

APLIKASI R PADA AKURASI ESTIMASI OBSERVASI HILANG DALAM PENELITIAN RANCANGAN ACAK KELOMPOK

Lisana SP¹, Soekardi Hadi Prabowo²

Program Studi Matematika F-Sainstek Universitas Islam As-Syafiiyah^{1,2}

email: lisanasp4128@gmail.com¹, sh.Prabowo59@gmail.com²

DOI: <https://doi.org/10.34005/ms.v2i1.3813>

ABSTRAK

Tujuan penelitian ini mengkaji aplikasi software *R* pada akurasi estimasi observasi yang hilang dalam penelitian Rancangan Acak Lengkap (RAK), serta analisis signifikansi hasil pengujian F-ANAVA. Proses Estimasi observasi hilang menggunakan metode Yates berdasarkan prinsip *Least Mean Square* (LMS). Sedangkan akurasi estimasi observasi hilang menggunakan kriteria Means Absolute Percentage Error (MAPE). Penelitian menghasilkan hasil akurasi estimasi observasi hilang menunjukkan masuk dalam kriteria sangat baik atau akurat. Observasi hilang berpengaruh pada proses perhitungan Analisis Variansi, kehilangan dua observasi mengakibatkan derajat kebebasan berkurang 2 dan mempengaruhi jumlah kuadrat residu sehingga menjadi kecil.

Kata Kunci : Observasi Hilang, Metode Yates-LMS, *Software R*, MAPE, *Two ay ANOVA*

ABSTRACT

The aim of this research is to examine the application of R Software on the Accuracy of Estimating Missing Observations in Completely Randomized Design (RAK) research, as well as the significance analysis of ANOVA test results/missing observation estimation process using the Yates method based on the Least Means Square (LMS) principle. Meanwhile, the accuracy of the estimation of missing observations uses the Means Absolute Percentage Error (MAPE) criteria. The research provides results the accuracy of the estimation of missing observations shows that it falls within the criteria of very good or accurate. The missing observations affect the variance analysis calculation process, losing two observations results in the degree of freedom being reduced by 2 and affecting the number of residual squares so that it becomes small.

Keywords: Loss of Observations, Yates-LMS Method, R Software, MAPE, Two ay ANOVA

P

PENDAHULUAN

Rancangan percobaan merupakan suatu tindakan yang dilakukan dalam rangka menemukan beberapa prinsip yang belum diketahui, serta menguji, menguatkan dan menjelaskan beberapa kebenaran yang diduga sementara atau dihipotesiskan. Di sisi lain rancangan percobaan merupakan salah satu teknik pengumpulan data yang berguna untuk menciptakan jenis perlakuan yang diinginkan, dengan cara mengamati perubahan-perubahan yang terjadi pada variabel terikat (*Dependent*) (Kismiantini, 2016). Berdasarkan percobaan ini muncul berbagai jenis rancangan, salah satu diantaranya Rancangan Acak Kelompok (RAK). Pengumpulan data pada penelitian RAK Seringkali ditemui adanya kasus observasi hilang (Kismiantini, 2016). Menurut Gomez and Gomez (2005) hilangnya observasi dapat disebabkan oleh perlakuan yang tidak tepat, kerusakan pada obyek percobaan dan observasi yang tidak logis. Permasalahan observasi hilang akan berpengaruh besar pada hasil uji variansi (ANAVA) dengan statistik F yang digunakan pada analisis data rancangan percobaan.

Untuk mengatasi Observasi hilang pada RAK salah satu diantaranya dapat diestimasi dengan metode Yates. Prinsip metode Yates adalah meminimumkan jumlah kuadrat *Least Means Square*. Penentuan nilai estimasi untuk dua atau lebih observasi yang hilang pada RAK dilakukan melalui proses iterasi sampai nilai estimasi konvergen suatu nilai tertentu. Dan lebih dahulu dilakukan nilai estimasi awal pada salah satu sel observasi hilang.

