

Teknik Mengatasi Data Hilang pada Kasus Rancangan Blok Lengkapacak

Rado Yendra¹, Muslimin²

^{1,2}Jurusan Matematika, Fakultas Sains dan Teknologi, UIN Sultan Syarif Kasim Riau
Jl. HR. Soebrantas No. 155 Simpang Baru, Panam, Pekanbaru, 28293
Email: rado.yendra@uin-suska.ac.id dan muslimin.pb@gmail.com

ABSTRAK

Data hilang pada rancangan blok lengkap acak (RBLA) merupakan informasi yang tidak tersedia untuk sebuah obyek (kasus) tertentu, dimana banyaknya data pengamatan dalam rancangan blok random lengkap kurang dari $k \times n$ data pengamatan. Metode Yates merupakan metode pendugaan untuk menangani data hilang yang dilakukan dengan meminimumkan jumlah kuadrat *error*, sedangkan metode Biggers merupakan metode pendugaan data hilang dengan pendekatan matrik. Kemudian untuk mengatasi error pada jumlah kuadrat perlakuan diperlukan tabel analisis varian alternatif sebagai alternatif analisis data hilang pada RBRL dan mengetahui pengaruh perlakuan terhadap respon pengamatan. Maka untuk satu dan dua data hilang dapat diselesaikan dengan menggunakan metode Yates. Untuk empat data hilang dapat diselesaikan dengan metode Biggers dan tiga data hilang dapat diselesaikan dengan menggunakan aturan Baten dan pendekatan matriks pada metode Biggers.

Kata Kunci: *analisis pendugaan, analisis varian alternatif, metode Yates, metode biggers, missing data*

ABSTRACT

Missing data on randomized complete block design (RCBD) is information that is not available for an object (case) specific, where the number of data observations in a randomized complete block design less than $k \times n$ observational data. analysis using Yates and Biggers method, because there is correspondence between formulas with the experimental design. Yates is a method of approximation methods for missing data handling with minimizing the sum of squared errors, while the Biggers method is a method of estimation of missing data with a matrix approach. Then to overcome the bias on the sum of squares treatment needed analysis of variants alternative table as an alternative to the analysis of missing data on RCBD and determine the effect of treatment on the response of the observation. Then for one and two missing data can be solved by using the method of Yates. For four missing data can be solved by the method of Biggers and three missing data can be solved by using Baten's rule and a matrix approach on Biggers method.

Keywords: *a randomized complete block design, approximation analysis, biggers method, missing data, the analysis of variants alternative*

Pendahuluan

Dalam suatu percobaan yang dilakukan seringkali dalam pelaksanaannya tidak berjalan lancar seperti yang diharapkan. Berbagai kendala dapat terjadi seperti kesalahan dalam penulisan jawaban atau dalam proses input data, sesuatu diluar kekuasaan kita seperti hewan atau tumbuhan percobaan mati (bukan karena perlakuan), alat ukur yang akan digunakan rusak, ataupun karena cuaca yang tidak memungkinkan. Kasus semacam ini sering disebut data hilang. Oleh karena itu, pentingnya dilakukan pendekatan terhadap data hilang. Pendekatan untuk menangani data hilang menurut Montgomery terdapat dua pendekatan analisis eksak dan analisis pendugaan. Analisis eksak yaitu tanpa menduga data hilang memang lebih mudah dan cepat untuk dilakukan, namun akan timbul masalah jika data hilang cukup besar (Little dan Rubin, 1987). Dan dilihat juga kesesuaian metode dengan rancangan percobaan yang digunakan. Kondisi ini menjadi salah satu alasan perlu dilakukan analisis pendugaan untuk menangani hilang, Pendugaan pada data hilang akan menghasilkan bias untuk jumlah kuadrat perlakuan. Sehingga diperlukan penanganan analisis varian sebagai alternatif untuk pengamatan data hilang yang lebih informatif untuk metode ini, yang dikenal dengan istilah analisis varian alternatif dan setelah diperoleh tabel analisis variansinya

juga untuk mengetahui adanya perbedaan pengaruh perlakuan terhadap respon pengamatan lakukan uji lanjut.

Metode dan Bahan Penelitian

1. Rancangan Blok Lengkap Acak (RBLA)

Rancangan Blok Lengkap Acak (RBLA) adalah suatu rancangan yang melakukan pengelompokan unit-unit percobaan ke dalam suatu blok-blok dan semua perlakuan dicobakan pada setiap blok yang ada. Tujuan memberikan blok adalah untuk memperoleh unit percobaan yang seragam mungkin dalam setiap bloknya, sehingga perbedaan yang diamati sebagian besar disebabkan oleh perlakuan.

