ISBN:1858-1382

J M S K

Jurnal Matematika, Statistika, & Komputasi
http://journal.unhas.ac.id/index.php/jmsk
Vol. 13, No. 2, 171-181, Januari 2017

ESTIMASI REGRESI *ROBUST* M PADA FAKTORIAL RANCANGAN ACAK LENGKAP YANG MENGANDUNG *OUTLIER*

Siswanto[1], Raupong[2], Annisa[3]

ABSTRAK

Dalam statistik, melakukan suatu percobaan adalah salah satu cara untuk mendapatkan suatu data. Pada perancangan percobaan ada beberapa hal yang membuat suatu data menjadi *outlier*, yaitu tidak berhasilnya suatu pengamatan pada salah satu unit percobaan atau kesalahan dalam pengambilan pengamatan. Pada perancangan percobaan tidak berhasilnya suatu pengamatan pada unit percobaan biasanya disebut data hilang. Dengan kata lain data hilang tersebut biasanya disebut *outlier*. Jika terjadi *outlier* atau suatu pengamatan gagal maka nilai datanya harus ditaksir atau melakukan percobaan ulang. Pada penelitian ini, untuk mengatasi *outlier* pada faktorial Rancangan Acak Lengkap (RAL) digunakan estimasi regresi *robust* M sehingga diperoleh nilai penduga baru yang *resistant* terhadap *outlier*.

Kata Kunci : Outlier, Estimasi Regresi Robust M, Rancangan Percobaan, Faktorial Rancangan Acak Lengkap.

1. Pendahuluan

Dalam statistik, melakukan percobaan adalah salah satu cara untuk mendapatkan suatu data. Terkait dengan data yang diperoleh dari suatu percobaan, ada hal penting yang perlu diperhatikan yaitu adanya *outlier* dalam data. Metode kuadrat terkecil tidak dapat digunakan jika asumsinya tidak terpenuhi karena adanya *outlier* maka diperlukan alternatif metode penduga parameter lain yang dapat mengatasi adanya pencilan dalam data amatan. Metode *Robust* dapat menjadi alternatif pilihan untuk menghasilkan model yang lebih baik dari hasil metode kuadrat terkecil.

Regresi *Robust* terdiri dari 5 metode penduga, yaitu estimasi *robust* M, estimasi *robust least median of square* (LMS), estimasi *robust least trimmed square* (LTS), estimasi *robust S* dan estimasi *robust* MM. Dari ke-5 metode diatas, penulis akan fokus pada metode estimasi *robust* M karena metode ini sering digunakan dan mempunyai keefektifan dalam mengatasi *outlier* karena dapat mengecilkan standar error.

Masalah *outlier* sering dibahas dalam analisis regresi akan tetapi tidak mendapatkan perhatian dalam konteks perancangan percobaan, dalam perancangan percobaan ada beberapa hal yang membuat suatu data menjadi *outlier*, yaitu tidak berhasilnya suatu pengamatan pada salah satu unit percobaan atau kesalahan dalam pengambilan pengamatan. Pada perancangan percobaan tidak berhasilnya suatu pengamatan pada unit percobaan biasanya disebut data hilang. Dengan kata lain data hilang tersebut biasanya disebut *outlier*.

Adapun tujuan dari penulisan ini adalah untuk mengestimasi parameter regresi *robust* menggunakan metode *robust* M dalam mengatasi *outlier*/data hilang pada perancangan percobaan faktorial RAL serta untuk menerapkan regresi *robust* pada perancangan faktorial RAL dengan metode *robust* M dalam mengatasi *outlier*/data hilang.[1]

[1,2,3] *Jurusan Matematika FMIPA Universitas Hasanuddin Makassar, Jl. Perintis Kemerdekaan Km.10 Makassar*

2. Tinjauan Pustaka

2.1. Rancangan Acak Lengkap (RAL)

Rancangan Acak Lengkap (RAL) merupakan rancangan yang paling sederhana jika dibandingkan dengan rancangan-rancangan lainnya. Dalam rancangan ini tidak terdapat lokal kontrol, sehingga sumber keragaman yang diamati hanya perlakuan dan galat. Ada dua hal yang perlu diperhatikan untuk RAL, yaitu :

a) Kecuali perlakuannya, semua (media percobaan dan keadaan-keadaan lingkungan lainnya) harus serba sama atau homogen.
b) Penempatan perlakuan ke dalam satuan-satuan percobaan dilakukan secara acak lengkap, yang artinya perlakukan semua satuan percobaan sebagai satu kesatuan di mana perlakuan-perlakuan (baik yang sama ataupun tidak) ditempatkan ke dalamnya secara acak.

2.1.1 Model Linier dan Penguraian Keragaman Total

Model linier aditif secara umum dari rancangan satu faktor dengan RAL dapat dibedakan menjadi dua yaitu model tetap dan model acak.