Pada estimasi beberapa observasi hilang untuk memperoleh hasil yang akurat dan efektif serta memudahkan perhitungan, dalam makalah ini digunakan paket software open Source R yang gratis dan tidak memerlukan lisensi bagi Penggunaanya [7 Prabowo H Soekardi]. R dikembangkan oleh Robert Ihaka dan Robert Gentleman dengan versi paling awal dibuat tahun 1992 (Prabowo H S dan Hardya, 2018), Sedangkan versi terakhir per Juli 2023 merupakan versi 4.32 yang digunakan sebagai alat bantu olah data pada makalah ini.

Berdasarkan latar belakang di atas, masalah penelitian difokuskan pada Akurasi estimasi dua observasi dan pengaruhnya pada hasil Analisis Variansi F berbasis *software R*. Tujuan penelitian untuk mengetahui Aplikasi Software R pada kriteria akurasi estimasi dua observasi hilang dan pengaruh observasi hilang pada hasil analisis variansi F.

1. KAJIAN TEORI

Rancangan Acak Kelompok (RAK)

Rancangan Acak Kelompok Lengkap (RAK) didefinisikan sebagai suatu percobaan yang kombinasi perlakuannya ditetapkan secara acak pada unit percobaan dalam satu blok. Secara umum, blok tidak dapat diacak karena Blok tersebut mewakili faktor-faktor yang memiliki keterbatasan dalam pengacakan seperti lokasi, tempat waktu, dan lain-lain.

Model linier Rancangan Acak Kelompok (RAK) dapat dituliskan dalam bentuk persamaan atau Formula sebagai berikut (Montgomery, 1976):

$$Y_{ij} = \mu + \tau_i + \beta_j + e_{ij} \quad (1)$$

ϵ

Agar Inferensi hasil pengujian valid, maka ε_{ij} menyebar bebas, Identik berdistribusi normal, $\varepsilon_{ij} \sim N(0, \sigma^2)$. Untuk Model tetap Diasumsikan $\sum_i \tau_i = 0$ dan $\sum_i \beta_j = 0$. untuk model acak diasumsikan setiap β_j menyebar bebas identik berdistribusi normal $(0, \sigma^2)$, τ_i menyebar bebas identik, berdistribusi Normal $(0, \sigma^2)$. Dengan metode kuadrat terkecil, diperoleh estimasi bag i blok ke-i dan perlakuan ke-j yaitu:

$$\hat{\beta}_i = \bar{Y}_{i.} - \bar{Y}_{..} \text{ dan } \hat{\tau}_i = \bar{Y}_{.j} - \bar{Y}_{..} \quad (2)$$

Sedangkan hipotesis (hipotesis nol) yang dirumuskan bahwa tidak ada pengaruh perlakuan terhadap respon. Melawan Hipotesis alternatif apengaruh perlakuan terhadap respon. Melawan Hipotesis alternatif ada pengaruh perlakuan terhadap respon. Secar umum Tabel Anova dua arah disajikn pada Tabel.1 berikut ini :

Tabel.1 Analisis Varian untuk RAKL

Sumbe Variasi	f	Sum of Square	Mean of Square	Nilai F	Konfirmasi
Blok	T	SS[B]	MS[B]	F Hitung	F_{Tabel} ε
Perlakuan	b(	SS[P]	MS[P]		
Residu	t)	SS[R]	MS [E]		
Total	rt-1	JK[T]			

Untuk pengujian dengan taraf uji α , $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ lawan H_a :sedikitnya ada $\mu_i \neq \mu_j$ untuk $i \neq j$ (1,2,...,t),digunakan statistik uji

Untuk pengujian dengan taraf uji α , $H_0: \mu_1 = \mu_2 = \dots = \mu_t$ lawan H_a :sedikitnya ada $\mu_i \neq \mu_j$ untuk $i \neq j$ (1,2,...,t),digunakan statistik uji

$$F_{Hitung} = \frac{SS[P]}{SS[E]} \sim F_{\alpha, (r-1), (r-1)(t-1)} \quad (3)$$