2. Data Hilang Pada RBLA

Data hilang merupakan informasi yang tidak tersedia untuk sebuah objek tertentu, dimana banyaknya data pengamatan dalam rancangan blok lengkap acak $b \times a$ data pengamatan.

2.1 Metode Yates

Metode ini adalah penduga terhadap data hilang pada rancangan blok lengkap acak (RBLA) sehingga kuadrat tengah error minimal. Penanganan data hilang dengan analisis penduga data hilang pertama kali dikembangkan oleh (Yates, 1933) prinsip dari metode Yates ini dengan meminimumkan jumlah kuadrat error.

2.2 Metode Biggers

Dimisalkan data hilangnya adalah X_{cd} dengan prinsip yang sama dengan metode Yates untuk satu data hilang, maka dilakukan penduga:

$$\begin{aligned}
 MS_E &= \sum_{i=1}^k \sum_{j=1}^n X_{ij}^2 - \frac{1}{n} \sum_{i=1}^k X_{ij}^2 - \frac{1}{k} \sum_{j=1}^n X_{ij}^2 + \frac{X_{\dots}^2}{nk} \\
 &= \sum_{i=1}^k \sum_{j=1}^n X_{ij}^2 + \sum_i \sum_j \mathcal{K}_{cd}^2 \frac{1}{n} \left[\sum_i \left(\sum_j X_{ij} \right)^2 + \sum_i \left(\sum_j X_{cj} + \sum_j X_{cj} \right)^2 \right] \\
 &\quad - \frac{1}{k} \left[\sum_j \left(\sum_i X_{ij} \right)^2 + \sum_j \left(\sum_i X_{id} + \sum_i X_{id} \right)^2 \right] + \frac{1}{nk} [G + \mathcal{K}_{cd}]^2 \\
 &= \mathcal{K}_{cd}^2 - \frac{1}{n} \left(\sum_j X_{cj} + X_c \right)^2 - \frac{1}{k} \left(\sum_i X_{id} + X_d \right)^2 + \frac{1}{nk} \left(G + \sum_i \sum_j X_{ij} \right)^2
 \end{aligned}$$

dimana,

G : jumlah semua nilai pengamatan dengan data hilang,

$$\frac{\partial MS_E}{\partial \mathcal{K}_{cd}} = 0$$

$$\mathcal{K}_{cd} - \frac{\left(\sum_j X_{cj} + \sum_{(j)} X_{cj} \right)}{n} - \frac{\left(\sum_i X_{id} + \sum_{(i)} X_{id} \right)}{k} + \frac{\left(G + \sum_{(i)} \sum_{(j)} X_{ij} \right)}{nk} = 0$$

$$nk\mathcal{K}_{cd} - k \left(\sum_j X_{cj} + \sum_{(j)} X_{cj} \right) - n \left(\sum_i X_{id} + \sum_{(i)} X_{id} \right) + \left(G + \sum_{(i)} \sum_{(j)} X_{ij} \right) = 0$$

$$nk\mathcal{K}_{cd} - k \sum_{(j)} X_{cj} - n \sum_{(i)} X_{id} + \sum_i \sum_j X_{ij} = k \sum_j X_{cj} + n \sum_i X_{id} - G \tag{1}$$

Persamaan diatas dikelompokkan dalam suku-suku yang berhubungan dengan kelompok sekutu, perlakuan sekutu dan tanpa sekutu sebagai berikut:

$$\begin{aligned}
 & nk\bar{X}_{cd} - k\left(\sum_{(j)} X_{ij} + \bar{X}_{cd}\right) - n\left(\sum_{(i)} X_{ij} + \bar{X}_{cd}\right) + \left(\sum_{(i)} \sum_{(j)} X_{ij} + \sum_{(i)} X_{id} + \sum_{(j)} X_{cj} + \bar{X}_{cd}\right) \\
 & = k\sum_j X_{cj} + n\sum_i X_{id} - G \\
 & (n-1)(k-1)\bar{X}_{cd} + (1-k)\sum_{(j)} X_{cj} + (1-n)\sum_{(i)} X_{id} + \sum_{(i)} \sum_{(j)} X_{ij} \\
 & = k\sum_j X_{cj} + n\sum_i X_{id} - G \tag{2}
 \end{aligned}$$

Analog untuk (p-1) data hilang yang lain. Sehingga diperoleh p buah persamaan yang analog dengan (1) dan (2). Bila ditulis dalam bentuk matriks:

$$A_{p \times p} X_{p \times 1} = Q_{p \times 1} \tag{3}$$

dengan

$A_{p \times p}$: matriks simetris dengan elemen elemen (n-1)(k-1) untuk kelompok dan perlakuan yang bersesuaian, (1-n) untuk kelompok yang bersesuaian, (1-k) untuk perlakuan yang bersesuaian dan 1 untuk lainnya.