Bentuk umum model linier aditif dapat dituliskan sebagai berikut :

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \text{ atau } Y_{ij} = \mu_i + \varepsilon_{ij}; i = 1,2,\dots,t; j = 1,2,\dots,r$$

dimana:

Y_{ij} = Pengamatan pada perlakuan ke$-i$ dan ulangan ke $-j$
μ = Rataan umum
τ_i = Pengaruh perlakuan ke$-i$
ε_{ij} = Error (pengaruh acak) pada perlakuan ke$-i$ dan ulangan ke $-j$

Asumsi untuk model tetap adalah $\sum_{i=1}^{t}\tau_i = 0$, $Var(\varepsilon_{ij}) = \sigma^2 \forall ij$ dan $\varepsilon_{ij} \sim N(0,\sigma^2)$, sedangkan untuk model acak adalah bahwa $E(\tau_i) = 0, Var(\tau_i) = \sigma_\tau^2,\ Var(\varepsilon_{ij}) = \sigma^2 \forall ij$ dan $\varepsilon_{ij} \sim N(0,\sigma^2)$.

2.1.2 Pengujian Hipotesis

Statistik uji $F_{hitung} = KTP/KTG$ mengikuti sebarang F dengan derajat bebas pembilang sebesar $t-1$ dan derajat bebas penyebut sebesar $t(r-1)$. Dengan demikian jika nilai $F_{hitung} > F_{tabel}$ $(F_{a;(t-1);t(r-1)})$ maka hipotesis nol ditolak dan berlaku sebaliknya. Penolakan hipotesis nol (H_0) berimplikasi bahwa perlakuan yang diberikan pada unit-unit percobaan memberikan pengaruh yang nyata terhadap respon yang diamati.

2.2 Model Linier Faktorial Rancangan Acak Lengkap (RAL)

Percobaan faktorial adalah suatu percobaan yang perlakuannya terdiri atas semua kemungkinan kombinasi taraf dari beberapa faktor. Percobaan faktorial dapat diterapkan secara langsung terhadap seluruh unit-unit percobaan jika unit percobannya relatif homogen Rancangan seperti ini disebut rancangan faktorial dengan rancangan dasar RAL atau lebih jauh disebut faktorial RAL. Bentuk umumnya dapat dirumuskan sebagai berikut :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\gamma)_{ij} + \varepsilon_{ijk}; i = 1,2,\dots,\alpha; j = 1,2,\dots,b; k = 1,2,\dots,r \quad (1)$$

Keterangan :

Y_{ijk} = Respon perlakuan pada taraf ke$-i$ faktor A, taraf ke$-j$ faktor B dan ulangan ke$-k$
μ = Nilai tengah populasi (rata-rata yang sesungguhnya)
α_i = Pengaruh utama taraf ke$-i$ dari faktor A
β_j = Pengaruh utama taraf ke$-j$ dari faktor B
$(\gamma)_{ij}$ = Pengaruh interaksi taraf ke$-i$ faktor A dan taraf ke$-j$ faktor B
ε_{ijk} = Pengaruh galat pada taraf ke$-i$ faktor A, taraf ke$-j$ faktor B dan ulangan ke$-k$

a, b, r = Banyaknya taraf faktor A, taraf faktor B dan ulangan.

2.3. Pendekatan *Robust* M

Model *Robust* linier berguna untuk memfilter hubungan linier ketika variasi acak pada data yang tidak normal atau ketika data mengandung *outlier* yang signifikan. Metode ini penting untuk menganalisa data yang dipengaruhi oleh *outlier* sehingga dihasilkan model yang *robust* atau *resistant* terhadap *outlier*.

2.3.1. Kuadrat Terkecil

Prosedur metode kuadrat terkecil adalah mendapatkan penduga b_0 dan b_1 yang menjadikan jumlah kuadrat error, yaitu $\sum_{i=1}^{n} \varepsilon_i^2$ minimum, yakni :

1. Membentuk $\sum_{i=1}^{n} \varepsilon_i^2$ dan fungsi b_0 dan b_1
$$S = f(b_0, b_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$$
2. Mendiferensialkan S terhadap b_0 dan b_1 sehingga diperoleh
$$\sum_{i=1}^{n} \varepsilon_i X_i = 0 \qquad (2)$$

Persamaan (2) dapat menggambarkan pengaruh yang disebabkan oleh titik eksperimen dengan residual yang ekstrim.

2.3.2. Estimasi M

Estimasi Mmerupakan metode regresi *robust* yang sering digunakan. Pendugaan parameter menggunakan metode ini disebut juga *Iteratively Reweighted Least Squares (IRLS)*. Metode ini menggunakan fungsi Huber berikut:

$$\psi(e^*) = \begin{cases} e_i^* \text{untuk} |e_i^*| \le r \\ r \text{ untuk} e_i^* > r \\ -r \text{ untuk} e_i^* < -r \end{cases}$$

Berikut ini adalah prosedur pendugaan :

