d_1)\rho_{xy}\rho_{xz}C_y - (\alpha d^* - d_1)(\alpha d^* + d_1)C_z}{2(\alpha d^* - d_1)C_y} \text{ ketika } 1 < (\alpha d^* - d_1) < 0. \quad (9)$$

Contoh Diketahui data pada Tabel 1 mengenai data pengeluaran, pendapatan dan tanggungan karyawan PT. Perkebunan Nusantara V Pekanbaru Tahun 2006. Misalkan data yang diketahui merupakan populasi, maka diambil sampel untuk menaksir rata-rata pengeluaran karyawan menggunakan syarat relatif efisiensi penaksir.

Tabel 1: Data Pengeluaran, Pendapatan dan Tanggungan Karyawan PT. Perkebunan Nusantara V Pekanbaru Tahun 2006.

No.	Nama	Pengeluaran (Rp)	Pendapatan (Rp)	Tanggungan (Jiwa)
1	Suradi	800.000	1.000.000	1
2	Erwin Saputra	900.000	1.000.000	1
3	Legiman	1.500.000	1.800.000	2
4	Iman	1.300.000	1.400.000	2
5	A. Khalidi	1.100.000	1.200.000	3
6	Candra	1.100.000	1.300.000	2

7	Arifin	1.200.000	1.400.000	3
8	Adi Sasono	1.400.000	1.500.000	1
9	Nirwan	1.300.000	1.600.000	2
10	Fadli	1.400.000	1.800.000	4
11	Rudi Ismanto	1.800.000	1.900.000	3
12	Heri Darma	1.500.000	2.000.000	4
13	Ridwan	1.700.000	2.000.000	3
14	Iwan	1.800.000	2.100.000	3
15	Edi Agus	2.000.000	2.100.000	2
16	Kelber	2.000.000	2.200.000	5
17	Rizaldy	2.000.000	2.200.000	2
18	Suyadi	2.000.000	2.300.000	1
19	Yohanes	2.100.000	2.300.000	3
20	Carles Purba	2.300.000	2.400.000	1
21	Arianto	2.000.000	2.400.000	1
22	Sugiarto	2.300.000	2.500.000	4
23	J. Hutajulu	2.500.000	2.600.000	2
24	Bambang	2.700.000	2.800.000	3
25	Laura Muslim	2.400.000	2.800.000	2
26	Heriawan	2.600.000	2.900.000	3
27	Rajad	2.500.000	3.000.000	2
28	Suprpto	2.800.000	3.200.000	6
29	Doni	3.000.000	3.400.000	2
30	N. Gultom	3.100.000	3.300.000	3
31	Tazul	3.000.000	3.400.000	2
32	Sudirman	3.000.000	3.700.000	3
33	Fernando	3.400.000	3.500.000	2
34	F. Butar-butur	3.300.000	3.500.000	3
35	R. Sianturi	3.500.000	3.600.000	2
36	E. Tarigan	3.400.000	3.700.000	3
37	Edi Suprianti	3.400.000	4.200.000	2
38	Bayu Lesmana	3.700.000	4.100.000	3
39	Krisna Setiawan	4.200.000	4.700.000	4
40	L. M. Silaban	3.800.000	4.900.000	2
41	Sukirman	4.700.000	5.200.000	5
42	Zulkifli	4.800.000	5.300.000	4
43	R. Lubis	5.300.000	5.500.000	2
44	Heri Agusman	5.000.000	5.800.000	2
45	Ade Huraina	5.400.000	5.600.000	4
46	Tuhun Bangun	5.500.000	5.700.000	2
47	Pandapotan	5.700.000	5.800.000	4

48	Ali Azhar	5.600.000	6.100.000	2
49	Abu Bakar	5.700.000	6.000.000	4
50	Sagala	6.400.000	6.100.000	3
51	Rafialdi	6.900.000	7.100.000	4
52	Jati Teguh	6.000.000	6.300.000	4
53	Maniruk	5.900.000	6.000.000	2
54	Kasmaliza	7.100.000	7.700.000	3
55	Sardoltua	6.600.000	7.000.000	4
56	Syahrial Nasution	7.000.000	7.500.000	4
57	Bete Napitopulu	7.500.000	8.000.000	6
58	Romaka Purba	7.200.000	8.000.000	4
Jumlah		200.100.000	218.400.000	165

Sumber: [2]

Untuk mengaplikasikan contoh dalam perbandingan *MSE* dari penaksir (1), (3) dan (5) dengan bantuan Microsoft Excel diperoleh nilai-nilai pada Tabel 2.

Tabel 2: Nilai-nilai yang diperlukan untuk membandingkan *MSE* dari penaksir dengan $\alpha = 0,1$, $Z_m = 1$, dan $Z_M = 6$.

N	58	S_z	1,197
n	25	S_{yx}	3,89754E+12
\bar{Y}	3.450.000	S_{yz}	1004385,965
\bar{X}	3.765.517,241	S_{xz}	1048941,319
\bar{Z}	2,845	ρ_{yx}	0,994
S_x	2015166	ρ_{yz}	0,431
S_y	1946138,777	ρ_{xz}	0,435

Substitusikan nilai-nilai yang diperoleh pada Tabel 2 ke persamaan (2), (4) dan (6), diperoleh nilai penaksir dan *MSE* dari masing- masing penaksir yang dimuat pada Tabel 3.

Tabel 3: Nilai *MSE* untuk ketiga penaksir.

Penaksir	<i>MSE</i>
$\hat{Y}_{RR}^{(1)}$	5,047E+10
$\hat{Y}_{RR}^{(2)}$	3,643E+09
$\hat{Y}_{RR}^{(3)}$	1,609E+10

Berdasarkan Tabel 3 dan kriteria penaksir yang relatif efisien pada persamaan (7), (8) dan (9) diperoleh nilai MSE penaksir $\hat{Y}_{RR}^{(2)} < \hat{Y}_{RR}^{(3)} < \hat{Y}_{RR}^{(1)}$. Selanjutnya, dilakukan perhitungan terhadap penaksir (1), (3) dan (5) untuk menunjukkan bahwa nilai taksiran $\hat{Y}_{RR}^{(2)}$ lebih mendekati nilai rata-rata populasi yang dimuat pada Tabel 4.

Tabel 4: Nilai taksiran masing-masing penaksir.

Penaksir	Nilai taksiran
$\hat{Y}_{RR}^{(1)}$	3.361.832
$\hat{Y}_{RR}^{(2)}$	3.434.204
$\hat{Y}_{RR}^{(3)}$	3.398.675

Pada Tabel 4 dapat dilihat bahwa penaksir $\hat{Y}_{RR}^{(2)}$ lebih mendekati nilai rata-rata populasi dengan nilai taksiran 3.434.204.

KESIMPULAN

Berdasarkan pembahasan yang telah dikemukakan pada artikel ini, penaksir $\hat{Y}_{RR}^{(2)}$ merupakan penaksir yang relatif efisien dari penaksir $\hat{Y}_{RR}^{(1)}$ dan $\hat{Y}_{RR}^{(3)}$.

DAFTAR PUSTAKA

- [1] Cochran, W.G. 1991. *Teknik Penarikan Sampel, Edisi ketiga*. Terj. Dari *Sampling Techniques*, oleh Radiansyah & E.R Osman. Penerbit Universitas Indonesia, Jakarta.
- [2] Sinaga, C. V. D. N. 2007. Pola Konsumsi Karyawan PT. Perkebunan Nusantara V (PTPN V) 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.