

PENAKSIR RASIO REGRESI LINEAR YANG EFISIEN UNTUK RATA-RATA POPULASI DENGAN MENGGUNAKAN DUA VARIABEL TAMBAHAN

Edi Jamilun^{1*}, Harison², Haposan Sirait²

¹Mahasiswa Program Studi S1 Matematika

²Dosen Jurusan Matematika

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

*edi_jamilun@yahoo.co.id

ABSTRACT

This paper discusses three ratio regression linear estimators by Singh, Upadhyaya and Premchandra for population mean $\bar{Y}$ using two auxiliary variables X and Z on simple random sampling which is a review of article "*An Improved Version of Regression Ratio Estimator with Two auxiliary Variables in Sample Surveys*". The three estimators are biased then the minimum of the (MSE) is determined. Furthermore, their MSE are compared to each other. This comparison shows that the estimator with the minimum MSE is the most efficiensi estimator.

Keywords: *ratio regression estimator, coefficient of variation, coefficient of correlation, biased and mean square error*

ABSTRAK

Artikel ini merupakan tinjauan ulang terhadap tiga penaksir rasio regresi linear yang diajukan oleh Singh, Upadhyaya dan Premchandra untuk rata-rata populasi $\bar{Y}$ dengan menggunakan dua variabel tambahan X dan Z pada sampling acak sederhana yang merupakan review dari artikel "*An Improved Version of Regression Ratio Estimator with Two auxiliary Variables in Sample Surveys*". Ketiga penaksir merupakan penaksir bias dan ditentukan *mean square error*. Selanjutnya, *mean square error* dari masing-masing penaksir dibandingkan. Perbandingan ini menunjukkan bahwa penaksir yang memiliki *mean square error* terkecil atau minimum adalah penaksir yang lebih efisien.

Kata kunci: *penaksir rasio regresi linear, koefisien variasi, koefisien korelasi, bias dan mean square error.*

1. PENDAHULUAN

Penaksir regresi 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

$$\hat{\bar{Y}}_R = \bar{y} \left(\frac{\bar{X}}{\bar{x}} \right),$$

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 rasio regresi linear untuk rata-rata populasi pada sampling acak sederhana dengan menggunakan variabel tambahan x dan z yang diajukan oleh Singh, Upadhyaya dan Premchandra [4], yaitu

$$\bar{y}_{RR}^{(1)} = \bar{y}_{lr} [1 + \alpha(u - 1)]^{-1} \quad (1)$$

$$\bar{y}_{RR}^{(2)} = \bar{y}_{lr} u^{-1} [1 + (1 - \alpha)u^{-1}] \quad (2)$$

$$\bar{y}_{RR}^{(3)} = \bar{y}_{lr} u [\alpha + (1 - \alpha)u] \quad (3)$$

dengan $\bar{y}_{lr} = \bar{y} + b_{yx}(\bar{X} - \bar{x})$, $u = \frac{\bar{z}}{\bar{Z}}$, α konstanta.

Ketiga penaksir rasio regresi linear untuk rata-rata populasi tersebut merupakan penaksir bias, kemudian ditentukan MSE . Berdasarkan ide dari Singh, Upadhyaya dan Premchandra [4], penulis membandingkan MSE dari masing-masing penaksir untuk memperoleh penaksir rasio regresi linear yang efisien. Penaksir yang memiliki nilai MSE terkecil merupakan penaksir yang efisien.

2. 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 $S_y^2 = \sum_{i=1}^N (y_i - \bar{Y})^2 / N - 1$ 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

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

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 tiga variabel.

3. PENAKSIR REGRESI UNTUK RATA-RATA POPULASI

Bentuk umum model regresi linear sederhana adalah

$$\bar{Y} = \alpha + \beta X + e, \quad (4)$$

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 (4) dapat ditulis

$$y_i = \alpha + \beta x_i + e_i, \quad i = 1, 2, \dots, n. \quad (5)$$

Diasumsikan $E(e_i) = 0$

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

Dari persamaan (5) maka jumlah kuadrat kesalahan pengamatan data terhadap garis regresi ditulis

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2. \quad (6)$$

Dengan mengganti α dan β pada persamaan (6) 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

$$a = \bar{y} - b\bar{x}. \quad (7)$$

Ketika garis regresi linear melalui titik pangkal $a = 0$, maka persamaan (7) menjadi

$$\bar{y} = b\bar{x}, \quad (8)$$

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

$$\hat{\bar{Y}} = b\bar{X}. \quad (9)$$

Dari pengurangan persamaan (8) dengan persamaan (9) secara aljabar, diperoleh

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

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

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

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

Bias dan *MSE* penaksir rasio regresi linear yang diajukan oleh Singh, Upadhyaya dan Premchandra [4] untuk rata-rata populasi pada sampling acak sederhana dari masing-masing penaksir sebagai berikut

Bias dan *MSE* dari persamaan (1)

diperoleh

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

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

dengan ρ adalah koefisien korelasi dan α konstan.