1. Dihitung penduga $\boldsymbol{\beta}$, dinotasikan $\boldsymbol{b}$ menggunakan *least square*, sehingga didapatkan $\hat{y}_{i,0}$ dan $\varepsilon_{i.0} = y_i - \hat{y}_{i,0}$, $(i = 1,2, \dots, n)$yang diperlakukan sebagai nilai awal ($y_i$adalah hasil eksperimen).
2. Dari nilai-nilai residual ini dihitung $\hat{\sigma}_0$, dan pembobot awal $w_{i,0} = \dfrac{\psi(\varepsilon_{i,0}^*)}{(\varepsilon_{i,0}^*)}$. Nilai $\psi(\varepsilon_i^*)$dihitung sesuai fungsi Huber, dan $\varepsilon_{i,0}^* = \frac{\varepsilon_{i,0}}{\hat{\sigma}_0}$.
3. Disusun matriks pembobot berupa matriks diagonal dengan elemen $w_{1,0}; w_{2,0}; \dots; w_{n,0}$dinamai W_0.
4. Dihitung penduga koefisien regresi $\boldsymbol{b}_{Robust\ ke\ 1} = (X^T W_0 X)^{-1} X^T W_0 Y$
5. Dengan menggunakan $\mathbf{b}_{Robust \text{ ke } 1}$ dihitung pula
$$\sum_{i=1}^{n} | y_i - \hat{y}_{i,1} | \quad \text{atau} \quad \sum_{i=1}^{n} | \varepsilon_{i.1} | .$$
6. Selanjutnya langkah 2 sampai dengan 5 diulang sampai didapatkan $\sum_{i=1}^{n} | \varepsilon_{i.m} |$konvergen.

3. Hasil dan Pembahasan

Dalam perancangan percobaan ada beberapa hal yang membuat suatu data menjadi *outlier*, yaitu tidak berhasilnya suatu pengamatan pada salah satu unit percobaan atau kesalahan dalam pengambilan pengamatan. Pada perancangan percobaan tidak berhasilnya suatu pengamatan pada unit percobaan biasanya disebut data hilang. Dengan kata lain data hilang tersebut biasanya disebut *outlier*. Jika terjadi *outlier* atau suatu percobaan gagal maka metode ANAVA tidak dapat digunakan dalam faktorial RAL, untuk itu nilai datanya harus ditaksir atau melakukan percobaan (*experiment*) ulang. Adapun model rancangan faktorial dalam Rancangan Acak Lengkap yang disingkat faktorial RAL adalah :

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\gamma)_{ij} + \varepsilon_{ijk}\,;\, i = 1,2,\dots,a\,;\, j = 1,2,\dots,b\,;\, k = 1,2,\dots,r \quad (3)$$

dimana :

Y_{ijk} = Respon perlakuan pada taraf ke$-i$ faktor A, taraf ke$-j$ faktor B dan ulangan ke$-k$
μ = Nilai tengah populasi (rata-rata yang sesungguhnya)
α_i = Pengaruh utama taraf ke$-i$ dari faktor A
β_j = Pengaruh utama taraf ke$-j$ dari faktor B
$(\gamma)_{ij}$ = Pengaruh interaksi taraf ke$-i$ faktor A dan taraf ke$-j$ faktor B
ε_{ijk} = Pengaruh galat pada taraf ke$-i$ faktor A, taraf ke$-j$ faktor B dan ulangan ke$-k$
a, b, r = Banyaknya taraf faktor A, taraf faktor B dan ulangan.

Bentuk faktorial RAL pada Persamaan (3) diubah ke dalam bentuk regresi linier berganda dengan menggunakan perkalian kronecker sebagai contoh yaitu : $a = 2\,, b = 3$ dan $r = 3$

Model regresi dari Persamaan (3) adalah :

$$\boldsymbol{Y} = \mathbf{1}_{18}\mu + (\boldsymbol{I}_2 \otimes \mathbf{1}_9)\boldsymbol{\alpha} + \begin{bmatrix} \boldsymbol{I}_3 \otimes \mathbf{1}_3 \\ \boldsymbol{I}_3 \otimes \mathbf{1}_3 \end{bmatrix}\boldsymbol{\beta} + (\boldsymbol{I}_6 \otimes \mathbf{1}_3)\boldsymbol{\gamma} + \boldsymbol{I}_{18}\boldsymbol{e}$$

$$\begin{aligned}\boldsymbol{Y} = {} & (\mathbf{1}_2 \otimes \mathbf{1}_3 \otimes \mathbf{1}_3)\mu + (\boldsymbol{I}_2 \otimes \mathbf{1}_3 \otimes \mathbf{1}_3)\boldsymbol{\alpha} + (\mathbf{1}_2 \otimes \boldsymbol{I}_3 \otimes \mathbf{1}_3)\boldsymbol{\beta} + (\boldsymbol{I}_2 \otimes \boldsymbol{I}_3 \otimes \mathbf{1}_3)\boldsymbol{\gamma} \\ & + (\boldsymbol{I}_2 \otimes \boldsymbol{I}_3 \otimes \boldsymbol{I}_3)\boldsymbol{e}\end{aligned}$$

