

Aplikasi Splus dalam Model Regresi dengan kesalahan Pengukuran Menggunakan Metode ICM

Hartatik

Program Studi D3 Teknik Informatika, Universitas Sebelas Maret
Surakarta

Jl. Ir Sutami, Surakarta

Email: hartatik89 @ymail.com

Abstract

Measurement data are often assumed to be without error is very rare indeed. Generally, any data obtained would contain an error due to measurement / factor measurement error (ME). For example in the regression modeling if X is a random variable or with measurement error. Then the complex calculations will not be separated from the role of computers and technology. As well as to estimate the following model, given the data (X_i, Y_i) , the regression models are

$$Y_i = g(X_i) + \varepsilon_i$$

where X_i is the i -th element of the predictor variable X and Y_i is the i -th element of the response variable Y . Variable X which is a predictor of the outcome variable is usually a certain constant observation, however, X is generally found that a random variable or variables where the value is not a fixed constant. For that reason in this case is called the regression models with regression models with measurement error. There are two methods, namely parametric and Nonparametric Approach. Parametric approach used Ordinary Least Square (OLS), and for Nonparametric measurement errors are ignored when used B-spline method and the measurement error is not negligible when used methods Iterative Conditional Modes (ICM). In the previous study, examined the effect of measurement errors on the estimation of the model. Furthermore, in this study the process of calculating a regression model using the S-Plus2000 Software as a computational tool to estimate the regression model. Data will be generated using the program will be carried out in this study, which in turn is used to estimate program regression models with measurement error.

Keywords: *Iterative Conditional Modes (ICM), regression, measurement error, Software S-Plus.*

Pendahuluan

Latar Belakang

Data hasil pengukuran atau pengamatan seringkali langsung diambil sebagai hasil pengamatan yang digunakan untuk estimasi model, demikian halnya didalam model regresi dengan spline. Data hasil pengukuran umumnya masih perlu dikaji ulang, dikarenakan adanya faktor *measurement error* (ME). Banyak faktor yang mempengaruhi hasil pengukuran suatu data, diantaranya : faktor alat, proses pengambilan data yang benar, faktor human, dan faktor alam. Faktor faktor tersebut, bisa

menyebabkan hasil pengukuran mengandung suatu kesalahan pengukuran (*measurement error*). Namun Sering dan kebanyakan peneliti mengambil dan mengasumsikan data sebagai nilai tetap dan mempercayai bahwa data tersebut sudah benar, tanpa melakukan pengecekan akan seberapa besar kesalahan pengukuran bisa diabaikan. Model regresi apabila kesalahan pengukuran itu signifikan tentunya akan berbeda apabila tidak ada kesalahan pengukuran. Misal

$$Y = g(X) + \varepsilon \quad (\text{Pers 1})$$

di mana X adalah variabel prediktor, Y variabel respon, g adalah suatu fungsi tertentu, dan ε adalah sesatan random independent dengan mean nol dan variansi σ_ε^2 .

Variabel X yang merupakan variabel prediktor dari hasil pengamatan biasanya diasumsikan sebagai variabel tetap (*fixed variable*). Namun kenyataannya, sering dijumpai X yang bukan *fixed variable* tetapi variabel random atau variabel X diukur dengan kesalahan (*error in variable*). Namun hal ini sering diabaikan untuk alasan praktis dan kemudahan perhitungan. Namun apabila hal ini dilakukan berulang dan pada variabel dengan tingkat resiko tinggi maka akibatnya akan fatal terhadap pengambilan keputusan. Kasus ini bisa dijumpai diantaranya dalam masalah epidemiologi, geologi, dan survival.

Kesalahan pengukuran (*measurement error*) adalah kesalahan yang muncul manakala suatu nilai dicatat tidak persis sama dengan nilai sebenarnya dalam kaitan dengan suatu proses pengukuran. Sehingga berkaitan dengan definisi ini, ada 3 variabel di dalam model kesalahan pengukuran, yaitu

$$W = X + U \quad (\text{Pers 2})$$

dengan W adalah variabel yang menyatakan hasil pengamatan yang disebut dengan variabel pengganti (*surrogate*), X adalah variabel prediktor yang tidak teramati (*latent variable*), dan U adalah variabel kesalahan pengukuran yang diasumsikan normal independen dengan mean 0 dan variansi σ_u^2 .