Estimasi Observasi Hilang

Sebelum melakukan Uji hipotesis menggunakan Tabel ANAVA dua arah di atass, lebih dahulu perlu dilakukan estimasi dua observasi Hilang dilakukan menggunakan metode Yates berdasarkan prinsip meminimumkan jumlah rata-rata jumlah kuadrat (*Least mean Square, LMS*) error. Pertama-tama menentukan nilai estimasi awal salah satu data hilang (h_{ij}) menggunakan persamaan (Sariayu Ni Made, dkk, 2015).

$$h_{ij} = \frac{\bar{Th}_{i.} + \bar{Th}_{.j}}{2} \quad (4)$$

$$\bar{Th}_{i.} = \frac{1}{k} \sum_{j=1}^k Y_{ij} \quad \bar{Th}_{.j} = \frac{1}{k} \sum_{i=1}^k Y_{ij} \quad (5)$$

selanjutnya melakukan proses estimasi observasi hilang lainnya menggunakan rumus :

$$h_{ij} = \frac{[r \cdot h_{i.} + ch_{.j} - (Ta + TE_{(i-1).})]}{[(r-1) \cdot (t-1)]} \quad (6)$$

AKURASI ESTIMASI OBSERVASI

Akurasi hasil estimasi Observasi dilakukan menggunakan kriteria *Mean Square Percentage Error* (MAPE) dengan persamaan sebagai berikut :

$$MAPE = \frac{1}{m} \sum_{i=1}^m \frac{|TY_i - YT'_i|}{TY_i} \quad (7)$$

Dalam hal ini TY_i nilai observasi baris-baris tanpa estimasi data hilang dan YT'_i nilai estimasi baris-baris dengan nilai estimasi.

Berdasarkan nilai MAPE dikembangkan oleh Lewis. (1982)

Menurut Lewis (19820, Semakin kecil nilai MAPE, semakin baik pula suatu model. Skala untuk menilai akurasi suatu model berdasarkan nilai. Kriteria Akurasi Estimasi observasi, dilakukan dengan mengadopsi kriteri interpretasi nilai MAPE yang dikembangkan oleh Lewis. (1982), disajikan dalam Tabel.3 sebagai berikut :

Tabel.3 Interpretasi Nilai MAPE untuk Akurasi Estimasi Observasi

Nilai MAPE (%)	Kriteria Akurasi
< (Kurang Dari) 10	Sangat Akurat
$10 \leq \text{MAPE} < 20$	Akurat (Baik)
$20 \leq \text{MAPE} < 50$	Wajar
>(lebih dari) 50	Tidak Akurat (Tidak Baik)

Sumber : Lewis (1982)

2. METODE PENELITIAN

Sumber Data pada penulisan makalah ini merupakan Data sekunder, berbasis referensi hasil penelitian Rancangan Percobaan Acak Kelompok yang dilakukan oleh Kismiantini (2016) dengan tujuan mengetahui keausan masa Pemakaian Ban. Populasi penelitian berbagai merek Ban dan Jenis Mobil, sedangkan sampel perlakuan penelitian merek Ban A, B, C, D sedangkan jenis mobil I, II, III, IV sebagai faktor kelompok (*Block*) Fokus kajian hanya membahas yang berkaitan dengan dua observasi hilang yang disajikan pada Tabel.1 sebagai berikut :

Tabel 4. Data Masa Pemakaian Ban Dua Observasi hilang

Merk Ban	Tipr Mobil				Total baris
	I	II	III	IV	
A	17	14	13	13	57
B	14	14	13	8	49
C	12	12	h_{33}	9	$33+h_{33}$
D	h_{41}	11	11	9	$31+h_{41}$
Total Kolom	$43+h_{41}$	51	$37+h_{33}$	39	$183+h_{33}+h_{41}$