Atau dapat dituliskan dalam bentuk umum sebagai berikut :

$$\begin{aligned}\boldsymbol{Y} = {} & (\mathbf{1}_a \otimes \mathbf{1}_b \otimes \mathbf{1}_r)\mu + (\boldsymbol{I}_a \otimes \mathbf{1}_b \otimes \mathbf{1}_r)\boldsymbol{\alpha} + (\mathbf{1}_a \otimes \boldsymbol{I}_b \otimes \mathbf{1}_r)\boldsymbol{\beta} + (\boldsymbol{I}_a \otimes \boldsymbol{I}_b \otimes \mathbf{1}_r)\boldsymbol{\gamma} \\ & + (\boldsymbol{I}_a \otimes \boldsymbol{I}_b \otimes \boldsymbol{I}_r)\boldsymbol{e}\end{aligned}$$

Dengan kata lain model linier umum bagi suatu rancangan percobaan dapat dituliskan dalam persamaan berikut:

$$\boldsymbol{Y} = \boldsymbol{X\theta} + \boldsymbol{\varepsilon} \quad (4)$$

dimana :

$\boldsymbol{Y}$: Vektor respon perlakuan berukuran $abr \times 1$
$\boldsymbol{X}$: Matriks berukuran $abr \times (1 + a + b + ab)$
$\boldsymbol{\theta}$: Vektor parameter model faktorial RAL berukuran $(1 + a + b + ab) \times 1$
$\boldsymbol{\varepsilon}$: Vektor galat berukuran $abr \times 1$
a, b, r : Banyaknya taraf faktor A, taraf faktor B dan ulangan

3.1. Metode Kuadrat Terkecil (*Ordinary Least Squares* = *OLS*) pada Faktorial RAL

Untuk menentukan penduga bagi parameter pada rancangan percobaan terdapat asumsi-asumsi yang harus dipenuhi yaitu :

Untuk model tetap : $\sum_{i=1}\alpha_i = 0, \sum_{j=1}\beta_j = 0, \sum_{i=1}(\gamma)_{ij} = 0, \sum_{j=1}(\gamma)_{ij} = 0,$
Untuk model acak : $\alpha_i \sim N(0, \sigma_\alpha^2);\ \beta_j \sim N\left(0, \sigma_\beta^2\right);\ (\gamma)_{ij} \sim N\left(0, \sigma_\gamma^2\right)$
Untuk memperoleh penduga bagi parameter $, \alpha_i,\ \beta_j$ dan $(\gamma)_{ij}$ maka Persamaan (3) diubah kedalam bentuk sebagai berikut :

$$\varepsilon_{ijk} = Y_{ijk} - \mu - \alpha_i - \beta_j - (\gamma)_{ij} \tag{5}$$

Persamaan (5) dikuadratkan dan dijumlahkan sehingga diperoleh :

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\varepsilon_{ijk}^2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(Y_{ijk} - \mu - \alpha_i - \beta_j - (\gamma)_{ij})^2 = R$$

Penduga $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j$ dan $(\hat{\gamma})_j$ bagi parameter μ,α_i, β_j dan $(\gamma)_{ij}$ diperoleh dengan meminimumkan nilai R dengan metode kuadrat terkecil yaitu melakukan turunan parsial pertama R terhadap parameter kemudian disamakan dengan nol sehingga diperoleh :

Penduga untuk parameter μ

$$\frac{\partial R}{\partial \mu} = 2\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \alpha_i - \beta_j - (\gamma)_{ij}\right)(-1) = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \alpha_i - \beta_j - (\gamma)_{ij}\right) = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} - abr\,\hat{\mu} - br\sum_{i=1}^{a}\alpha_i - ar\sum_{j=1}^{b}\beta_j - r\sum_{i=1}^{a}\sum_{j=1}^{b}(\gamma)_{ij} = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} - abr\,\hat{\mu} = 0$$

$$abr\hat{\mu} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} \ => \hat{\mu} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk}}{abr} = \bar{Y}_{\ldots} \tag{6}$$

Penduga untuk parameter $\boldsymbol{\alpha_i}$

$$\frac{\partial R}{\partial \alpha_i} = 2\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \beta_j - (\gamma)_{ij}\right)(-1) = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \beta_j - (\gamma)_{ij}\right) = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} - br\,\hat{\mu} - br\hat{\alpha}_i - r\sum_{j=1}^{b}\beta_j - r\sum_{j=1}^{b}(\gamma)_{ij} = 0$$

$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} - br\,\hat{\mu} - br\hat{\alpha}_i = 0$$

$$br\hat{\alpha}_i = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} - br\,\hat{\mu} => \hat{\alpha}_i = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}Y_{ijk} - br\,\hat{\mu}}{br} = \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\ldots} \tag{7}$$

Penduga untuk parameter $\boldsymbol{\beta_j}$

$$\frac{\partial R}{\partial \beta_j} = 2\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\gamma)_{ij}\right)(-1) = 0$$
$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\gamma)_{ij}\right) = 0$$
$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - ar\,\hat{\mu} - r\sum_{i=1}^{a}\hat{\alpha}_i - ar\hat{\beta}_j - r\sum_{i=1}^{a}(\gamma)_{ij} = 0$$
$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - ar\,\hat{\mu} - ar\hat{\beta}_j = 0$$
$$ar\hat{\beta}_j = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - ar\,\hat{\mu} \;=>\; \hat{\beta}_j = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - ar\,\hat{\mu}}{ar}$$
$$= \bar{Y}_{.j.} - \bar{Y}_{...} \qquad (8)$$