Berdasarkan model regresi (1) dan (2), ada dua pendekatan digunakan untuk menduga kurva regresi yaitu metode regresi parametrik dan nonparametrik. Tujuan dari penelitian ini adalah untuk melakukan langkah langkah estimasi model regresi dengan menggunakan software S-Plus, yaitu dengan metode ICM.

B-Spline

Model regresi nonparametrik, jika terdapat spline dengan orde m dan kumpulan knots yang memenuhi $a < \xi_1 < \dots < \xi_k < b$. Dapat didefinisikan sejumlah $2m$ knot tambahan $\xi_{-(m-1)}, \dots, \xi_1, \xi_0, \xi_{k+1}, \dots, \xi_{k+m}$ dimana $\xi_{-(m-1)}, \dots, \xi_0 = a$ dan $\xi_{k+1}, \dots, \xi_{k+m} = b$.

B-spline orde m yang sesuai untuk knot $\xi_1, \dots, \xi_k$ dinyatakan dengan

$$B_{j,m}(X) = \frac{X - \xi_j}{\xi_{j+m-1} - \xi_j} B_{j,m-1}(X) + \frac{\xi_{j+m} - X}{\xi_{j+m} - \xi_{j+1}} B_{j+1,m-1}(X), \quad (\text{Pers 3})$$

dengan $B_{j,l}(X) = \begin{cases} 1, & X \in [\xi_j, \xi_{j+1}) \\ 0, & \text{untuk yang lain.} \end{cases}$

Misal diasumsikan bahwa

$$\hat{g}(X) = \sum_{j=1}^p B_{j,m}(X) \beta_j,$$

dengan $B_1(X), \dots, B_p(X)$, adalah basis untuk vektor spline berderajat m dengan titik knot $\xi_1, \dots, \xi_k$ dan $p=m+k$. Maka

$$S(\mathbf{g}) = \sum_{i=1}^n \{Y_i - \hat{g}(X_i)\}^2 + \alpha \int_a^b \{g^{(2)}(x)\}^2 dx$$

Sehingga estimator *Penalized Least Square* di atas dapat dituliskan dengan

$$\mathbf{S}(\mathbf{g}) = (\mathbf{Y} - \mathbf{B}(X)\boldsymbol{\beta})^T (\mathbf{Y} - \mathbf{B}(X)\boldsymbol{\beta}) + \boldsymbol{\alpha} \boldsymbol{\beta}^T \mathbf{D} \boldsymbol{\beta}, \quad (\text{Pers 4})$$

dengan elemen ke- ij dari $\mathbf{D}$ adalah

$$\mathbf{D} = \int_a^b \{B_i''(x) B_j''(x)\} dx.$$

Sehingga penduga fungsi g adalah

$$\hat{g}(X) = \sum_{j=1}^p \mathbf{B}_{j,m}(X) \beta_j, \quad p=m+k \quad (\text{Pers 5})$$

dengan

$$\boldsymbol{\beta} = \left((\mathbf{B}(X))^T \mathbf{B}(X) + \boldsymbol{\alpha} \mathbf{D} \right)^{-1} (\mathbf{B}(X))^T \mathbf{Y}.$$

Metode ICM untuk Model ME

Dalam mengestimasi fungsi g dalam model kesalahan pengukuran dengan Metode ICM diperlukan informasi prior sebelum menentukan distribusi posterior yang selanjutnya nanti digunakan untuk iterasi dalam Metode ICM.

