

PENAKSIR RASIO REGRESI LINEAR SEDERHANA UNTUK RATA-RATA POPULASI MENGGUNAKAN KARAKTER TAMBAHAN

Astari Rahmadita^{1*}, Harison², Haposan Sirait²

¹ Mahasiswa Program Studi S1 Matematika

² Dosen Jurusan Matematika

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

*astari_rahmadita11@yahoo.com

ABSTRACT

Estimators discussed here are three regression ratio estimators of population mean \bar{Y} using information on two auxiliary variables X and Z under simple random sampling without replacement. They are proposed by Singh, Upadhyaya and Premchandra [4] which is a review of the article "An Improved Version of Regression Ratio Estimator with Two Auxiliary Variables in Sample Surveys." All estimators are biased. The efficient estimator is one with the smallest Mean Square Error (MSE), determined by comparing each type of estimator.

Keywords: *regression ratio estimator, simple random sampling, biased estimator and mean square error.*

ABSTRAK

Penaksir yang dibahas adalah tiga penaksir rasio regresi untuk rata-rata populasi \bar{Y} menggunakan dua karakter tambahan X dan Z pada sampling acak sederhana tanpa pengembalian. Penaksir tersebut diajukan oleh Singh, Upadhyaya dan Premchandra [4] yang merupakan *review* dari artikel "An Improved Version of Regression Ratio Estimator with Two Auxiliary Variables in Sample Surveys." Ketiga penaksir yang diajukan merupakan penaksir bias. Berdasarkan nilai *Mean Square Error (MSE)* masing-masing penaksir, penaksir yang memiliki *MSE* yang terkecil merupakan penaksir yang efisien.

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

1. PENDAHULUAN

Populasi berkarakter Y berukuran N akan diteliti untuk mendapatkan taksiran nilai parameternya. Informasi tambahan yang berhubungan dengan Y diperlukan untuk mendapatkan hasil yang lebih akurat. Informasi tambahan ini harus diketahui besar populasinya yang dinotasikan dengan X dan Z .

Metode yang digunakan untuk menaksir rata-rata populasi pada sampling acak sederhana diantaranya adalah penaksir rasio regresi. Regresi linear merupakan metode statistik yang digunakan untuk mengetahui hubungan antara populasi berkarakter X dan populasi berkarakter Y . Metode lainnya untuk menaksir rata-rata populasi adalah penaksir rasio regresi. Tujuan metode rasio regresi adalah untuk meningkatkan ketelitian penelitian dengan mengambil manfaat hubungan antara penaksir regresi linear Y dan X yang berkorelasi positif terhadap Z .

Artikel ini membahas tiga penaksir rasio regresi menggunakan karakter tambahan yang diajukan oleh Singh, Upadhyaya dan Premcandra [4] yaitu \hat{Y}_{nr1} , \hat{Y}_{nr2} dan \hat{Y}_{nr3} . Berdasarkan ide dari Singh, Upadhyaya dan Premcandra [4] penulis mendetailkan bias dan *Mean Square Error (MSE)* dari masing-masing penaksir. Selanjutnya penulis membandingkan *MSE* penaksir. Penaksir yang efisien untuk penaksir bias adalah penaksir yang memiliki *MSE* terkecil.

2. SAMPLING ACAK SEDERHANA

Pengambilan sampel acak sederhana merupakan suatu metode untuk mengambil n unit sampel dari N unit populasi, sehingga setiap elemen ${}_N C_n$ sampel yang berbeda mempunyai kesempatan yang sama untuk dipilih sebagai unit sampel. Pengambilan sampel ini adalah pengambilan sampel acak tanpa pengembalian agar karakteristik unit-unit lebih akurat. Probabilitas terambilnya 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 ${}_N C_n^{-1}$.

Bias dan *MSE* pada sampling acak sederhana ditentukan dengan menggunakan teorema variansi dan kovariansi.

Teorema 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

$$V(\bar{y}) = \frac{S_y^2}{n} \left(\frac{N-n}{N} \right).$$

Bukti dari Teorema ini dapat dilihat pada [1: h.27].