Penduga untuk parameter $(\boldsymbol{\gamma})_{ij}$

$$\frac{\partial R}{\partial (\gamma)_{ij}} = 2\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\hat{\gamma})_{ij}\right)(-1) = 0$$
$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}\left(Y_{ijk} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - (\hat{\gamma})_{ij}\right) = 0$$
$$\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - r\hat{\mu} - r\hat{\alpha}_i - r\hat{\beta}_j - r(\hat{\gamma})_{ij} = 0$$
$$r(\hat{\gamma})_{ij} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - r\hat{\mu} - r\hat{\alpha}_i - r\hat{\beta}_j$$
$$(\hat{\gamma})_{ij} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r} Y_{ijk} - r\hat{\mu} - r\hat{\alpha}_i - r\hat{\beta}_j}{r}$$
$$(\hat{\gamma})_{ij} = \bar{Y}_{ij.} - \bar{Y}_{...} - (\bar{Y}_{i..} - \bar{Y}_{...}) - (\bar{Y}_{.j.} - \bar{Y}_{...})$$
$$(\hat{\gamma})_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...} \qquad (9)$$

3.2. Metode Kuadrat Terkecil (*Ordinary Least Squares* = *OLS*) pada Regresi

Ide dasar dari metode kuadrat terkecil adalah dengan membuat $\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X\theta}$sekecil mungkin yaitu dengan meminimalkan $\boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\theta})^t(\boldsymbol{Y} - \boldsymbol{X\theta})$.Untuk meminimalkannya dapat diperoleh dengan melakukan turunan parsial pertama terhadap $\boldsymbol{\theta}$kemudian menyamakannya dengan nol maka,

$$\boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon} = (\boldsymbol{Y} - \boldsymbol{X\theta})^t(\boldsymbol{Y} - \boldsymbol{X\theta})$$
$$= (\boldsymbol{Y}^t - \boldsymbol{X}^t\boldsymbol{\theta}^t)(\boldsymbol{Y} - \boldsymbol{X\theta})$$
$$= \boldsymbol{Y}^t\boldsymbol{Y} - \boldsymbol{Y}^t\boldsymbol{X\theta} - \boldsymbol{X}^t\boldsymbol{\theta}^t\boldsymbol{Y} + \boldsymbol{X}^t\boldsymbol{\theta}^t\boldsymbol{X\theta}$$
$$= \boldsymbol{Y}^t\boldsymbol{Y} - 2\boldsymbol{X}^t\boldsymbol{\theta}^t\boldsymbol{Y} + \boldsymbol{X}^t\boldsymbol{\theta}^t\boldsymbol{X\theta}$$
$$\frac{d\boldsymbol{\varepsilon}^t\boldsymbol{\varepsilon}}{d\boldsymbol{\theta}} = -2\boldsymbol{X}^t\boldsymbol{Y} + 2\boldsymbol{X}^t\boldsymbol{X\theta} = 0$$

$\boldsymbol{X}^t\boldsymbol{X\theta} = \boldsymbol{X}^t\boldsymbol{Y}$ masing-masing ruas dikalikan $(\boldsymbol{X}^t\boldsymbol{X})^-$

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{X}^t\boldsymbol{X})^-\boldsymbol{X}^t\boldsymbol{Y}$$

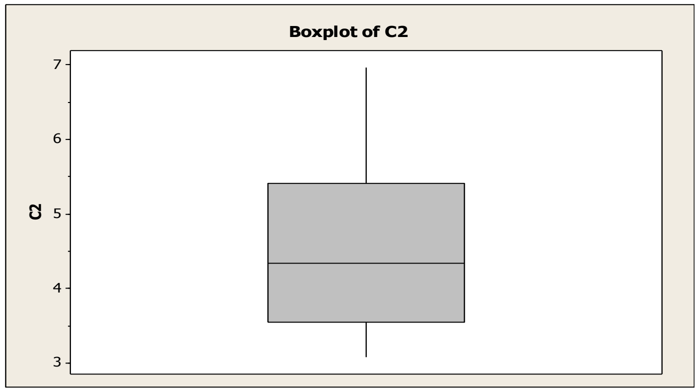
Persamaan (4) yang diperoleh dengan merekonstruksi bentuk umum faktorial RAL ke bentuk regresi berganda dalam bentuk matriks memiliki penduga sebagai berikut :

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\theta}}$$

3.3. Penerapan Regresi *Robust* pada Rancangan Faktorial RAL

3.3.1 Rancangan Faktorial RAL yang Tidak Mengandung *Outlier*/Data Hilang

Salah satu cara yang digunakan untuk mengidentifikasi *outlier* adalah metode *boxplot*. Untuk mengidentifikasi data RAL digunakan metode grafis *boxplot* dengan Minitab 16 hasilnya adalah sebagai berikut :



Gambar 1 Data pengamatan nilai kadar air briket arang yang tidak mengandung *outlier*

Berdasarkan Gambar 1 diketahui bahwa tidak terdapat titik yang terdapat di luar kotak *boxplot* sehingga dapat disimpulkan bahwa tidak mengalami *outlier*.