1. Pemilihan Prior

Misalkan data observasi diasumsikan berdistribusi normal,

$$Y \sim N(g(X_i), \sigma_Y^2)$$

$$W \sim N(X_i, \sigma_U^2)$$

Distribusi Prior untuk $\mathbf{X}$ dan $\mathbf{g}$ yang digunakan untuk model kesalahan pengukuran (1) adalah sebagai berikut

$$X \sim N(\mu_X, \sigma_X^2) \quad (\text{Pers 6})$$

dimana μ_X dan σ_X^2 adalah suatu konstanta yang ditentukan.

Sedangkan untuk distribusi *prior* g dilakukan pendekatan “*partially improper*” (Green and Silverman, 1995) yaitu:

$$p(\mathbf{g}) \propto \exp - \frac{\alpha}{2} (\mathbf{g}^T \mathbf{K} \mathbf{g}),$$

dimana untuk dua matrik Q dan R maka matrik K didefinisikan sebagai $K = QR^{-1}Q^T$

Langkah-langkah menentukan matrik K:

- Definisikan: $h_i = t_{i+1} - t_i$, $i=1, \dots, n-1$
- Tentukan matrik Q adalah matrik yang berukuran $n \times (n-2)$ dengan elemen q_{ij} adalah sebagai berikut

$$q_{j-1,j} = h_{j-1}^{-1}, q_{j,j} = -h_{j-1}^{-1}, q_{j+1,j} = h_j^{-1},$$

untuk $i=1, \dots, n$, dan $j=2, \dots, n-1$, dan $q_{ij} = 0$ untuk $|i - j| \geq 2$.

Prior untuk varians-variannya dalam Metode ICM merupakan hasil dari estimasinya, yaitu sebagai berikut:

$$\hat{\sigma}_U^2 = \frac{\sum_{i=1}^n (m_i - 1) s_i^2}{\sum_{i=1}^n (m_i - 1)}, \quad (\text{Pers 7})$$

dan menurut Green and Silverman (1995), estimasi dari varians *error* adalah

$$\sigma_\varepsilon^2 = \frac{\sum_{i=1}^n \{Y_i - g(X)\}^2}{\text{tr}\{I - A(\hat{\alpha})\}}, \quad (\text{Pers 8})$$

dengan $\hat{\alpha}$ diestimasi dengan metode GCV.

2. Pembentukan Densitas Posterior dengan metode ICM

Berdasarkan pada prior pada sub bab 4.2.2.2, selanjutnya ditentukan distribusi posterior dari parameter-parameter dalam regresi nonparametrik. Parameter dalam model (1) yaitu $\mathbf{X}, \mathbf{g}, \sigma_X^2, \sigma_\varepsilon^2$, dan σ_u^2 . Misalkan, $\boldsymbol{\theta} = (\mathbf{X}, \mathbf{g}, \sigma_\varepsilon^2, \sigma_X^2, \sigma_u^2)$ maka berdasarkan persamaan (6) maka densitas posterior dari $\boldsymbol{\theta}$ adalah

$$\begin{aligned} p(\boldsymbol{\theta} | \mathbf{Y}, \mathbf{W}) &\propto p(\boldsymbol{\theta}) p(\mathbf{Y} | \boldsymbol{\theta}) p(\mathbf{W} | \boldsymbol{\theta}) \\ &\propto p(\mathbf{X}, \mathbf{g}, \sigma_X^2, \sigma_\varepsilon^2, \sigma_u^2) p(\mathbf{Y} | \mathbf{g}, \mathbf{X}, \sigma_\varepsilon^2) p(\mathbf{W} | \mathbf{X}, \sigma_u^2) \\ &\propto p(\mathbf{Y} | \mathbf{g}, \mathbf{X}, \sigma_\varepsilon^2) p(\mathbf{W} | \mathbf{X}, \sigma_u^2) p(\mathbf{X}) p(\mathbf{g}) p(\sigma_X^2) p(\sigma_u^2) p(\sigma_\varepsilon^2) p(\alpha) \end{aligned}$$

Pendekatan metode ICM digunakan untuk mengestimasi parameter $\boldsymbol{\theta}$. Estimasi parameter $\boldsymbol{\theta}$ dengan metode ICM yaitu menentukan distribusi posterior persamaan di atas, yang hampir sama dengan menentukan posterior dengan metode Bayesian di mana semua parameternya mempunyai distribusi prior. Namun dalam metode ICM, yang

disebut juga sebagai Bayesian sebagian, tidak semua parameternya mempunyai distribusi, dalam hal ini σ_X^2 , σ_ε^2 , dan σ_u^2 diestimasi dari data.

