MODIFIKASI PENAKSIR
UNTUK RASIO PADA SAMPLING BERPERINGKAT

Devani Erwin[1], **Arisman Adnan**[2], **Rustam Efendi**[2]

Devanierwin@yahoo.com

[1]Mahasiswa Program S1 Matematika
[2]Dosen Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Kampus Binawidya Pekanbaru,28293,Indonesia

ABSTRACT

Ratio estimator discussed are two proposed by Samawi Muttlak [4], for ratio estimator rank set sampling (RSS) and ratio estimator using regression coefficients has been reviewed, each estimator is a biased estimator. so as to determine an efficient estimator by comparing Mean Square Error (MSE) of each estimator.

Key words: *Ranked Set Sampling, regression coefficients, biased, Means Square Error*

1. PENDAHULUAN

Penaksir untuk rasio pada sampling berperingkat didefinisikan dengan

$$\hat{R}_{rss} = \frac{\bar{y}_{(i)}}{\bar{x}_{(i)}} \tag{1}$$

penaksir rasio (1) menjadi modifikasi penaksir untuk rasio pada sampling berperingkat

$$\hat{R}_{mrss} = \frac{\bar{y}_{i(n)} + b\left(\bar{X} - x_{i(n)}\right)}{\bar{x}_{i(n)}} \tag{2}$$

Kedua penaksir untuk rasio ini merupakan penaksir bias. Untuk mendapatkan penaksir rasio yang efisien adalah dengan membandingkan *MSE* untuk masing-masing penaksir dengan kriteria bahwa semakin kecil *MSE* yang diperoleh maka penaksir semakin efisien.

2. SAMPLING ACAK SEDERHANA

Sampling acak sederhana adalah suatu metode untuk mengambil n unit dari populasi berukuran N, dimana setiap elemen mempunyai kesempatan yang sama untuk diambil menjadi anggota sampel. Pengambilan sampel dapat dilakukan dengan pengembalian atau tanpa pengembalian.

Pada sampling acak sederhana tanpa pengembalian, banyaknya sampel yang akan terbentuk adalah C_n^N. Probabilitas suatu unit akan terpilih menjadi sampel pada pengambilan pertama adalah n/N, pada pengambilan kedua adalah $(n-1)/(N-1)$ dan seterusnya, maka probabilitas seluruh n unit-unit tertentu yang terpilih dalam n pengambilan adalah $(C_n^N)^{-1}$.

Teorema 2.2 [1 :h. 28] Variansi dari rata-rata $(\bar{y})$ untuk sampel acak sederhana adalah

$$V(\bar{y}) = \frac{S^2}{n}\frac{N-n}{N} = \frac{S^2}{n}(1-f) \tag{3}$$

dengan $f = \dfrac{n}{N}$ adalah fraksi penarikan sampel.

Bukti : Bukti dari teorema ini dapat dilihat pada **[1 :h. 28]**

Teorema 2.3 [1:h. 29] Jika y_i, x_i adalah sebuah pasangan yang bervariasi pada unit dalam populasi dan $\bar{y}, \bar{x}$ adalah rata-rata dari sampel acak sederhana berukuran n, maka kovariansinya adalah

$$Cov(\bar{y}, \bar{x}) = \frac{1-f}{n}\frac{1}{N-1}\sum_{i=1}^{n}(y_i - \bar{Y})(x_i - \bar{X})\,. \tag{4}$$

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

3. TAKSIRAN PARAMETER PADA MODEL REGRESI LINEAR SEDERHANA

Dalam beberapa penelitian sering kali ingin mengetahui hubungan antara satu variabel dengan beberapa variabel lainnya yang disebut dengan regresi. Salah satu model regresi yang digunakan adalah model regresi linear. Regresi linear adalah suatu persamaan yang menyatakan hubungan antara variabel tak bebas y_i dari variabel bebas x_i dalam bentuk persamaan linear. Model regresi linear sederhana adalah

$$y = \alpha + \beta x + e. \tag{5}$$

dengan α dan β adalah parameter yang akan ditaksir atau disebut dengan koefisien regresi, dan e adalah kesalahan pengamatan.

Salah satu metode yang digunakan untuk menentukan taksiran dari parameter α dan β pada model regresi linear sederhana adalah metode kuadrat terkecil, yaitu suatu metode penaksir dengan prinsip meminimumkan jumlah kuadrat kesalahan pengamatan.

