

Pendugaan Amatan Yang Hilang Pada Rancangan Acak Kelompok (RAK)

Kismiantini *

Abstrak

Pengumpulan data dengan metode percobaan seringkali ditemui adanya data hilang. Hal ini disebabkan diantaranya oleh kecerobohan peneliti atau kerusakan unit percobaan yang tidak dapat dihindari. Masalah data hilang pada rancangan acak lengkap dengan ulangan sama dapat diatasi dengan ulangan tidak sama. Namun data hilang pada rancangan acak kelompok tidak dapat diatasi dengan cara tersebut, karena ulangan pada rancangan ini berperan sebagai kelompok. Salah satu metode untuk menduga data hilang pada rancangan acak kelompok adalah dengan metode *Least Mean Square* (LMS). Masalah data hilang pada rancangan acak kelompok akan berakibat pada hasil analisis ragamnya, yaitu derajat bebas galat akan berkurang sebanyak total data hilang, dan besarnya jumlah kuadrat galat akan semakin kecil seiring dengan semakin banyaknya data hilang.

Kata Kunci : Rancangan acak kelompok, data hilang, Least Mean Square (LMS), analisis ragam.

1. Pendahuluan

Metode percobaan merupakan salah satu metode pengumpulan data yang berguna untuk menciptakan jenis perlakuan yang diinginkan, dengan mengamati perubahan-perubahan yang terjadi pada responsnya (Aunuddin, 2005). Berdasarkan metode percobaan ini muncullah berbagai jenis rancangan. Secara garis besar rancangan percobaan diklasifikasikan menjadi dua, yaitu rancangan perlakuan, misalnya rancangan satu faktor, dua faktor, tiga faktor atau lebih, dan rancangan lingkungan seperti Rancangan Acak Lengkap (RAL), Rancangan Acak Kelompok Lengkap (RAKL) atau yang lebih dikenal dengan Rancangan Acak Kelompok (RAK), Rancangan Bujur Sangkar Latin (RBSL), dan sebagainya.

Seringkali pada pengumpulan data percobaan, terdapat pengamatan yang hilang yang disebabkan oleh kecerobohan atau kerusakan unit percobaan yang tidak dapat dihindari (Montgomery, 1995), dimana hal tersebut akan berpengaruh pada teknik analisis ragamnya. Data hilang ini akan memberikan hasil analisis yang tidak valid, analisis ragam akan menyimpang, dan koefisien keragaman (KK) menjadi lebih besar atau di atas 50% (Sutarno, 2004). Bila masalah data hilang ini dijumpai pada RAL yang semula dirancang dengan ulangan sama, maka hal tersebut dapat diatasi dengan ulangan tidak sama. Namun bila data hilang terjadi pada RAK, maka cara mengatasi sebagaimana yang dilakukan pada RAL tidak lagi dapat digunakan, karena ulangan pada rancangan ini digunakan sebagai kelompok. Ada berbagai metode untuk mengatasi hal tersebut, di antaranya dengan cara transformasi data, dan *Least Mean Square* (LMS). Pada makalah ini akan dibahas masalah pendugaan data hilang pada RAK dengan menggunakan *Least Mean Square* (LMS).

* Staf Pengajar pada Jurusan Pendidikan Matematika, FMIPA Universitas Negeri Yogyakarta

1. Rancangan Acak Kelompok (RAK)

Rancangan Acak Kelompok (RAKL) baik digunakan jika keheterogenan unit percobaan berasal dari satu sumber keragaman (Mattjik & Sumertajaya, 2006). Rancangan ini digunakan bila unit percobaan tidak lagi relatif homogen, dimana unit percobaan tersebut selanjutnya dikelompokkan dengan maksud untuk mengurangi galat atau error dari percobaan yang mungkin terjadi. Montgomery (1995) memberikan model linier aditif dari RAK, dimana rancangan perlakuan yang digunakan adalah rancangan satu faktor, dalam bentuk persamaan berikut :

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, b \end{cases} \quad (1)$$

dimana

μ : rata-rata keseluruhan

τ_i : pengaruh perlakuan ke- i

β_j : pengaruh kelompok ke- j

ε_{ij} : galat dengan $\varepsilon_{ij} \sim N(0, \sigma^2)$.