Sehingga bersyarat pada $\hat{\sigma}_X^2$, $\hat{\sigma}_\varepsilon^2$, dan $\hat{\sigma}_u^2$, distribusi posterior dengan metode ICM adalah

$$p(\theta|\mathbf{Y}, \mathbf{W}) \propto p(\mathbf{Y}|\mathbf{g}, \mathbf{X}, \sigma_\varepsilon^2) p(\mathbf{W}|\mathbf{X}, \sigma_u^2) p(\mathbf{X}) p(\mathbf{g}) p(\sigma_X^2) p(\sigma_U^2) p(\sigma_\varepsilon^2) p(\alpha)$$

$$\propto \exp \left[-\frac{1}{2\sigma_\varepsilon^2} (\mathbf{Y}_i - \mathbf{g}(\mathbf{X}_i))^2 - \frac{1}{2\sigma_W^2} (\mathbf{W}_i - \mathbf{X}_i)^2 - \frac{1}{2\sigma_X^2} (\mathbf{X}_i - \mu_X)^2 + \right.$$

$$\left. - \frac{\hat{\alpha}}{2\sigma_\varepsilon^2} \mathbf{g}^T \mathbf{K} \mathbf{g} \right]$$

(Pers 9)

Pendekatan ICM seperti yang dijelaskan oleh Besag (1986), yaitu dengan menemukan mode posterior dari (9). Dalam metode ICM parameter diperbarui (*update*) dalam setiap iterasinya. Berbeda dengan metode fully Bayesian, dalam metode Bayesian posterior dari masing-masing parameternya diestimasi dari kondisi penuh.

Algoritma ICM

1. *Initialize* $\hat{\sigma}_X^2$, $\hat{\sigma}_\varepsilon^2$, dan $\hat{\sigma}_u^2$, dan $\mathbf{g}^{(0)}$
 $\mathbf{g}^{(0)}$ merupakan hasil estimasi dengan menggunakan Naïve Methode, yaitu meregresikan antara nilai Y dengan rata-rata dari W. Nilai awal dari $\hat{\sigma}_X^2$, $\hat{\sigma}_\varepsilon^2$, dan $\hat{\sigma}_u^2$ diturunkan dari persamaan (6), (7), (8).
2. Berdasarkan pada kondisi 1, menentukan data $\mathbf{X}^{(i)}$ bersyarat $\mathbf{g}^{(i-1)}$ yang memaksimalkan (9)
 Nilai $\mathbf{X}^{(i)}$ yang memberikan nilai yang maksimum pada (4.17) tidak dapat diselesaikan secara analitik. Untuk itu, dalam hal ini, digunakan metode grid uniform untuk memaksimumkan (9).
3. Menentukan vektor $\mathbf{g}^{(i)}$ bersyarat pada $\mathbf{X}^{(i)}$, dengan menggunakan Spline.
 Berdasarkan pada persamaan (4.17), dengan bersyarat pada $\mathbf{X}^{(i)}$, memaksimumkan densitas posterior (9) sama artinya dengan meminimumkan

$$\left[\frac{1}{2\hat{\sigma}_\varepsilon^2} \sum_{i=1}^n (\mathbf{Y}_i - \mathbf{g}(\mathbf{X}_i))^2 + \frac{\hat{\alpha}}{2\hat{\sigma}_\varepsilon^2} \mathbf{g}^T \mathbf{K} \mathbf{g} \right].$$

Bentuk (8) merupakan persamaan *smoothing Spline*. Untuk itu, mengestimasi $\mathbf{g}$ dapat digunakan B-Spline, dengan bersyarat pada $\mathbf{X}^{(i)}$.