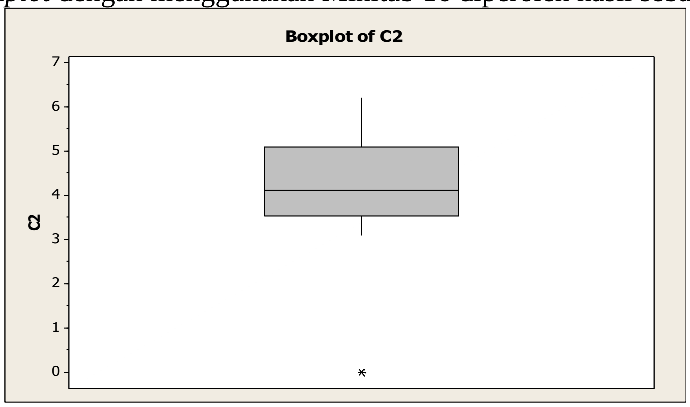
Tabel 1 Hasil analisis variansi data pengamatan nilai kadar air briket arang

Sumber Keragaman	Derajat Bebas (db)	Jumlah Kuadrat (JK)	Kuadrat Tengah (KT)	F Hitung	F Tabel ($F_{0,05}$)
Perlakuan	5	20,4	4,08		
– A	1	6,4	6,4	51,2*	4,75
– B	2	12,5	6,25	50*	3,89
– AB	2	1,4	0,7	5,6*	3,89
Galat	12	1,5	0,125		
Total	17	21,9			

Berdasarkan Tabel 1 pengaruh utama faktor jenis kayu memiliki ($F_{hitung} = 51{,}2$) $>$ nilai $F_{0,05(db1=1,db2=12)} = 4{,}75$ sehingga H_0 ditolak pada taraf α (signifikan) berarti faktor jenis kayu berpengaruh terhadap respon yang diamati. Pengaruh utama faktor Persentase Penambahan Bahan Aditif Lilin Lebah memiliki ($F_{hitung} = 50$) $>$ nilai $F_{0,05(db1=2,db2=12)} = 3{,}89$ sehingga H_0 ditolak pada taraf α (signifikan) berarti faktor persentase penambahan bahan aditif lilin lebah berpengaruh terhadap respon yang diamati. Pengaruh sederhana (interaksi) faktor jenis kayu dengan faktor persentase penambahan bahan aditif lilin lebah memiliki ($F_{hitung} = 5{,}6$) $>$ nilai $F_{0,05(db1=2,db2=12)} = 3{,}89$ sehingga H_0 ditolak pada taraf α (signifikan) berarti interaksi faktor jenis kayu dengan faktor persentase penambahan bahan aditif lilin lebah adalah berpengaruh.

3.3.2 Rancangan Faktorial RAL yang Mengandung *Outlier*/Data Hilang

Dalam perancangan percobaan ada beberapa hal yang membuat suatu data tersebut menjadi *outlier*, yaitu tidak berhasilnya suatu pengamatan pada salah satu unit percobaan atau kesalahan dalam pengambilan pengamatan (Suntoyo Yitnosumarto.1991). Pada data perancangan percobaan tidak berhasilnya unit pengamatan dapat dituliskan nol,yang menyebabkan pengaruh utama atau pengaruh sederhana (interaksi) menjadi tidak signifikan. Identifikasi *outlier* metode grafis, yaitu metode *boxplot* dengan menggunakan Minitab 16 diperoleh hasil sebagai berikut :



Gambar 2 Data pengamatan nilai kadar air briket arang yang mengandung *outlier*

Berdasarkan Gambar 2 dapat disimpulkan bahwa titik yang terdapat di luar kotak *boxplot* merupakan *outlier*/data hilang.

Tabel 2 Hasil analisis variansi data pengamatan nilai kadar air briket arang

Sumber Keragaman	**Derajat Bebas (db)**	**Jumlah Kuadrat (JK)**	**Kuadrat Tengah (KT)**	**F Hitung**	**F Tabel ($F_{0,05}$)**
Perlakuan	5	8,6	1,72		
– A	1	0,8	0,8	0,375	4,75
– B	2	4,8	2,4	1,125	3,89
– AB	2	3,0	1,5	0,703125	3,89
Galat	12	25,6	2,133		
Total	17	34,2			

Berdasarkan Tabel 2 pengaruh utama faktor jenis kayu memiliki ($F_{hitung} = 0{,}375$) $<$ nilai $F_{0,05(db1=1,db2=12)} = 4{,}75$ sehingga H_0 diterima pada taraf α (signifikan) berarti faktor jenis kayu tidak berpengaruh terhadap respon yang diamati. Pengaruh utama faktor Persentase Penambahan Bahan Aditif Lilin Lebah memiliki ($F_{hitung} = 1{,}125$) <nilai $F_{0,05(db1=2,db2=12)} = 3{,}89$ sehingga H_0 diterima pada taraf α (signifikan) berarti faktor persentase penambahan bahan aditif lilin lebah tidak berpengaruh terhadap respon yang diamati. Pengaruh sederhana (interaksi) faktor jenis kayu dengan faktor persentase penambahan bahan aditif lilin lebah memiliki ($F_{hitung} = 0{,}703125$) $<$ nilai $F_{0,05(db1=2,db2=12)} = 3{,}89$ sehingga H_0 diterima pada taraf α (signifikan) berarti interaksi faktor jenis kayu dengan faktor persentase penambahan bahan aditif lilin lebah tidak berpengaruh.