$X_{p \times 1}$: matriks dari data hilang.

$Q_{p \times 1}$: matriks nilai $kX_c + nX_d - G$ dari persamaan yang bersesuaian

Dari Persamaan (3) diperoleh:

$$X_{p \times 1} = A^{-1} Q_{p \times 1} \tag{4}$$

Untuk memperjelas matriks $A_{p \times p}$ misalkan dalam percobaan ini ada empat data hilang yaitu X_{kk}, X_{kl}, X_{mk} dan X_{st} . Elemen-elemen dari $A_{p \times p}$ ditentukan sebagai berikut: $AX = Q$

$$\begin{bmatrix} (k-1)(n-1) & 1-k & 1-n & 1 \\ 1-k & (k-1)(n-1) & 1 & 1 \\ 1-n & 1 & (k-1)(n-1) & 1 \\ 1 & 1 & 1 & (k-1)(n-1) \end{bmatrix} \begin{bmatrix} X_{kk} \\ X_{kl} \\ X_{mk} \\ X_{st} \end{bmatrix} = \begin{bmatrix} kX_k + nX_k - G \\ kX_k + nX_l - G \\ kX_m + nX_k - G \\ kX_s + nX_t - G \end{bmatrix}$$

jika terdapat tiga data hilang, dari pendugaan empat data hilang kemudian dapat digunakan untuk tiga data hilang pada RBLA. (W.D. Baten, 1939), ketika terdapat tiga data hilang atau empat pada rancangan blok lengkap acak, maka aturan-aturan tertentu sebagai berikut:

I. Kasus 1

$$\begin{bmatrix} (k-1)(n-1) & 1 & 1 \\ 1 & (k-1)(n-1) & 1 \\ 1 & 1 & (k-1)(n-1) \end{bmatrix} \begin{bmatrix} X_{kk} \\ X_{ll} \\ X_{mm} \end{bmatrix} = \begin{bmatrix} Q_{kk} \\ Q_{ll} \\ Q_{mm} \end{bmatrix}$$

II. Kasus 2

Untuk 3 data hilang pada perlakuan yang sama, jika x_i dan x_j maka elemen ij adalah (1-k),

$$\begin{bmatrix} (k-1)(n-1) & (1-k) & (1-k) \\ (1-k) & (k-1)(n-1) & (1-k) \\ (1-k) & (1-k) & (k-1)(n-1) \end{bmatrix} \begin{bmatrix} X_{kk} \\ X_{ll} \\ X_{mm} \end{bmatrix} = \begin{bmatrix} Q_{kk} \\ Q_{ll} \\ Q_{mm} \end{bmatrix}$$

Sedangkan untuk 3 data hilang pada kelompok yang sama jika x_i dan x_j maka elemen ij adalah (1-n).

III. Kasus 3

$$\begin{bmatrix} (k-1)(n-1) & (1-k) & 1 \\ (1-k) & (k-1)(n-1) & (1-k) \\ 1 & (1-k) & (k-1)(n-1) \end{bmatrix} \begin{bmatrix} X_{kk} \\ X_{ll} \\ X_{mm} \end{bmatrix} = \begin{bmatrix} Q_{kk} \\ Q_{ll} \\ Q_{mm} \end{bmatrix}$$

IV. Kasus 4

$$\begin{bmatrix} (k-1)(n-1) & (1-k) & 1 \\ (1-k) & (k-1)(n-1) & 1 \\ 1 & 1 & (k-1)(n-1) \end{bmatrix} \begin{bmatrix} X_{kk} \\ X_{ll} \\ X_{mm} \end{bmatrix} = \begin{bmatrix} Q_{kk} \\ Q_{ll} \\ Q_{mm} \end{bmatrix}$$

V. Kasus 5

$$\begin{bmatrix} (k-1)(n-1) & (1-n) & 1 \\ (1-n) & (k-1)(n-1) & 1 \\ 1 & 1 & (k-1)(n-1) \end{bmatrix} \begin{bmatrix} X_{kk} \\ X_{ll} \\ X_{mm} \end{bmatrix} = \begin{bmatrix} Q_{kk} \\ Q_{ll} \\ Q_{mm} \end{bmatrix}$$

Hasil dan Pembahasan

1. **Penanganan Dua Data Hilang pada Kasus Rancangan Blok Lengkap Acak**

Data dibawah diperoleh dari buku Design And Analysis of Experiments edisi ke delapan, sebuah artikel di jurnal keselamatan kebakaran (efek dari desain nosel pada stabilitas dan kinerja jet air yang bergejolak.