4. Iterasi $i=i+1$ diulang sampai i tertentu.

Perhitungan Dengan Program S-Plus

Perhitungan model regresi B-Spline dan ICM dengan menggunakan software SPSS untuk membuat model regresi dari data dengan kesalahan pengukuran adalah sbb:

Program ICM

```
icm _ function(w,y,regress=F,ak=T,iter=10,spar=0,grid=500,sigw=0){
  n _ length(y)
  k _ dim(w)[2]
  wbar _ apply(w,1,mean)

  # initial values of x:
  x _ wbar

  # initial estimate of sigw:
  if (sigw ==0){
    ssw _ apply(w,1,var)
    sigw _ sqrt(mean(ssw))
  }

  # xp is the grid values over x...keep track of the spline values
  xp _ seq(min(wbar),max(wbar),length=grid)

  # Initial smoothing
  fit _ smooth.spline(x,y,all.knots=ak,spar=spar,cv=T)
  spar _ fit$spar
  fit2 _ predict(fit,xp)
  res <- (fit$yin - fit$y)/(1-fit$lev)

  # initial estimate of sigy:
  sigy _ sqrt(var(res))

  for (i in 1:iter){

    # find the condition post. value for each x_i
    for (ii in 1:n){
      val _ (-1/(2*sigy^2))*((y[ii] - fit2$y)^2)
      val _ val - (k/(2*sigw^2))*((mean(w[ii,]) - xp)^2)
      if (regress){
        val _ val - (1/(2*var(x)^2))*((mean(xp) - xp)^2)
      }

      # Set x_i to its max
      x[ii] _ xp[val==max(val)]
    }

    fit _ smooth.spline(x,y,spar=spar,all.knots=ak)
    # this saves the predicted values for each grid point (the height of y)
    fit2 _ predict(fit,xp)

    fit21 _ predict(fit,xp)
```

```

res1 <- (fit$yin - fit$y)/(1-fit$lev)

# estimate of sigy icm:
sigy1 _ sqrt(var(res1))
out _ list(x=xp,y=fit2$y,s,xfit=x,sigy=sigy,sigy1=sigy,sigw=sigw,spar=spar)
out
}

```

Simulasi Data Untuk :

Kasus 1: fungsi eksponensial

```

#####here is an example simulation: (uncomment and run)
ffl _ function(x)
{
  d <- Exp(x)
  d
}

##### generate the x-y-w
n _ 100
x _ rnorm(n,0.5,0.25)
y _ ffl(x) + rnorm(n,0,.015)
repeats _ 2
w _ matrix(0,nrow=n,ncol=repeats)
for (j in 1:repeats){
  w[,j] _ x + rnorm(n,0,(sqrt(3/7)*0.25) )}
wbar _ apply(w,1,mean)
xp _ seq(min(wbar),max(wbar),length=500)
yp _ ffl(xp)

##### neat pictures (prints (x,y) points, (wbar,y) , true curve,
##### naive, and icm spline
plot(x,y,pch=5,xlim=c(min(wbar),max(wbar)),ylim=c(min(y),max(y)),
  xlab="",ylab="")
points(wbar,y,pch=18,col=2)
lines(xp,yp,lty=1,lwd=2)
ddd _ smooth.spline(wbar,y)
lines(ddd$x,ddd$y,lty=1,col=2,lwd=2)
im5 _ icm(w,y,iter=5)
lines(im5$x,im5$y,lty=1,lwd=2,col=3)

```

kasus 2 : fungsi truncated

```

ffl <- function(x)
{
  d <- trunc(x) + trunc(1 - x)
  d
}
> for(p in 1:10) {
  sim <- function(w, y)
  {
    ##### generate the x-y-w
    n <- 250
    x <- rnorm(n, 0, 1)
    y <- ffl(x) + rnorm(n, 0, 0.01)