Pada RAK, jika faktor pada model linier aditif pada (1) diasumsikan tetap, maka model linier aditifnya disebut sebagai model tetap (*fix model*). Sehingga model pada persamaan (1) di atas di atas dianggap memenuhi asumsi model tetap, yaitu

$$\sum_{i=1}^a \tau_i = 0, \sum_{j=1}^b \beta_j = 0, \text{ dan } \varepsilon_{ij}^{bsi} \sim N(0, \sigma^2).$$

Sedangkan jika faktor pada model linier aditif yang digunakan pada persamaan (1) di atas diasumsikan acak, maka model linier aditifnya disebut sebagai model acak (random number), sehingga model pada (1) dianggap memenuhi

$$\tau_i \sim N(0, \sigma_\tau^2), \beta_j \sim N(0, \sigma_\beta^2), \text{ dan } \varepsilon_{ij}^{bsi} \sim N(0, \sigma^2).$$

Sebagaimana yang dikemukakan sebelumnya, masalah data atau amatan yang hilang akan berpengaruh besar pada hasil analisis ragam dari model yang digunakan. Makalah ini difokuskan pada simulasi pendugaan data/amatan yang hilang pada beberapa kondisi untuk menjelaskan pengaruh data/amatan yang hilang itu, dalam bentuk ilustrasi kasus pada berbagai kondisi tersebut. Ilustrasi yang dimaksud adalah gugus data awal tanpa adanya data/amatan yang hilang, gugus data awal dimana terdapat satu data/amatan yang hilang, dan gugus data awal dimana terdapat dua data/amatan yang hilang. Selanjutnya, akan dilihat bagaimana hasil tabel analisis ragam pada masing-masing kondisi, serta melakukan pendugaan data/amatan yang hilang dengan metode LMS. Untuk itu, bagaimana uraian mengenai penggunaan LMS akan diberikan langsung pada ilustrasi yang berhubungan.

2. Ilustrasi Pendugaan Data Hilang pada Rancangan Acak Kelompok (RAK)

3.1 Ilustrasi Model RAK Tanpa Ada Data Hilang

Kismiantini

Suatu percobaan dilakukan untuk mengetahui keausan suatu ban. Telah dilakukan penelitian dengan menggunakan Rancangan Acak Kelompok (RAK). Sebagai perlakuan didasarkan pada jenis ban yang digunakan, dan taraf atau level dari perlakuan yang digunakan dimisalkan dengan jenis ban A, B, C, dan D. Sedangkan jenis mobil dianggap sebagai faktor kelompok, dimana dalam penelitian ini dimisalkan dengan jenis mobil I, II, III dan IV. Sedangkan yang menjadi respon adalah masa pemakaian ban dalam 20000 mil. Pada tabel berikut diberikan data yang digunakan dalam makalah ini.

Tabel 1. Data Lengkap untuk Masa Pemakaian Ban

Mobil	Ban				y_j
	A	B	C	D	
I	17	14	12	13	56
II	14	14	12	11	51
III	13	13	10	11	47
IV	13	8	9	9	39
y_i	57	49	43	44	193

Teknik analisis ragam untuk data pada ilustrasi tersebut dapat diselesaikan dengan menggunakan bantuan software SAS 9.1. Program yang digunakan diberikan pada kotak berikut.

```

data ban;
input mobil merk$ pemakaian@@;
datalines;
1 A 17 1 B 14 1 C 12 1 D 13
2 A 14 2 B 14 2 C 12 2 D 11
3 A 13 3 B 13 3 C 10 3 D 11
4 A 13 4 B 8 4 C 9 4 D 9
;
proc print;
run;
proc anova;
class mobil merk;
model pemakaian=mobil merk;
run;
```

Selanjutnya, diperoleh ringkasan tabel analisis ragam dari hasil pengolahan SAS 9.1, yang diberikan pada Tabel 2 berikut.

Tabel 2. Analisis Ragam pada Data Lengkap untuk Masa Pemakaian Ban

Sumber Keragaman	db	JK	KT	F_{hit}	$Pr > F$
Kelompok	3	38.6875	12.89583333	10.04	0.0031
Perlakuan	3	30.6875	10.22916667	7.96	0.0067
Galat	9	11.5625	1.28472222		
Total	15	80.9375			

$$R^2 = 0.857143, KK = 9.396525, RMSEP = 1.133456.$$

Dari ilustrasi data di atas, diketahui bahwa model linier aditif untuk perlakuan penggunaan jenis ban mobil merupakan model tetap (*fix model*), karena perlakuan hanya didasarkan pada empat level jenis ban yang dicobakan, sehingga kesimpulan akhir juga akan berkisar hanya pada keempat jenis ban tersebut. Dari tabel di atas terlihat bahwa nilai Kuadrat Tengah Galat (KTG)

Kismiantini

yang diperoleh adalah 1.284, dengan nilai statistik uji F yang diperoleh adalah 7.96, dengan signifikansi 0.0067. Maka dapat dikatakan bahwa ada pengaruh yang signifikan penggunaan keempat jenis ban yang berbeda, yaitu A, B, C dan D terhadap aus tidaknya ban dalam pemakaian.