Tabel 1 Tabel Kinerja Desain Nosel Terhadap Tingkatan Jet Air dengan Dua Data Hilang

Perlakuan	Blok						Total
	1	2	3	4	5	6	
1	0.78	0.80	0.81	0.75	0.77	0.78	4.69
2	0.85	X_{22}	0.92	0.86	0.81	0.83	4.27
3	0.93	0.92	0.95	0.89	0.89	0.83	5.41
4	1.14	0.97	0.98	0.88	X_{45}	0.83	4.8
5	0.97	0.86	0.78	0.76	0.76	0.75	4.88
Total	4.67	3.55	4.44	4.14	3.23	4.02	24.05

Sumber data: *Design and Analysis of Experiments* edisi ke delapan

Pada tabel diatas terdapat dua data hilang. Kita tidak bisa mencari dua data hilang sekaligus pada Metode Yates tapi harus dicari satu persatu, pertama kita cari terlebih dahulu X_{22} , setelah itu kita cari X_{45} sebagai berikut:

$$\hat{X}_{22} = \frac{(k-1)(n-1)(k \sum_j X_{2j} + n \sum_i X_{i2}) - k \sum_j X_{4j} - n \sum_i X_{i5} - (nk - k - n) \sum_{i,j} X_{ij}}{(nk - k - n + 2)(nk - k - n)}$$

$$\begin{aligned}
 &= \frac{(4 \times 5)(5 \times 4.27 + 6 \times 3.55) - (5 \times 4.8 - 6 \times 3.23) - 19 \times 24.05}{21 \times 19} \\
 &= \frac{20(21.35 + 21.3) - (24 - 19.38) - 456.95}{399} \\
 &= \frac{352.67}{399} \\
 &= 0.884
 \end{aligned}$$

Selanjutnya akan dicari data hilang yang kedua X_{45} .

$$\begin{aligned}
 \bar{X}_{45} &= \frac{(k-1)(n-1)(k \sum_j X_{4j} + n \sum_i X_{i5}) - k \sum_j X_{2j} - n \sum_i X_{i2} - (nk - k - n) \sum_{i,j} X_{ij}}{(nk - k - n + 2)(nk - k - n)} \\
 &= \frac{(4 \times 5)(5 \times 4.8 + 6 \times 3.23) - (5 \times 4.27 - 6 \times 3.55) - 19 \times 24.05}{21 \times 19} \\
 &= \frac{20(43.38) - (42.65) - 456.95}{399} \\
 &= \frac{368}{399} = 0.92
 \end{aligned}$$

Kita peroleh dua data hilang $\bar{X}_{22} = 0.88$ dan $\bar{X}_{45} = 0.92$.

2. Penanganan Empat Data Hilang pada Kasus Rancangan Blok lengkap Acak

Pada kasus empat data hilang dibawah kita menduga data hilang menggunakan metode biggers, data yang digunakan masih menggunakan data sebelumnya.

Tabel 2 Tabel Kinerja Desain Nosel Terhadap Tingkatan Jet Air dengan Empat Data Hilang

Perlakuan	Blok						Total
	1	2	3	4	5	6	
1	X_{11}	0.80	0.81	0.75	0.77	0.78	3.91
2	0.85	0.85	0.92	0.86	0.81	X_{26}	4.29
3	0.93	X_{32}	0.95	0.89	0.89	0.83	4.49
4	1.14	0.97	0.98	0.88	0.86	0.83	5.66
5	0.97	0.86	X_{53}	0.76	0.76	0.75	4.1
Total	3.89	3.48	3.66	4.14	4.09	3.19	22.45

Sumber data: *Design and Analysis of Experiments* edisi ke delapan.

Pada tabel 2 diatas masih menggunakan data yang sebelumnya namun kali ini terdapat empat buah data yang hilang, pada kasus empat data hilang ini kita menggunakan metode Biggers.