Teorema 2.4 [1 :h. 42] Misalkan $y = \alpha + \beta x + e$ merupakan penaksir tak bias untuk model regresi linear sederhana, dengan metode kuadrat terkecil diperoleh penaksir dari α dan β yaitu:

$$\alpha = \frac{\sum_{i=1}^{n} y_i}{n} - \beta \frac{\sum_{i=1}^{n} x_i}{n}, \tag{6}$$

dan

$$\beta = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}, \tag{7}$$

nilai taksiran b untuk β pada Teorema 2.4 dapat juga ditulis dengan

$$b = \frac{s_{xy}}{s_x^2}$$

dalam bentuk sampling berperingkat dapat ditulis

$$b = \frac{s_{xi[n]yi[n]}}{s_{xi[n]}^2} \tag{8}$$

dengan

$s_{xi(n)}^2$ menyatakan variansi sampel berkarakter x.

sedangkan $s_{xi(n)yi(n)}$ menyatakan kovariansi antara variabel x dan y.

4. SAMPLING BERPERINGKAT

Sampling berperingkat pertama kali ditemukan oleh McIntyre [3]. McIntyre menggunakan sampling berperingkat sebagai pengganti dari sampling acak sederhana untuk menaksir rata-rata populasi.

Penarikan sampel untuk sampling berperingkat adalah suatu metode untuk mengambil n unit sampel dari populasi yang berukuran N, dimana setiap elemen memiliki kesempatan yang sama untuk terpilih menjadi anggota sampel. Sampel berukuran n tersebut dibagi kedalam sejumlah m bagian dengan ukuran yang sama, kemudian unit terkecil dipilih dari bagian pertama dan unit terkecil ke dua dipilih dari bagian ke dua, prosedur ini berkelanjutkan hingga unit dengan tingkatan terbesar terpilih. Siklus ini diulangi sebanyak r kali, sehingga unit mr akan diukur selama proses sampling berperingkat.

$$\bar{Y}_{i[n]} = \frac{\sum_{i-1}^{N} y_{(i)}}{N}\ ,\ \bar{y}_{rss} = \frac{\sum_{i-1}^{n} y_{(i)}}{n}\ ,\ S_y^2 = \frac{\sum_{i=1}^{N}\left(y_{(i)} - \bar{Y}_{i[n]}\right)^2}{N}$$

Teorema 2.2.1 Apabila sampling yang digunakan merupakan sampling berperingkat, maka variansi dari rata-rata sampel sampling berperingkat $\left(\bar{y}_{rss}\right)$adalah

$$Var(\bar{y}_{RSS}) = \frac{1}{n}\left(S_y^2 - \frac{1}{n}\sum_{i=1}^{n} S_{[i]}^2\right) \tag{9}$$

Bukti:

$$Var(\bar{y}_{rss}) = \mathrm{var}\left(\frac{1}{n}\sum_{i=1}^{n} y_{(i)}\right)$$

$$= \frac{1}{n^2}\left(\mathrm{var}\sum_{i=1}^{n} y_{(i)}\right)$$

$$= \frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\left(E\left(y_{(i)} - \bar{Y}_{i[n]}\right)^2\right)\right)$$

$$Var(\bar{y}_{rss}) = \frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{i=1}^{N} P\left(y_{(i)} - \bar{Y}_{i[n]}\right)\left(y_{(i)} - \bar{Y}_{i[n]}\right)\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\left(y_{i[n]} - \bar{Y}_{i[n]}\right)^2\right)\right)$$

atau

$$Var(\bar{y}_{rss}) = \frac{1}{n}\left(\frac{1}{N}\sum_{i=1}^{N}\left(y_{(i)} - \bar{Y}_{i[n]}\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\left(y_{i[n]} - \bar{Y}_{i[n]}\right)^2\right) \right)$$

$$Var(\bar{y}_{rss}) = \frac{1}{n}\left(S_y^2 - \frac{1}{n}\sum_{i=1}^{n} S_{[i]}^2 \right).$$

Teorema 2.2.2 Adanya variabel tambahan *X* yang dinotasikan $x_{[i]}$, jika $(y_{[i]}, x_{[i]})$ adalah sebuah pasangan yang bervariasi ditetapkan pada unit dalam populasi dan $\bar{y}_{rss}$, $\bar{x}_{rss}$ adalah rata-rata dari sampling berperingkat berukuran *n*, maka kovariansinya adalah

$$\text{cov}(y_{rss}, x_{rss}) = \frac{1}{n}\left(S_{xy} - \frac{1}{n}\sum_{i=1}^{n} S_{xy[i]} \right) \tag{10}$$

Bukti:

Berdasarkan Defenisi 2.3 maka diperoleh

$$\text{cov}(y_{rss}, x_{rss}) = E\left(\bar{y}_{rss} - \bar{Y}_{i[n]}\right)\left(\bar{x}_{rss} - \bar{X}_{i[n]}\right)$$

Misalkan $u_i = x_i + y_i$ dengan rata-rata sampel u_i adalah $\bar{u}_{rss} = \bar{x}_{rss} + \bar{y}_{rss}$. Rata-rata populasi dari u_i adalah $\bar{U} = \bar{X}_{i[n]} + \bar{Y}_{i[n]}$.