Teorema 2[1:h.29] Jika y_i, x_i adalah sebuah pasangan yang bervariasi ditetapkan pada unit dalam populasi berukuran N dan \bar{y}, \bar{x} adalah rata-rata dari sampel acak sederhana berukuran n , maka kovariansinya adalah

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

Bukti dari Teorema ini dapat dilihat pada [1:h.29].

3. PENAKSIR REGRESI LINEAR

Bentuk umum model regresi linear sederhana Y atas X sebagai berikut

$$Y = A + BX + e,$$

dengan Y adalah variabel tak bebas, X adalah variabel bebas, A dan B adalah parameter (koefisien regresi), dan e adalah kesalahan penarikan sampel [2: h. 20].

Metode yang digunakan untuk menentukan taksiran dari parameter A dan B adalah metode kuadrat terkecil (*least square method*), yaitu suatu metode penaksir dengan prinsip meminimumkan jumlah kuadrat kesalahan penarikan sampel. Misalkan (x_i, y_i) adalah n pasangan data pengamatan $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

$$y_i = A + Bx_i + e_i, \text{ untuk } i = 1, 2, 3, \dots, n.$$

Jumlah kuadrat kesalahan dapat ditulis

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - A - Bx_i)^2, \quad (1)$$

dengan meminimumkan jumlah kuadrat kesalahan pada persamaan (1) maka didapat taksiran untuk A dan B berturut-turut yaitu

$$a = \bar{y} - b\bar{x},$$

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

Garis regresi linear yang melalui titik pangkal artinya jika $a = 0$, maka

$$\bar{y} = b\bar{x}, \quad (2)$$

jika koefisien regresi b berlaku untuk rata-rata sampel, maka b juga berlaku untuk rata-rata populasi, sehingga

$$\hat{Y} = b\bar{X}. \quad (3)$$

Selisih persamaan (3) dengan persamaan (2) secara aljabar, diperoleh

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

\hat{Y} disebut penaksir regresi linear untuk rata-rata populasi yang dinotasikan dengan \hat{Y}_{LR} dan dirumuskan sebagai berikut

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

4. BIAS DAN MSE PENAKSIR RASIO REGRESI LINEAR UNTUK RATA-RATA POPULASI

Selanjutnya akan dibahas bias dan MSE penaksir rasio regresi linear pada sampling acak sederhana yang diajukan oleh Singh, Upadhyaya dan Premchandra [4].

Pertama, penaksir rasio regresi Singh dan Ruis Espejo [4]

$$\hat{Y}_{nr1} = \left(\bar{y} + b_{yx} (\bar{X} - \bar{x}) \right) \left(\alpha \frac{\bar{Z}}{\bar{z}} + (1 - \alpha) \frac{\bar{z}}{\bar{Z}} \right). \quad (4)$$

Kedua, penaksir rasio regresi Naik dan Gupta (1991 a) [4]

$$\hat{Y}_{nr2} = \left(\bar{y} + b_{yx} (\bar{X} - \bar{x}) \right) \left(\frac{\bar{z}}{\bar{Z} \left(1 + (1 - \alpha) \left(\frac{\bar{z}}{\bar{Z}} - 1 \right) \right)} \right). \quad (5)$$

Ketiga, penaksir rasio regresi Naik dan Gupta (1991 b) [4]

$$\hat{Y}_{nr3} = \left(\bar{y} + b_{yx} (\bar{X} - \bar{x}) \right) \left(\frac{\bar{Z} + p\bar{z} - p\bar{Z}}{\bar{Z} + q\bar{z} - q\bar{Z}} \right), \quad (6)$$

dengan $0 < \alpha < 1, 0 < p < 1$ dan $0 < q < 1$.

Bias dan *MSE* penaksir rasio regresi linear pada sampling acak sederhana yang diajukan oleh Singh, Upadhyaya dan Premcandra [4].