```

```

repeats <- 2
w <- matrix(0, nrow = n, ncol =
  repeats)
for(j in 1:repeats) {
  w[, j] <- x + rnorm(n, 0, 0.5)
}
#print(w[,1])
#print(w[,2])
wbar <- apply(w, 1, mean)
xp <- seq(min(wbar), max(wbar), length
  = 500)
yp <- ffl(xp)
##### neat pictures (prints (x,y) points, (wbar,y) , true curve,
##### naive, and icm spline
plot(x, y, pch = 5, xlim = c(min(wbar),
  max(wbar)), ylim = c(min(y),
  max(y)), xlab = "", ylab = "")
points(wbar, y, pch = 18, col = 2)
lines(xp, yp, lty = 1, lwd = 2)
dw1 <- smooth.spline(w[, 1], y)
lines(dw1$x, dw1$y, lty = 1, col = 5,
  lwd = 2)
dw2 <- smooth.spline(w[, 2], y)
lines(dw2$x, dw2$y, lty = 1, col = 8,
  lwd = 2)
ddd <- smooth.spline(wbar, y)
lines(ddd$x, ddd$y, lty = 1, col = 4,
  lwd = 2)
im5 <- icm(w, y, iter = 5)
lines(im5$x, im5$y, lty = 1, lwd = 2,
  col = 6)
out <- list(im5)
out
}
h[p] <- sim(w, y)
s <- h[p]
}

```

kasus 3:fungsi trigonometri

```

#####here is an example simulation: (uncomment and run)
ffl _ function(x)
{
  d <- sin(0.5 * pi * x)/( 1 + (sign(x) + 1)*(2*x^2))
  d
}
##### generate the x-y-w
n _ 100
x _ rnorm(n,0.5,0.25)
y _ ffl(x) + rnorm(n,0,.015)
repeats _ 2
w _ matrix(0,nrow=n,ncol=repeats)
for (j in 1:repeats){

```

```

w[,j] _ x + rnorm(n,0,(sqrt(3/7)*0.25) )}
wbar _ apply(w,1,mean)
xp _ seq(min(wbar),max(wbar),length=500)
yp _ ffl(xp)

##### neat pictures (prints (x,y) points, (wbar,y) , true curve,
##### naive, and icm spline
plot(x,y,pch=5,xlim=c(min(wbar),max(wbar)),ylim=c(min(y),max(y)),
     xlab="",ylab="")
points(wbar,y,pch=18,col=2)
lines(xp,yp,lty=1,lwd=2)
ddd _ smooth.spline(wbar,y)
lines(ddd$x,ddd$y,lty=1,col=2,lwd=2)
im5 _ icm(w,y,iter=5)
lines(im5$x,im5$y,lty=1,lwd=2,col=3)

```

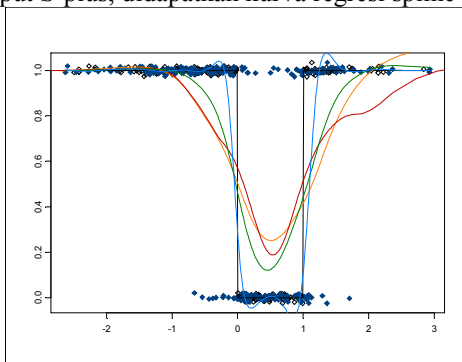
SIMULASI

Simulasi dilakukan untuk 3 model data yang mewakili proses dalam model data statistic di bidang bisnis, yaitu eksponen, truncated dan trigonometri. Selanjutnya dilakukan simulasi untuk masing-masing model 1, model 2, dengan variasi nilai $n = 25, 50, 100, 250$, $\sigma_{\epsilon}^2 = 0.01^2, 0.1^2, 1^2$, $\sigma_U^2 = 0.01^2, 0.1^2, 0.5^2$ dan 1^2 , $\mu_x = 0$, dan $\sigma_x^2 = 1$.

Model 1: Fungsi *Truncated*: $g(x) = x_+ + (1 - x)_+$.

dengan $n=250$, $\sigma_{\epsilon}^2 = 0.01^2$, $\sigma_U^2 = 0.5^2$, $\mu_x = 0$, dan $\sigma_x^2 = 1$.