3.4 Regresi *Robust* Estimasi M pada Data Faktorial RAL

Berdasarkan hasil identifikasi *outlier* disimpulkan terdapat *outlier* pada data. Selanjutnya, untuk mengatasi permasalahan tersebut digunakan regresi *robust* dengan estimasi M.

Tabel 3 Hasil iterasi memperoleh nilai koefesien parameter*Robust*

Iterasi						
OLS	$\boldsymbol{b}_{robust\,1}$	$\boldsymbol{b}_{robust\,2}$	$\boldsymbol{b}_{robust\,3}$	$\boldsymbol{b}_{robust\,4}$	$\boldsymbol{b}_{robust\,5}$	$\boldsymbol{b}_{robust\,6}$

μ	2.085	2.152	2.195	2.205	2.208	2.208	2.208
α_1	0.884	0.816	0.773	0.763	0.761	0.761	0.761
α_2	1.201	1.336	1.422	1.443	1.447	1.447	1.447
β_1	0.806	1.009	1.138	1.167	1.175	1.175	1.175
β_2	1.051	0.983	0.94	0.93	0.928	0.928	0.928
β_3	0.228	0.159	0.117	0.107	0.105	0.105	0.105
γ_{11}	0.845	0.642	0.514	0.483	0.477	0.477	0.477
γ_{12}	-0.026	0.041	0.085	0.095	0.097	0.097	0.097
γ_{13}	0.064	0.131	0.175	0.185	0.187	0.187	0.187
γ_{21}	-0.039	0.366	0.624	0.686	0.698	0.698	0.698
γ_{22}	1.077	0.941	0.856	0.835	0.831	0.831	0.831
γ_{23}	0.163	0.028	-0.058	-0.078	-0.082	-0.082	-0.082
$\sum_{i=1}^{n}\lvert\varepsilon_{i,m}\rvert$	11,42	10,60	10,09	9,97	9,95	9,95	9,95

Berdasarkan Tabel 3 terlihat bahwa selisih $\sum_{i=1}^{n}\lvert\varepsilon_{i,m}\rvert$ $\boldsymbol{b}_{robust\ 4}$ dan $\boldsymbol{b}_{robust\ 5}$ sama dengan nol. Hal ini menunjukkan bahwa estimasi parameter telah konvergen. Setelah mendapatkan penduga yang resistant terhadap *outlier*, nilai tersebut digunakan pada persamaan $\widehat{\boldsymbol{Y}} = \boldsymbol{X}\widehat{\boldsymbol{\theta}}$sehingga diperoleh nilai $\widehat{\boldsymbol{Y}}$sebagai berikut :

$\hat{Y}_{111} = 4{,}62$	$\hat{Y}_{131} = 3{,}26$	$\hat{Y}_{221} = 5{,}41$
$\hat{Y}_{112} = 4{,}62$	$\hat{Y}_{132} = 3{,}26$	$\hat{Y}_{222} = 5{,}41$
$\hat{Y}_{113} = 4{,}62$	$\hat{Y}_{133} = 3{,}26$	$\hat{Y}_{223} = 5{,}41$
$\hat{Y}_{121} = 3{,}99$	$\hat{Y}_{211} = 5{,}53$	$\hat{Y}_{231} = 3{,}68$
$\hat{Y}_{122} = 3{,}99$	$\hat{Y}_{212} = 5{,}53$	$\hat{Y}_{232} = 3{,}68$
$\hat{Y}_{123} = 3{,}99$	$\hat{Y}_{213} = 5{,}53$	$\hat{Y}_{233} = 3{,}68$

Berdasarkan nilai pendugaan yang telah diperoleh diatas, akan dicari ANAVA pada data yang mengandung *outlier* dengan mengganti nilai pada respon Y_{211}yang bernilai nol dengan nilai taksiran $\hat{Y}_{211} = 5{,}53$.

Tabel 4 Hasil analisis variansi data pengamatan nilai kadar air briket arang

Sumber Keragaman	**Derajat Bebas (db)**	**Jumlah Kuadrat (JK)**	**Kuadrat Tengah (KT)**	**F Hitung**	**F Tabel ($\boldsymbol{F_{0,05}}$)**
Perlakuan	5	15,8	3,16		
– A	1	4,8	4,8	48*	4,75
– B	2	10,1	5,05	50,5*	3,89
– AB	2	0,9	0,45	4,5*	3,89
Galat	12	1,2	0,1		
Total	17	17,0			