3.2 Ilustrasi Model RAK dengan Satu Amatan yang Hilang

Untuk menjelaskan penggunaan metode *Least Mean Square* pada data hilang satu pengamatan pada rancangan acak kelompok, digunakan ilustrasi yang sama dengan ilustrasi di atas, namun terdapat satu data yang hilang seperti yang ditunjukkan pada tabel berikut. Pada Tabel 3, amatan yang hilang dinotasikan dengan huruf x .

Tabel 3. Data untuk Masa Pemakaian Ban

Mobil	Ban				y_j
	A	B	C	D	
I	17	14	12	13	56
II	14	14	12	11	51
III	13	13	x	11	$37 + x$
IV	13	8	9	9	39
y_i	57	49	$33 + x$	44	$183 + x$

Secara umum, untuk mencari jumlah kuadrat galat (JKG) dari suatu model linier aditif sebagaimana yang diberikan pada persamaan (1), dan dengan menggunakan rancangan perlakuan RAK yang sering disebut juga RAK lengkap atau RAKL (*Randomized Completely Block Design*), dengan a perlakuan dan b kelompok adalah

$$JKG = \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2 \quad (2)$$

$$JKG = \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \frac{1}{b} \sum_{i=1}^a \left(\sum_{j=1}^b y_{ij} \right)^2 - \frac{1}{a} \sum_{j=1}^b \left(\sum_{i=1}^a y_{ij} \right)^2 + \frac{1}{ab} \left(\sum_{i=1}^a \sum_{j=1}^b y_{ij} \right)^2 \quad (3)$$

$$JKG = x^2 - \frac{1}{b} (y_{i.}' + x)^2 - \frac{1}{a} (y_{.j}' + x)^2 + \frac{1}{ab} (y_{..}' + x)^2 + R \quad (4)$$

dimana

$y_{..}'$: total pengamatan dengan satu data hilang

$y_{i.}'$: total perlakuan dengan pengamatan satu data hilang

$y_{.j}'$: total kelompok dengan satu pengamatan data hilang

R : konstanta yang tidak memuat data hilang x .

Selanjutnya, jika JKG diturunkan terhadap x dan disamakan dengan nol, maka dapat ditulis

$$\frac{d(JKG)}{dx} = 0 \quad (5)$$

Sehingga diperoleh rumus umum untuk data hilang satu pengamatan sebagai berikut :

Kismiantini

$$x = \frac{ay'_i + by'_j - y'_..}{(a-1)(b-1)} \quad (6)$$

Jika nilai amatan yang hilang sudah dapat diduga atau diketahui, maka tabel analisis ragam untuk data sudah dapat ditentukan menggunakan cara yang biasa digunakan pada permasalahan Perancangan Percobaan lainnya.

Diketahui bahwa jumlah kuadrat galat dari Rancangan Acak Kelompok (RAK) dengan model linier aditif pada persamaan (1), sebagaimana yang diuraikan pada persamaan (2) sampai (4), maka diperoleh :

$$\begin{aligned} JKG &= JKT - JKP - JKK \\ &= \sum_{i=1}^a \sum_{j=1}^b y_{ij}^2 - \frac{1}{b} \sum_{i=1}^a y_i^2 - \frac{1}{a} \sum_{j=1}^b y_j^2 + \frac{y_{..}^2}{N} \end{aligned}$$

Berdasarkan nilai pada Tabel 3 di atas dan dengan menggunakan (4), maka diperoleh jumlah kuadrat galat atau JKG menjadi :

$$JKG = x^2 - \frac{1}{4}(33+x)^2 - \frac{1}{4}(37+x)^2 + \frac{(183+x)^2}{16} + R,$$

dengan R adalah konstanta yang tidak memuat x . Selanjutnya JKG tersebut diturunkan terhadap x dan disamakan dengan nol untuk mendapatkan nilai amatan yang hilang. Atau,

$$\frac{d(JKG)}{dx} = 2x - \frac{2}{4}(33+x) - \frac{2}{4}(37+x) + \frac{2}{16}(183+x)$$

dan

$$\frac{d(JKG)}{dx} = 0$$

sehingga

$$x = 10.78$$

Sebagai akibat dari proses pendugaan amatan hilang di atas, maka derajat bebas bagi galat berkurang satu.