Sumber : Data penelitian Kismiantini, 2016

Prosedur estimasi dua observasi Hilang pada Tabel.4 di atas dilakukan menggunakan metode Yates berdasarkan prinsip meminimumkan jumlah rata-rata jumlah kuadrat (*Least mean Square, LMS*) error. Berikut ini langkah - langkah estimasi dua observasi hilang metode Yates disertai format sintaks atau *script software open source R* sebagai berikut

langkah pertama, mengkonversi Data Tabel.1 pada format Input *Software R*

```
A<-c(17,14,13,13)      C<-c(14,14,13,0)      B<-c(12,12,h33,9)      D<-c(h41,11,11,9)
Data<-matrix(c(A.,B,C,D),nr=4,ncol=4)
```

langkah kedua, Menentukan nilai estimasi awal salah satu data hilang (h_{ij}) menggunakan formula (4) dan (5), dan konversi dalam format *Software R* :

```
h.3<-sum(MB[,3])/3  h3<-sum(MB[3,])/3  h33<-(MR.3+MR3)/2
```

langkah ketiga, Gabung nilai estimasi awal h_{33} langkah kedua pada Tabel.1, selanjutnya lakukan proses estimasi data hilang lainnya h_{41} menggunakan Formula (6), dikonversikan ke dalam format Skript *Software R* :

```
MBE<-matrix(c(A,B,CE,Dh),nr=4,ncol=4)      DE<-t(MBE)      nnTE<-sum(DE)
SBh.1<-sum(DE[,1])      SBh4<-sum(DE[4,])  H.1<-r*SBh.1  H4<-r*SBh4
```

#estimasi observasi Hilang data DE[4,1]#

$$DE_{41} <- ((H_4 + H_{.1}) - TE) / ((r-1) * (c-1))$$

Langkah keempat sampai langkah tertentu, dengan melakukan iterasi atau mengulang langkah kedua dan ketiga sampai diperoleh nilai estimasi dari kedua observasi hilang yang konvergen kesuatu nilai tertentu.

Sedangkan Perhitungan akurasi hasil estimasi Observasi dilakukan menggunakan kriteria *Mean Square Percentage Error (MAPE)* berdasarkan persamaan (7), kemudian dikonversikan ke dalam Format Skript R :

```
A<matrix(c(mean(C),mean(D)),ncol=2)
```

```
CE<-c(12,12,E33b,9)
```

```
DE<-c(E41b,11,11,9)
```

```
ES<-matrix(c(mean(CE),mean(DE)),ncol=2)
```

```
k<-2
```

```
PE<-sum((abs(ES-A))/A)
```

```
MAPE<- (1/k)*(PE*100)
```

```
print(MAPE<=30)
```

untuk memudahkan penentuan criteria pada sintaks R hanya diambil dua criteria sbb:

#Jika TRUE artinya cKriteria Akurasi estimasi data Hilang Cukup Baik#

#Jika FALSE artinya Kriteria Akurasi Estmasi Data Hilang kurang baik#

Kriteria akurasi estimasi observasi hilang dikonsultasikan dengan kriteria MAPE yang diajukan oleh (Lewis, 1982), tersaji pada Tabel,3.

Perhitungan ANAVA Dua Jalur (Two Way ANOVA)

Penentuan statistik F pada uji ANAVA Dua Jalur (*Two Way ANOVA*) untuk data yang mengandung Data Hilang, diperoleh dengan bantuan *Software Open Source R*, melalui konversi data pada Tabel.1 ke dalam format sintaks *R* sebagai berikut :

```
paint<-
```

```
data.frame(mpb=c(17,14,13,13,14,14,13,8,12,12,0,9,0,11,11,9),jmb1=factor(rep(1:4,4)),mkb=factor(rep(c("A","B","C","D"),c(4,4,4,4))))
```

```
anova(lm(mpb~mkb,data=paint))
```

```
anova(lm(mpb~jmb1,data=paint))
```

```
anova(lm(mpb~mkb+jmb1,data=paint))
```

```
F1.hitung<-1.4877
```

```
F1.tabel<-qf(0.95,3,7)
```

```
#Signifikansi Uji Jenis Mobil-masa pakai Ban Two Way Anova #
print(F1.hitung>=F1.tabel)
```

#Jika TRUE artinya Merk Ban berpengaruh Signifikan terhadap Masa pemakaian Ban#

#Jika FALSE artinya Merk Ban tidak berpengaruh Signifikan terhadap Masa pemakaian Ban#

```
F2.hitung<-0.3605
```

```
F2.tabel<-qf(0.95,3,7)
```

```
#Signifikansi Uji Merk Ban-masa pakai Ban Two Way Anova#
print(F2.hitung>=F2.tabel)
```