Berdasarkan Teorema 2.2.1 maka diperoleh

$$\text{var}(\bar{u}_{rss}) = E\left(\bar{u}_{rss} - \bar{U}\right)^2 = \frac{1}{n}\left(S_u^2 - \frac{1}{n}\sum_{i=1}^{n} S_{u[i]}^2 \right)$$

dengan

$$S_u^2 = \frac{\sum_{i=1}^{N}\left(u_{[i]} - \bar{U}\right)^2}{N}$$

maka diperoleh

$$E\left(\bar{u}_{rss}-\bar{U}\right)^2=\frac{1}{n}\left(\frac{1}{N}\sum_{i=1}^{N}\left(u_{[i]}-\bar{U}\right)^2-\frac{1}{n}\sum_{i=1}^{n}\left(\bar{U}_i-\bar{u}\right)^2\right)$$

atau

$$E\left(\left(\bar{x}_{rss}+\bar{y}_{rss}\right)-\left(\bar{X}_{iI[n]}+\bar{Y}_{i[n]}\right)\right)^2=\frac{1}{n}\left(\left(\frac{\sum_{i=1}^{N}\left(\left(x_{(i)}+y_{(i)}\right)-\left(\bar{X}+\bar{Y}\right)\right)^2}{N}\right)\right.$$

$$\left.-\frac{1}{n}\sum_{i=1}^{n}\left(\left(x_{i[n]}+y_{i[n]}\right)-\left(\bar{X}_{i[n]}+\bar{Y}_{i[n]}\right)\right)^2\right)$$

$$E\left(\left(\bar{y}_{rss}-\bar{Y}_{i[n]}\right)+\left(\bar{x}_{rss}-\bar{X}_{i[n]}\right)\right)^2=\frac{1}{n}\left(\left(\frac{\sum_{i=1}^{N}\left(\left(y_{(i)}-\bar{Y}_{i[n]}\right)+\left(x_{(i)}-\bar{X}_{i[n]}\right)\right)^2}{N}\right)\right.$$

$$\left.-\frac{1}{n}\sum_{i=1}^{n}\left(\left(y_{i[n]}-\bar{Y}_{i[n]}\right)+\left(x_{i[n]}-\bar{X}_{i[n]}\right)\right)^2\right)$$

$$E\left(\left(\bar{y}_{rss}-\bar{Y}_{i[n]}\right)\left(\bar{x}_{rss}-\bar{X}_{i[n]}\right)\right)=\frac{1}{n}\left(\left(\frac{\sum_{i=1}^{N}\left(\left(y_{(i)}-\bar{Y}_{i[n]}\right)-\left(x_{(i)}-\bar{X}_{i[n]}\right)\right)}{N}\right)\right.$$

$$\left.-\frac{1}{n}\sum_{i=1}^{n}\left(\left(y_{i[n]}-\bar{Y}_{i[n]}\right)\left(x_{i[n]}-\bar{X}_{i[n]}\right)\right)\right)$$

sehingg a:

$$\operatorname{cov}\left(y_{rss},x_{rss}\right)=\frac{1}{n}\left(S_{xy}-\frac{1}{n}\sum_{i=1}^{n}\left(x_{i[n]}-\bar{X}_{i[n]}\right)\left(y_{i[n]}-\bar{Y}_{i[n]}\right)\right)$$

$$\operatorname{cov}\left(y_{RSS},x_{RSS}\right)=\frac{1}{n}\left(S_{xy}-\frac{1}{n}\sum_{i=1}^{n}S_{xy[i]}\right)$$

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

Bias yang dihasilkan dapat dihitung untuk melihat seberapa besar kesalahan taksiran tersebut. Sebelumnya telah dibuktikan bahwa kedua penaksir yang diajukan merupakan penaksir bias.