Untuk mendapatkan tabel analisis ragam pada data dimana terdapat satu amatan yang hilang, maka akan digunakan bantuan *software* SAS 9.1 dengan perintah sebagai berikut.

```
data ban;
input mobil merk$ pemakaian@@;
datalines;
1 A 17 1 B 14 1 C 12 1 D 13
2 A 14 2 B 14 2 C 12 2 D 11
3 A 13 3 B 13 3 C . 3 D 11
4 A 13 4 B 8 4 C 9 4 D 9
;
proc print;
run;
proc anova;
class mobil merk;
```

Kismiantini

```

model pemakaian=mobil merk;
run;
```

Hasil ringkasan tabel analisis ragam dari output SAS 9.1, dimana amatan yang hilang belum diketahui/diduga nilainya, diperlihatkan pada Tabel 4 berikut.

Tabel 4. Analisis Ragam untuk Satu Amatan Hilang, Data Masa Pemakaian Ban.

Sumber Keragaman	db	JK	KT	F _{hit}	Pr > F
Kelompok	3	38.23333333	12.74444444	9.05	0.0060
Perlakuan	3	26.90000000	8.96666667	6.37	0.0163
Galat	8	11.26666667	1.40833333		
Total	14	76.40000000			

$$R^2 = 0.852531, KK = 9.727313, RMSEP = 1.186732$$

Terlihat dari tabel di atas bahwa nilai derajat bebas total berkurang 1 dari nilai derajat bebas total pada Tabel 2. Terlihat juga bahwa nilai KTG atau RMSEP untuk data dengan 1 amatan hilang meningkat dibandingkan dengan data dimana tidak ada amatan yang hilang pada Tabel 2, yaitu nilai 1.4083 dan 1.1867 dibandingkan nilai 1.2847 dan 1.3334.

3.3 Ilustrasi Model RAK dengan Amatan yang Hilang Lebih dari Satu

Untuk menjelaskan penggunaan metode *Least Mean Square* pada data dengan amatan hilang lebih dari satu, dalam hal ini ada 2 amatan yang hilang pada rancangan acak kelompok, digunakan ilustrasi yang sama dengan ilustrasi di atas, seperti yang ditunjukkan pada tabel berikut. Pada Tabel 5, amatan yang hilang dilihat pada sel yang kosong, atau pada nilai respon Y_{33} dan Y_{14} .

Tabel 5. Data untuk Masa Pemakaian Ban

Mobil	Ban				y_j
	A	B	C	D	
I	17	14	12		56
II	14	14	12	11	51
III	13	13		11	$37 + x$
IV	13	8	9	9	39
y_i	57	49	$33 + x$	44	$183 + x$

Pendugaan amatan yang hilang, dimana terdapat lebih dari satu amatan yang hilang, maka prosedur pendugaan dengan *Least Mean Square* (LMS), dilakukan dengan beberapa langkah berikut.

Langkah 1. Menduga salah satu data hilang dengan rata-rata.

Misalkan yang akan diduga pertama kali adalah Y_{33} , maka rumus yang digunakan adalah

$$Y_{33} = \left(\frac{12+12+9}{3} + \frac{13+13+11}{3} \right) / 2 = 11.667$$

Langkah 2. Substitusikan nilai dugaan Y_{33} ke dalam Tabel 5 di atas, sehingga diperoleh tabel berikut.

Kismiantini

Mobil	Ban				y_j
	A	B	C	D	
I	17	14	12		43
II	14	14	12	11	51
III	13	13	11.667	11	48.667
IV	13	8	9	9	39
y_i	57	49	44.667	31	181.667

Langkah 3. Menduga nilai amatan yang hilang untuk sel Y_{14} . Rumus yang digunakan adalah

$$Y_{14} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 31) + (4 \times 43) - 181.667}{(4-1)(4-1)} = 12.704$$

Langkah 4. Substitusikan nilai Y_{14} ke dalam Tabel 5, dan hilangkan nilai Y_{33} yang dihasilkan pada Langkah 1, kemudian lakukan kembali pendugaan nilai Y_{33} dengan rumus yang digunakan pada Langkah 3, sehingga diperoleh

Mobil	Ban				y_j
	A	B	C	D	
I	17	14	12	12.704	55.704
II	14	14	12	11	51
III	13	13		11	37
IV	13	8	9	9	39
y_i	57	49	33	31	182.704

$$Y_{33} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 33) + (4 \times 37) - 182.704}{(4-1)(4-1)} = 10.811$$

Langkah 5. Ulangi kembali Langkah 4 untuk menduga Y_{14} , menggunakan penduga Y_{33} yang baru

$$Y_{14} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 31) + (4 \times 43) - 180.811}{(4-1)(4-1)} = 12.799$$