Berdasarkan Tabel 4.4 pengaruh utama faktor jenis kayu memiliki ($F_{hitung} = 48$) $>$ nilai $F_{0,05(db1=1,db2=12)} = 4{,}75$ sehingga H_0 ditolak pada taraf α (signifikan) yang berarti bahwa faktor jenis kayu berpengaruh terhadap respon yang diamati. Pengaruh utama faktor Persentase Penambahan Bahan Aditif Lilin Lebah memiliki ($F_{hitung} = 50{,}5$) $>$ nilai $F_{0,05(db1=2,db2=12)} = 3{,}89$ sehingga H_0 ditolak pada taraf α (signifikan) yang berarti bahwa faktor persentase penambahan bahan aditif lilin lebah berpengaruh terhadap respon yang diamati. Pengaruh sederhana (interaksi) faktor jenis kayu dengan faktor persentase penambahan bahan aditif lilin lebah memiliki $(F_{hitung} = 4{,}5) >$ nilai $F_{0,05(db1=2,db2=12)} = 3{,}89$ sehingga H_0 ditolak pada taraf α (signifikan) yang berarti bahwa interaksi faktor jenis kayu dengan faktor persentase penambahan bahan aditif lilin lebah adalah berpengaruh.

4.1 KESIMPULAN

Berdasarkan hasil analisis yang telah dilakukan maka dapat diambil beberapa kesimpulan sebagai berikut:

1. Untuk mengatasi *outlier/*data hilangpada faktorial Rancangan Acak Lengkap digunakan metode *robust* M dengan menghitung penduga koefisien regresi yaitu $\boldsymbol{b}_{Robust\ ke\ 1} = (X^T W_0 X)^{-1} X^T W_0 Y$
   Dimana W_0 dengan elemen $w_{1,0}; w_{2,0}; \ldots; w_{n,0}$ adalah matriks pembobot berupa matriks diagonal, yang diperoleh dengan $w_{i,0} = \frac{\psi(\varepsilon_{i,0}^*)}{(\varepsilon_{i,0}^*)}$. Nilai $\psi(\varepsilon_{i,0}^*)$ dihitung sesuai fungsi Huber dan $(\varepsilon_{i,0}^*) = \frac{\varepsilon_{i,0}}{\hat{\sigma}_0}$ sehingga diperoleh nilai penduga yang *resistant* terhadap *outlier*.
2. Dalam kasus "Sifat Briket Arang dari Kayu Bayam (*Intsia bijuga*) dan Kayu Palapi (*Heritiera sp*) dengan Penambahan Berbagai Konsentrasi Lilin Lebah sebagai Bahan Aditif" merupakan data hilang pada pengamatan adalah faktor A taraf ke 2 yaitu kayu Palapi, faktor B (persentase lilin lebah) taraf ke 1 yaitu persentase 5% dan ulangan pertama bila unit percobaan tersebut dibuat nol menyebabkan pengaruh utama faktor A atau pengaruh utama faktor B atau pengaruh sederhana (interaksi) menjadi tidak signifikan sehingga perlu diduga dengan metode *robust* M, setelah diduga dengan metode *robust* M pengaruh utama faktor A atau pengaruh utama faktor B atau pengaruh sederhana (interaksi) menjadi signifikan.

4.2 SARAN

Penelitian ini membahas tentang penggunaan regresi *robust* pada data faktorial RAL yang mengandung *outlier* dengan Metode Estimasi M. Untuk penelitian selanjutnya dapat dilakukan penelitian atau kajian lebih sifat-sifat teoritis pada beberapa estimator *robust* pada rancangan percobaan lainnya untuk mengatasi data yang mengandung *outlier*.

DAFTAR PUSTAKA

[1] Cook,R.D.,1977. "Detection of Influential Observation in Linear Regression", *Technometrics*, 19: 15-18.

[2] Draper, N.R dan Harry S. 1992. *Analisis Regresi Terapan (edisi kedua)*. Alih Bahasa: Bambang Sumantri. Jakarta : PT. Gramedia Pustaka Utama.

[3] Gaspersz, Vincent. 1991. *Metode Perancangan Percobaan*. Bandung : Armico.

[4] John, J.A., 1978. "Outliers in Factorial Experiments", *J. Series C,* 27: 111-119.

[5] Mattjik, A. A & Sumertajaya, I. M. 2002. *Perancangan Percobaan dengan Aplikasi SAS dan MINITAB Jilid I Edisi Kedua*. Bogor : IPB Press.

[6] Montgomery, D. C. 2003. *Design And Analysis of Experiments 5th Edition*. Singapore: John Wiley & Sons.

[7] Nurbaity. 1999. *Sifat Briket Arang dari Kayu Bayam (Intsia bijuga) dan Kayu Palapi (Heritiera sp) dengan Pembahasan Berbagai Konsentrasi Lilin Lebah sebagai Bahan Aditif*. Skripsi : Universitas Hasanuddin, Ujung Pandang.

[8] Raymond, J.C., 1979. *Robust Methods for Factorial Experiments with Outliers*, University of North Carolina at Chapel Hill.

[9] Shayle, R.S., George, C., and Charles, E.Mc. *Variance Components*. John Wiley & Sons, INC. 1992.

[10] Yitnosumarto, S. *Percobaan perancangan, Analisis dan Interpretasinya*. Penerbit PT. Gramedia Pustaka Umum. Jakarta, 1991