#Jika TRUE artinya Jenis Mobil berpengaruh Signifikan terhadap Masa pemakaian Ban#

#Jika FALSE artinya Jenis Mobil tidak berpengaruh Signifikan terhadap Masa pemakaian Ban#

Selanjutnya melakukan uji Hipotesis setelah nilai hasil estimasi observasi hilang yang konvergen yaitu h_{33} dan h_{41} berturut-turut ke nilai tertentu, dimasukkan ke dalam Tabel1 dan kemudian dikonversi lagi ke dalam sintaks R

```
paint<-
```

```
data.frame(mpb=c(17,14,13,13,14,14,13,8,12,12,10,8,9,12,8,11,11,9),jmb1=factor(rep(1:4,4)),mkb=factor(rep(c("A","B","C","D"),c(4,4,4,4))))
```

lakukan kembali operasi olah data algoritma sintaks R uji hipotesis ANAVA dua Jalur di atas dengan cara menggantu sintaks input data R sebelum nilai estimasi dengan konversi data setelah nilai estimasi observasi hilang dimasukkan ke dalam Tabel .4

Metode penelitian menjelaskan rancangan kegiatan, ruang lingkup atau objek, bahan dan alat utama, tempat, teknik pengumpulan data, definisi operasional variable penelitian, dan teknik analisis. [Arial, 10, normal].

3. HASIL DAN PEMBAHASAN

, Berdasarkan perhitungan melalui proses olah data menggunakan Algoritma sintaks software R di atas, diperoleh hasil sebagai berikut

Nilai Estimasi Observasi Hilang

Hasil perhitungan menggunakan metode *Yates* dengan prinsip *Least mean Square (LMS)*, setelah dilakukan iterasi perhitungan sampai ke kesembilan disajikan ke dalam Tabel.3 sebagai berikut :

TABEL. 3 KRITERIA *MAPE*

Sel Observasi Hilang	Nilai Estimasi Observasi Hilang		
	Awal	Hasil iterasi 6-9	Akhir
h₃₃	11.666 7	Konvergen ke	11.8
h₄₁	12.703 7	Konvergen ke	12.8

Sumber : Hasil Olah Data menggunakan **Software R** versi 4.12

Dari Tabel.3 terlihat bahwa nilai estimasi dua observasi hilang pada Tabel.1 untuk sel **h₃₃** dan **h₄₁** hasil iterasi menghasilkan nilai yang konvergen berturut-turut ke nilai 10.8 dan 12.8.

Akurasi Hasil Estimasi Observasi Hilang

Perhitungan ukuran ketepatan atau akurasi nilai observasi hilang pada baris ketiga dan keempat setelah sebelum dan setelah memasukkan nilai observasi hilang pada Tabel.1, menggunakan formula *MAPE* Hasilnya disajikan dalam Tabel.6 berikut ini

Tabel.4 Nilai *Means Square Error (MAPE)*

	<i>MAPE (%)</i>		
	Nilai	Cakupan interval	Kriteria
Dua observasi Hilang	3.211	< 20 (True)	Sangat baik Baik

Sumber : Hasil Olah Data menggunakan **Software R** versi 4.12

Berdasarkan Tabel.4 terlihat nilai *MAPE* yang diperoleh sebesar 3,211% masuk dalam cakupan interval kurang dari (<) 20%. Dengan demikian akurasi estimasi dua observasi hilang tersebut dinyatakan masuk ke dalam kriteria sangat baik.