Bias dan *MSE* dari persamaan (4) adalah

$$\begin{aligned} B(\hat{Y}_{nr1}) &\approx \frac{\alpha \bar{Y}}{\bar{Z}^2} S_z^2 + \left(\frac{-2\alpha + 1}{\bar{Z}} \right) (\bar{y}_z - B_{yx} S_{xz}) \\ MSE(\hat{Y}_{nr1}) &\approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + (1-2\alpha)^2 C_z^2 - 2(1-2\alpha) \rho_{yx} \rho_{xz} C_y C_z \\ &\quad + 2(1-2\alpha) \rho_{yz} C_y C_z) \end{aligned} \quad (7)$$

Bias dan *MSE* dari persamaan (5) adalah

$$\begin{aligned} B(\hat{Y}_{nr2}) &\approx \frac{\alpha \bar{Y}}{\bar{Z}^2} S_z^2 + \left(\frac{-2\alpha + 1}{\bar{Z}} \right) (\bar{y}_z - B_{yx} S_{xz}) \\ MSE(\hat{Y}_{nr2}) &\approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + \alpha^2 C_z^2 + 2\alpha \rho_{yx} \rho_{xz} C_y C_z \\ &\quad - 2\alpha \rho_{yz} C_y C_z) \end{aligned} \quad (8)$$

Bias dan *MSE* dari persamaan (6) adalah

$$\begin{aligned} B(\hat{Y}_{nr3}) &\approx \frac{\bar{Y}}{\bar{Z}^2} \alpha (1-q) (\bar{y} - 1) (\bar{z} - q) - 2\alpha q \bar{y}_z + \frac{\alpha (1-q)}{\bar{Z}} (\bar{y}_z - B_{yx} S_{xz}) \\ MSE(\hat{Y}_{nr3}) &\approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + \alpha^2 (1-q)^2 C_z^2 - 2\alpha (1-q) \rho_{yx} \rho_{xz} C_y C_z \\ &\quad + 2\alpha (1-q) \rho_{yz} C_y C_z) \end{aligned} \quad (9)$$

5. PENAKSIR RASIO REGRESI YANG EFISIEN

Penaksir yang efisien dari penaksir bias dapat ditentukan dengan cara membandingkan *MSE* masing-masing penaksir tersebut.

1. Perbandingan antara persamaan (7) dan persamaan (8) diperoleh penaksir \hat{Y}_{nr1} lebih efisien dari penaksir \hat{Y}_{nr2} jika

$$\rho_{yz} < \frac{-3\alpha^2 C_z + 4\alpha C_z + C_z - 2\alpha \rho_{yx} \rho_{xz} C_y + 2\rho_{yx} \rho_{xz} C_y}{2C_y (-\alpha)} \quad (10)$$

2. Perbandingan antara persamaan (7) dan persamaan (9) diperoleh penaksir \hat{Y}_{nr1} lebih efisien dari penaksir \hat{Y}_{nr3} ketika $-2 < (-2\alpha - \alpha(p-q)) < 0$ maka

$$\rho_{yz} > \frac{(-2\alpha - \alpha(p-q))C_z + \alpha^2 (p-q)C_z + 2(-2\alpha - \alpha(p-q))\rho_{yx}\rho_{xz}C_y - 2\alpha (p-q)\rho_{yx}\rho_{xz}C_y}{2C_y (-2\alpha - \alpha(p-q))} \quad (11)$$

dan ketika $0 < (-2\alpha - \alpha(p-q)) < 2$ maka

$$\rho_{yz} < \frac{(-2\alpha - \alpha(p-q))C_z + \alpha^2 (p-q)C_z + 2(-2\alpha - \alpha(p-q))\rho_{yx}\rho_{xz}C_y - 2\alpha (p-q)\rho_{yx}\rho_{xz}C_y}{2C_y (-2\alpha - \alpha(p-q))}$$