Bias dan *MSE* dari persamaan (2) diperoleh

$$B(\bar{y}_{RR}^{(2)}) \approx 4\bar{Y} - \frac{1}{\bar{Z}} \rho_{lr} S_y S_z + \bar{Y} \cdot \bar{Z} - \frac{4(1+\alpha)}{\bar{Z}} \bar{Y} - 4\alpha \bar{Y}$$

$$MSE(\bar{y}_{RR}^{(2)}) \approx \frac{1-f}{n} (C_y^2 \alpha^2 - 4C_y^2 \alpha + 4C_z^2 - \rho_{yx}^2 C_y^2 \alpha^2 + 4\rho_{yx}^2 C_y^2 \alpha - 4\rho_{yx}^2 C_y^2$$

$$+ 4C_z^2 \alpha^2 - 12C_z^2 \alpha + 9C_z^2 - 4\rho_{yz} C_y C_z \alpha^2 + 14\rho_{yz} C_y C_z \alpha - 12\rho_{yz} C_y C_z$$

$$+ 4\rho_{yx} \rho_{xz} C_z C_y \alpha^2 - 14\rho_{yx} \rho_{xz} C_z C_y \alpha - 12\rho_{yx} \rho_{xz} C_z C_y)$$

Bias dan *MSE* dari persamaan (3) diperoleh

$$B(\bar{y}_{RR}^{(3)}) \approx \frac{\alpha}{\bar{Z}} \rho_{lr} S_y S_z + \frac{1-\alpha}{\bar{Z}^2}$$

$$MSE(\bar{y}_{RR}^{(3)}) \approx \frac{1-f}{n} \bar{Y}^2 (C_y^2 - \rho_{yx}^2 C_y^2 + C_z^2 \alpha^2 - 4C_z^2 \alpha + 4C_z^2 - 2\rho_{yz} C_y C_z \alpha$$

$$+ 4\rho_{yz} C_y C_z + 2\rho_{yx} \rho_{xz} C_z C_y \alpha - 4\rho_{yx} \rho_{xz} C_z C_y)$$

5. PENAKSIR RASIO REGRESI LINEAR 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 rasio regresi linear $\bar{Y}_{RR}^{(1)}$ dengan penaksir rasio regresi linear $\bar{Y}_{RR}^{(2)}$ diperoleh $MSE(\bar{Y}_{RR}^{(1)}) < MSE(\bar{Y}_{RR}^{(2)})$

a. Jika $-3 < \alpha < -1$,

$$\rho_{yz} < \frac{C_y^2(\alpha^2 - 4\alpha + 3) + \rho_{yx}^2 C_y^2(\alpha^2 - 4\alpha - 3) + C_z^2(3\alpha^2 + 12\alpha - 9) - \rho_{yx}\rho_{xz}C_yC_z(4\alpha^2 + 16\alpha + 12)}{C_yC_z(-4\alpha^2 - 16\alpha - 12)}$$

b. Jika $-3 > \alpha > -1$,

$$\rho_{yz} > \frac{C_y^2(\alpha^2 - 4\alpha + 3) + \rho_{yx}^2 C_y^2(\alpha^2 - 4\alpha - 3) + C_z^2(3\alpha^2 + 12\alpha - 9) - \rho_{yx}\rho_{xz}C_yC_z(4\alpha^2 + 16\alpha + 12)}{C_yC_z(-4\alpha^2 - 16\alpha - 12)}$$

2. Perbandingan antara penaksir rasio regresi linear $\bar{Y}_{RR}^{(1)}$ dengan penaksir rasio regresi linear $\bar{Y}_{RR}^{(3)}$ diperoleh $MSE(\bar{Y}_{RR}^{(3)}) < MSE(\bar{Y}_{RR}^{(1)})$ jika

$$\rho_{yz} > \frac{C_z\alpha - C_z + \rho_{yx}\rho_{xz}C_y}{C_y}$$

3. Perbandingan antara penaksir rasio regresi linear $\bar{Y}_{RR}^{(2)}$ dengan penaksir rasio regresi linear $\bar{Y}_{RR}^{(3)}$ diperoleh $MSE(\bar{Y}_{RR}^{(3)}) < MSE(\bar{Y}_{RR}^{(2)})$ jika $\forall \alpha \in R$