Uji hipotesis F-ANAVA 2 Jalur (*Two Way ANOVA*)

Hasil ringkasan Tabel Analisis Varian dua jalur (*Two way ANOVA*) untuk data sebelum dan sesudah dimasukkan nilai observasi hilang disajikan pada Tabel.5 berikut ini :

**Tabel.5 HASIL UJI F Analsis Variansi 2 Jalur Sebelum dan Sesudah
Memasukkan nilai observasi Hilang**

Sumber Variasi		Nilai F Hasi Uji ANAVA			
		Hitung	Db.error	Tabel	Keputusan
Perlakuan (Merk Ban)	Sebelum	1.487	7	4.347	Tidak signifikan
	Sesudah	7.811	9	3.863	Signifikan
Kelompok (Jenis Mobil)	Sebelum	0.361	7	4.347	Tidak Signifikan
	Sesudah	10.061	9	3.863	Signifikan

Sumber : Hasil Olah Data menggunakan **Software R** versi 4.12

Pada rancangan percobaan acak Kelompok lengkap (RAKL) menjadi focus kajian uji Hipotesis ANAVA F hanya mengenai faktor perlakuan Merek Ban, sedangkan Jenis mobil hanya sebagai faktor kelompok yang sengaja dilibatkan dengan tujuan mengurangi variasi error (*Sum Square Error*). Berdasarkan Tabel 5 Kolom F.Hitung terlihat perbedaan nilai derajat bebas sumber variasi error pada Tabel ANAVA dua jalur sebelum dan sesudah memasukkan nilai estimasi observasi hilang lebih kecil dengan selisih 2. berikutnya pada kolom F-Hitung untuk baris sumber variasi perlakuan merek Ban terlihat perubahan nilai dari sebelum menyertakan nilai observasi hilang sebesar 1.487 yang ternyata lebih kecil dari nilai F-tabel = 4.347 sehingga uji ANAVA dua jalur dinyatakan tidak signifikan. Dibandingkan dengan F-Hitung setelah menyertakan nilai estimasi observasi hilang Menjadi 7.811 dan nilai ini lebih besar dari F-Tabel = 4.347, sehingga dinyatakan signifikan. Dengan demikian dapat disimpulkan bahwa observasi hilang berpengaruh pada hasil analisis Variansi dua Jalur.

4. KESIMPULAN

Berdasarkan hasil pembahasan terkait Penelitian rancangan percobaan acak kelompok lengkap dari referensi data sekunder mengenai Estimasi Amatan yang hilang dengan faktor perlakuan merek Ban dan jenis mobil sebagai faktor kelompok (Block), perhitungan olah data berbasis *Software R* yang efektif, dapat disimpulkan bahwa;

- 1) Akurasi estimasi observasi hilang dinyatakan masuk ke dalam kriteri sangat baik (Akurat).
- 2) Observasi hilang berpengaruh pada hasil Analisis Variansi dua Jalur terutama untuk faktor perlakuan merk Ban.

REFERENSI

- Gomez, K.A and Gomez, A.A. (2005). Statistical Prosedure for Agriculture research, An international Rice Research Institute, New York : john Wiley and sons.
- Lewis, C.D. (1982). Industrial and business forecasting methods. London: ButterworthsIN
- Juan José Montaña Moreno, Alfonso Palmer Pol, Albert Sesé Abad and Berta Cajal Blasco, 2013“Using the R-MAPE index as a resistant measure of forecast accuracy”,*Psicothema* 2013, Vol. 25, No. 4, 500-506, doi: 10.7334/psicothema2013.23, SSN 0214 - 9915 CODEN PSOTEG Copyright © 2013 Psicothema, www.psicothema.com
- Montgomery, D.C. 2009. Design and Analysis of Experiments: International S eventh Edition USA: John Wiley & Sons.

- Prabowo, H.S (2017), Aplikasi perangkat lunak R pada Uji Hipotesis analisis Variansi satu Faktor, Prosiding SESIOMADIKA, Statistika. 52-59
- Prabowo, H. Soekardi dan Hardyaningwati (2018), Pemodelan Vector Autoregressive pada Peramalan Kadar Beban Cemar COD di Sungai Citarum menggunakan R
- Sariayu Ni Made, Made susilawati dan Kartika Sari (2015), Pendugaan Data Hilang Dengan Metode Yated dan algoritma FM Pada Rancangan Percobaan *Lattice* Seimbang, Ejournal Matematika., vol.4 (2), Mei hal.74-82
- Lewis, C.D., 1982. International and Business Forecasting Methods. London: Butterworths