Efisiensi dari beberapa penaksir tersebut dapat dilihat dengan membandingkan *MSE* dari setiap penaksir, dengan asumsi bahwa semakin kecil *MSE* dari penaksir maka penaksir tersebut lebih efisien. Bentuk dikemukakan beberapa penaksir untuk rasio yang diajukan oleh Samawi Muttlak [5]. Ditentukan *MSE* dari:

Bias dan *MSE* penaksir untuk rasio pada sampling berperingkat $\hat{R}_{rss}$ adalah

$$B(\hat{R}_{rss}) = \frac{1-f}{n}\left(RC_{x[i]}^2 - RC_{y[i]x[i]}\right) \approx \frac{1-f}{n}\left(C_{x[i]}^2 - C_{y[i]x[i]}\right)R. \tag{11}$$

$$MSE\left(\hat{R}_{rss}\right) \cong \frac{R^2}{n}\left\{C_{x[i]}^2 + C_{y[i]}^2 - 2\rho C_{x[i]}C_{y[i]}\right\}. \tag{12}$$

Bias dan *MSE* modifikasi penaksir untuk rasio pada sampling berperingkat $\hat{R}_{mrss}$ adalah

$$B(\hat{R}_{mrss}) \approx \frac{1-f}{n}\left(RC_{x[i]}^2\right) = \frac{1-f}{n}\left(C_{x[i]}^2\right)R. \tag{13}$$

$$MSE\left(\hat{R}_{mrss}\right) \cong \frac{R^2}{n}\left\{C_{x[i]}^2 + C_{y[i]}^2\left(1-\rho^2\right)\right\}. \tag{14}$$

Selanjutnya akan ditentukan penaksir yang efisien diantara kedua penaksir yang diajukan, yaitu

Penaksir rasio $\hat{R}_{rss}$ dengan penaksir rasio $\hat{R}_{mrss}$.

Diperoleh bahwa penaksir rasio $\hat{R}_{mrss}$ lebih efisien dari penaksir rasio $\hat{R}_{rss}$ jika

$B < B_1$ dan $B > B_2$

atau

$B_2 < B < B_1$

6. CONTOH

Dengan menggunakan data dari Kadilar dan Chingi [2]. Data dari daerah Marmara Turkey. Data berkaitan dengan jumlah dari pohon apel pada 106 desa di daerah Marmara tahun 1999.

Untuk mengilustrasikan bagaimana mengaplikasikan sampling berperingkat pada data ini, dipilih sampel berukuran 36 dari populasi dengan cara sampling acak tanpa pengembalian. Data sampel tersebut dikelompokkan kedalam 3 set masing-masing berukuran 3 ($m = 3$) dan diulang sebanyak 4 kali perlakuan yang sama ($r = 4$). Mengikuti metode sampling berperingkat maka seluruhnya dilakukan sebanyak 12 kali penelitian ($mr = 12$). Untuk menghitung *MSE* dari masing-masing penaksir terlebih dahulu ditentukan nilai yang dibutuhkan. Dengan bantuan Microsoft Excel diperoleh nilai-nilai sebagaimana yang tertera pada Tabel 1.

Tabel 1 Nilai-nilai yang diperlukan untuk membandingkan *MSE* dari kedua penaksir

N	106
n	12
m	3
r	4
B	0.82
ρ	0.82
$\overline{X}$	-6.97×10^{-17}
$\overline{Y}$	1.67×10^{-16}
R	-2.398
$S_{x[i]}$	1
$S_{y[i]}$	1
$C_{x[i]}$	-0.14×10^{-17}
$C_{y[i]}$	6×10^{-17}

Dengan menggunakan informasi pada Tabel 1, diperolehlah bahwa penaksir untuk rasio $\hat{R}_{mrss}$ merupakan penaksir yang paling efisien dibandingkan dengan penaksir untuk rasio $\hat{R}_{rss}$

DAFTAR PUSTAKA

[1] Cochran, W.G. 1991. *Teknik Penarikan Sampel, Edisi ketiga.* Terj. Dari *Sampling Techniques,* oleh Rudiansyah & E.R Osman. Penerbit Universitas Indonesia, Jakarta.

[2] Kadilar, C., Unyazici, Y., Cingi H. 2007. Ratio Estimator for the Population Mean Using Ranked Set Sampling. *Department of Statistics*, *acettepe University*, Ankara, Turkey

[3] Mcintyre, G. A. 1952. A Method for Unbiased Selective Sampling Using Ranked Sets. *Australian Journal*. 3, 385-390.

[4] Montgomery, D.C & G. C. Runger. 1999. *Applied Statistics and Probability for Engineers, Second Edition*. John Wiley & Sons, Inc, New York.

[5] Samawi, Hani M dan Muttlak, Hassen A. 1996. Estimation of Ratio Using Rank Set Sampling. *Akademi Verlag*. Biom J 753-764