$$\rho_{yz} > \frac{-C_y^2(\alpha^2 - 4\alpha + 3) + \rho_{yx}^2 C_y^2(\alpha^2 - 4\alpha + 3) - C_z^2(3\alpha^2 - 8\alpha + 5)}{(C_yC_z(-4\alpha^2 + 6\alpha - 16))} - \frac{\rho_{yx}\rho_{xz}C_yC_z(4\alpha^2 - 16\alpha + 16)}{(C_yC_z(-4\alpha^2 + 6\alpha - 16))}$$

6. CONTOH

Perhatikan data tentang rata-rata berat ikan, panjang ikan dan berat telur ikan bunga air (*cupeichthys bleekeri*) dari Desrita [2]. Penelitian dilakukan di Hulu sungai Kampar Kanan kecamatan XIII koto kampar provinsi Riau di Stasiun Muara Takus 2006. Data panjang ikan, berat ikan dan berat telur ikan. Dengan menggunakan data tersebut akan ditentukan penaksir rasio regresi linear yang efisien untuk menaksir rata-rata berat ikan 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 berat ikan digunakan panjang ikan (x) dan berat telur ikan (z). Untuk menghitung MSE dari masing-masing penaksir terlebih dahulu ditentukan nilai yang dibutuhkan. Informasi yang diperoleh dari berat ikan, panjang ikan dan berat telur ikan dengan menggunakan *Microsoft Excel*, yaitu

$\bar{Y} = 0,5367$	$C_y = 0,176$	$\rho_{yx} = 0,7902$	$N = 84$
$\bar{X} = 39,3214$	$C_x = 0,0503$	$\rho_{yz} = 0,3240$	$n = 20$
$\bar{Z} = 0,00557$	$C_z = 1,0504$	$\rho_{xz} = 0,4464$	$\alpha = 0,5$

Dengan menggunakan informasi sebelumnya, diperoleh bahwa

(i) $MSE(\bar{Y}_{RR}^{(1)}) < MSE(\bar{Y}_{RR}^{(2)})$ jika $\rho_{yz} > -1,0026$.

(ii) $MSE(\bar{Y}_{RR}^{(3)}) < MSE(\bar{Y}_{RR}^{(1)})$ jika $\rho_{yz} < 8,8146$.

(iii) $MSE(\bar{Y}_{RR}^{(3)}) < MSE(\bar{Y}_{RR}^{(2)})$ jika $\rho_{yz} < 0,9745$ dengan $\alpha > 0$

Selanjutnya nilai MSE dari masing-masing penaksir diberikan pada Tabel 1.

Tabel 1. Nilai MSE untuk ketiga penaksir

Penaksir	MSE
$\bar{Y}_{RR}^{(1)}$	0,003214
$\bar{Y}_{RR}^{(2)}$	0,029088
$\bar{Y}_{RR}^{(3)}$	0,000307

Berdasarkan Tabel 1, dapat dilihat bahwa penaksir rasio $\bar{Y}_{RR}^{(3)}$ memiliki nilai MSE yang terkecil dengan syarat bahwa kondisi lebih efisien dapat terpenuhi.

7. KESIMPULAN

Dari pembahasan di atas, dapat disimpulkan bahwa penaksir rasio regresi linear $\bar{Y}_{RR}^{(3)}$ merupakan penaksir yang paling efisien dari penaksir rasio regresi linear $\bar{Y}_{RR}^{(1)}$ dan penaksir rasio regresi linear $\bar{Y}_{RR}^{(2)}$ jika syarat efisien terpenuhi.

DAFTAR PUSTAKA

- [1] Cochran, W. G. 1991. *Teknik Penarikan Sampel, Edisi Ketiga*. Terj. Dari *Sampling Techniques*, oleh Rudiansyah & E. R Osman. UI Press, Jakarta.
- [2] Desrita. 2006. Aspek biologi ikan bunga air di Hulu Sungai Kampar kanan Kecamatan XIII Koto Kampar. Skripsi Fakultas Perikanan Universitas Riau, Pekanbaru.
- [3] Myers, R. H. 1986. *Classical and Modern Regression with Applications*. Wadsworth, Inc, America.
- [4] Singh, H.P, Upadhyaya, L.N, Premchandra. An Improved Version of Regression Ratio Estimator With Two Auxiliary Variables in Sample Surveys, *Statistics In Transition*, vol. 10, pp 85-100, 2009.