3. Perbandingan antara persamaan (8) dan persamaan (9) diperoleh penaksir \hat{Y}_{nr3} lebih efisien dari penaksir \hat{Y}_{nr2} jika

$$\rho_{yz} < \frac{(\alpha^2 C_z + \alpha^2 (p-q)C_z - 2\alpha \rho_{yx} \rho_{xz} C_y - 2\alpha (p-q)\rho_{yx} \rho_{xz} C_y)}{-2\alpha C_y (+p-q)} \quad (12)$$

6. CONTOH

Diberikan data mengenai permintaan tanaman hias (\hat{Y}) di Kecamatan Marpoyan [3] dengan sampel $n = 35$ menggunakan karakter tambahan yaitu harga tanaman hias (\hat{X}) dan pendapatan responden (\hat{Z}). *MSE* dari masing-masing penaksir dihitung dengan menentukan terlebih dahulu nilai yang dibutuhkan dengan bantuan Microsoft Excel dan diperoleh nilai-nilai sebagai berikut dengan $\alpha = 0,5$, $p = 0,4$ dan $q = 0,2$.

$N = 100$	$S_x = 1397122$	$S_{yz} = 24835549$	
$\bar{Y} = 21,84$	$S_y = 10,5665$	$S_{xz} = 1,9725710^{12}$	$\rho_{xz} = 0,14087$
$\bar{X} = 1793100$	$S_z = 10022132$	$\rho_{yx} = 0,86347$	$C_y = 0,483814$
$\bar{Z} = 2922960$	$S_{yx} = 12747127$	$\rho_{yz} = 0,23542$	$C_z = 0,342876$

Untuk menentukan penaksir yang efisien dapat dilihat dari *MSE* masing-masing penaksir yang diberikan pada tabel berikut ini

Tabel 1 : Nilai *Mean Square Error* dari Masing-masing Penaksir.

Penaksir Rasio	<i>MSE</i>
\hat{Y}_{r1}	0,527547
\hat{Y}_{r2}	0,622024
\hat{Y}_{r3}	0.532468

Nilai-nilai yang diperoleh dari permintaan tanaman hias, harga dan pendapatan respondendisubstitusikan ke persamaan (10), (11) dan (12),maka diperoleh

1. $MSE(\hat{Y}_{r1}) < MSE(\hat{Y}_{r2})$ jika $\rho_{yz} < 0.955901$
2. $MSE(\hat{Y}_{r1}) < MSE(\hat{Y}_{r3})$ jika $\rho_{yz} > 0.125069$.
3. $MSE(\hat{Y}_{r3}) < MSE(\hat{Y}_{r2})$ jika $\rho_{yz} < 0.263379$.

Berdasarkan Tabel 1 dan syarat efisiensi ρ_{yz} , maka penaksir \hat{Y}_{r1} lebih efisien dari penaksir \hat{Y}_{r2} dan penaksir \hat{Y}_{r3} .

7. KESIMPULAN

Nilai *MSE* dari penaksir rasio regresi linear yang diperoleh dengan menggunakan dua karakter tambahan untuk rata-rata populasi yang diajukan pada sampling acak sederhana, kemudian membandingkan *MSE* dari masing-masing penaksir, sehingga dapat disimpulkan bahwa untuk data permintaan tanaman hias di Kecamatan Marpoyan, Pekanbaru penaksir rasio regresi \hat{Y}_{r1} lebih efisien dibandingkan dengan penaksir \hat{Y}_{r2} dan penaksir \hat{Y}_{r3} jika syarat efisiensi ρ_{yz} terpenuhi.

DAFTAR PUSTAKA

- [1] Cochran, W.G. 1977. *Sampling Techniques*, Third Edition. John Wiley & Sons, New York.
- [2] Draper, R.N & H, Smith. 1998. *Applied Regression Analysis*, Third Edition. John Wiley & Sons, New York.
- [3] Sinaga, Ruth. 2011. Analisis Permintaan Masyarakat Terhadap Tanaman Hias di Kecamatan Marpoyan Damai Kota Pekanbaru. Skripsi Fakultas Ekonomi Universitas Riau, Pekanbaru.

- [4] 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.