Untuk mencari data hilang pada kasus diatas maka, dengan percobaan lima perlakuan dan enam blok, dan terdapat empat data hilang data hilang adalah X_{11} , X_{26} , X_{32} , X_{53} , kita akan menduga dengan menggunakan metode Biggers.

Sehingga persamaan $A_{pxp} X_{px1} = Q_{px1}$, maka dapat ditulis:

$$\begin{bmatrix} 12 & 1 & 1 & 1 \\ 1 & 12 & 1 & 1 \\ 1 & 1 & 12 & 1 \\ 1 & 1 & 1 & 12 \end{bmatrix} \begin{bmatrix} X_{11} \\ X_{26} \\ X_{32} \\ X_{44} \end{bmatrix} = \begin{bmatrix} Z_{11} \\ Z_{25} \\ Z_{32} \\ Z_{53} \end{bmatrix}$$

Kemudian dapat ditulis kembali dengan persamaan:

$$X_{px1} = A^{-1} Q_{px1}$$

$$\begin{bmatrix} X_{11} \\ X_{26} \\ X_{32} \\ X_{53} \end{bmatrix} = \begin{bmatrix} 0,050343 & -0,00229 & -0,00229 & -0,00229 \\ -0,00229 & 0,050343 & -0,00229 & -0,00229 \\ -0,00229 & -0,00229 & 0,050343 & -0,00229 \\ -0,00229 & -0,00229 & -0,00229 & 0,050343 \end{bmatrix} \times \begin{bmatrix} 20,44 \\ 18,14 \\ 20,88 \\ 20,01 \end{bmatrix}$$

$$\begin{bmatrix} X_{11} \\ X_{26} \\ X_{32} \\ X_{53} \end{bmatrix} = \begin{bmatrix} 0,893936 \\ 0,772883 \\ 0,917094 \\ 0,871304 \end{bmatrix}$$

Maka diperoleh dua nilai pengganti data hilang $X_{11} = 0.89$, $X_{26} = 0.77$, $X_{32} = 0.92$ dan $X_{53} = 0.87$.

kita lakukan uji selanjutnya seperti Rancangan Blok Lengkap, sehingga kita dapatkan tabel anava seperti berikut ini:

Tabel 3 Tabel Anava pada RBLA dengan Empat Data Hilang

Sumber Variasi	Derajat Bebas	Jumlah Kuadrat	Kuadrat Tengah	F
Rata-rata	1	0,21747		
Perlakuan	4	0,08103	0,02026	7,9
Blok	5	0,09539	0,01908	
Error	16	0,04105	0,00256	
Jumlah	26	0,21747		

Jika terdapat pengaruh pada perlakuan lakukan uji lanjut dengan menggunakan uji lanjut LSD.

Kesimpulan

Data hilang pada Rancangan Blok lengkap Acak menyebabkan tabel anava berubah dimana derajat bebas dari galat dan total berkurang sebanyak data hilang. Untuk menduga data hilang yang lebih dari dua kita bisa menggunakan Metode Biggers. Jika terdapat perbedaan pengaruh pada perlakuan kita lakukan uji lanjut dengan Metode *Least Significance Difference* (LSD) untuk melihat secara detail pengaruh perlakuan satu dan lainnya.

Daftar Pustaka

- [1] Arnata, I wayan, dkk. "Rancangan Percobaan: Teori, Aplikasi Spss dan Excel". Malang: Lintaskata. 2011.
- [2] B.Rubin, Donald. " Inference and Missing Data". *Biometrika*, Volume 63, issue(Desember, 1976), 581-529.

- [3] Biggers, J.D. (1959) The Estimation of Missing and Mixed-up Observation in several Experimental Design. *Biometrika* Vol. 46 no 1/2 pp. 91-105.
- [4] Kusriningrum. "Data Hilang". <https://www.slideshare.net/JauharAnam/08-data-hilang-missing-data>, diakses 18 Mei 2017.
- [5] Montgomery, Douglas C. "Design and Analysis of Experiments". Edisi 8, hal 139-182, Arizona State university, 2013.
- [6] Sari, kartika, dkk. "Penduga Data Hilang dengan Metode Yates dan Algoritma Em Pada Rancangan Lattice Seimbang". *E-Jurnal Matematika* Vol 4(2), pp. 74-82. 2015.
- [7] Sudjana. "Desain dan Analisis Eksperimen". Bandung: Tarsito. 2002.
- [8] Widiaroh. "Estimasi Data hilang Pada Rancangan Acak Kelompok Lengkap". Jurusan Matematika Fmipa Undip, 2007.