Langkah 6. Ulangi kembali Langkah 5 untuk menduga Y_{33} , menggunakan penduga Y_{14} yang baru, sehingga diperoleh

$$Y_{33} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 33) + (4 \times 37) - 182.799}{(4-1)(4-1)} = 10.8$$

Langkah 7. Ulangi kembali Langkah 6 untuk menduga Y_{14} , menggunakan penduga Y_{33} , sehingga diperoleh

Kismiantini

$$Y_{14} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 31) + (4 \times 43) - 180.8}{(4-1)(4-1)} = 12.8$$

Langkah 8. Ulangi kembali Langkah 7 untuk menduga Y_{33} , menggunakan penduga Y_{14} yang baru, sehingga diperoleh

$$Y_{33} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 33) + (4 \times 37) - 182.8}{(4-1)(4-1)} = 10.8$$

Langkah 9. Ulangi kembali Langkah 8 untuk menduga Y_{14} , menggunakan penduga Y_{33} , sehingga diperoleh

$$Y_{14} = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{(a-1)(b-1)} = \frac{(4 \times 31) + (4 \times 43) - 180.8}{(4-1)(4-1)} = 12.8$$

Langkah 4-9 tersebut diulangi terus sampai diperoleh nilai dugaan untuk amatan hilang yang konvergen (nilainya tidak mengalami perubahan lagi). Dari hasil pendugaan dengan LMS diperoleh nilai dugaan kedua amatan yang hilang pada Tabel 5 adalah $Y_{33} = 10.8$ dan $Y_{14} = 12.8$. Sehingga Tabel 5 beserta nilai dugaannya dituliskan dalam tabel berikut.

Mobil	Ban				y_j
	A	B	C	D	
I	17	14	12	12.8	55.8
II	14	14	12	11	51
III	13	13	10.8	11	47.8
IV	13	8	9	9	39
y_i	57	49	43.8	43.8	193.6

Sedangkan tabel analisis ragam untuk data di atas dengan adanya dua amatan yang hilang diperoleh dengan bantuan *software* SAS 9.1, diperoleh dengan perintah berikut.

```

data ban;
input mobil merk$ pemakaian@@;
datalines;
1 A 17 1 B 14 1 C 12 1 D .
2 A 14 2 B 14 2 C 12 2 D 11
3 A 13 3 B 13 3 C . 3 D 11
4 A 13 4 B 8 4 C 9 4 D 9
;
proc print;
run;
proc anova;
class mobil merk;
model pemakaian=mobil merk;
run;

```

Kismiantini

Hasil ringkasan tabel analisis ragam dari output SAS 9.1 diberikan pada tabel berikut.

Tabel 6. Analisis Ragam pada Data Masa Pemakaian Ban, Dua Amatan Hilang.

Sumber Keragaman	db	JK	KT	F _{hit}	Pr > F
Kelompok	3	38.88095238	12.96031746	17.16	0.0013
Perlakuan	3	31.54761905	10.51587302	13.93	0.0025
Galat	7	5.28571429	0.75510204		
Total	13	75.71428571			

$$R^2 = 0.930189, KK = 7.156191, RMSEP = 0.868966.$$

Dari tabel di atas terlihat bahwa nilai derajat bebas total berkurang 2, sesuai dengan banyaknya amatan yang hilang pada data. Sedangkan nilai KTG dan RMSEP terlihat lebih kecil dari data lengkap atau data dengan satu amatan hilang, yaitu 0.7551 dan 0.8689.

4. Kesimpulan

Pendugaan data hilang satu, dua atau lebih pada Rancangan Acak Kelompok dapat dilakukan dengan metode *Least Mean Square* (LMS). Data hilang akan berpengaruh pada analisis ragam, yaitu bila data hilang ada sebanyak k maka derajat bebas akan berkurang sebanyak k pula. Amatan yang hilang juga dapat mempengaruhi jumlah kuadrat galat menjadi lebih kecil.

Daftar Pustaka

- [1] Aunuddin. 2005. *Statistika: Rancangan dan Analisis Data*. Bogor: IPB Press.
- [2] A. Gelman. 2005. Analysis of variance - Why it is more important than ever. *The Annals of Statistics*, 33: 1-33.
- [3] A. A. Mattjik dan I. M. Sumertajaya. 2006. *Perancangan Percobaan*. Bogor: IPB Press.
- [4] D. C. Montgomery. 1995. *Design and Analysis of Experiments*. New York : John Wiley & Sons.
- [5] Sutarno. 2004. Pemanfaatan MStat-C dalam analisa faktorial data hasil penelitian pertanian. *Informatika Pertanian* 13 : 721-734.