Dan dari hasil output S-plus, didapatkan kurva regresi spline dan nilai MSE sbb:



Gambar 1 Estimasi Kurva untuk Fungsi *Truncated*.

Tabel 2 Nilai MSE Fungsi *Truncated*.

Metode	Nilai MSE
Non ME[1]	0.157338
Non ME[2]	0.139422
Naïve	0.102883

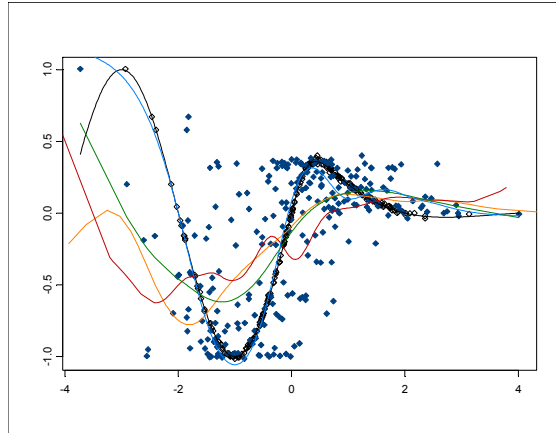
ICM	0.00143
-----	---------

Model 2 : Fungsi Trigonometri

$$g(x) = \frac{\sin(\pi x / 2)}{1 + 2x^2(\sin g(x) + 1)},$$

$$n=250, \sigma_{\varepsilon}^2 = 0.01^2, \sigma_u^2 = 0.1^2, \mu_x = 0, \text{ dan } \sigma_x^2 = 1.$$

Dan dari hasil output S-plus, didapatkan kurva regresi spline dari simulasi di atas adalah sebagai berikut:



Gambar 2 Estimasi Kurva untuk Fungsi Trigonometri

Tabel 3 Nilai MSE untuk Fungsi Trigonometri

Metode	Nilai MSE
Non ME[1]	0.158436
Non ME[2]	0.195413
Naïve	0.147683
ICM	0.030937

Analisa Perbandingan Model Me Dan Non Me

Berdasarkan hasil simulasi dengan SPLUS untuk ketiga model di atas, Gambar 1 dan 2 menunjukkan kurva regresi dengan metode ICM.

Dilihat dari Gambar 1 dan 2, pengabaian kesalahan pengukuran (model non-ME) berpengaruh terhadap kurva regresi. Terlihat bahwa kurva regresi bila kesalahan pengukuran diabaikan jauh dari kurva sebenarnya. Nampak juga dari Gambar 1 dan 2, kurva regresi dengan ICM yang paling mendekati dengan kurva sebenarnya, $g(x)$. Dilihat dari hasil MSE untuk ketiga model simulasi, pengabaian kesalahan pengukuran (model non-ME) berpengaruh terhadap hasil MSE. Terlihat bahwa estimasi kurva regresi bila kesalahan pengukuran diabaikan jauh lebih besar nilai MSE nya dibandingkan dengan model ME dengan metode ICM. Nampak juga dari Tabel 1 dan 2,

estimasi kurva regresi untuk data simulasi dengan metode ICM memberikan nilai MSE paling kecil dibandingkan metode Naive maupun dengan model non-ME.

Hal ini membuktikan bahwa adanya kesalahan pengukuran dalam variabel prediktor, X , berpengaruh terhadap estimasi kurva regresi. Dan pengabaian kesalahan pengukuran ini bisa mengakibatkan ketidakakuratan prediksi model regresi, sehingga berpengaruh terhadap hasil peramalan.

Simpulan

Berdasarkan pembahasan di atas, maka dapat disimpulkan bahwa: Aplikasi Software S-Plus telah dikembangkan guna perhitungan dan estimasi model regresi baik dengan kesalahan pengukuran maupun apabila kesalahan pengukuran diabaikan, dan hal ini sangat membantu guna menunjang pengambilan keputusan di berbagai bidang di antaranya bidang bisnis.

Daftar Pustaka

- Amemiya, Y and Fuller, W. A .1984. *Estimation for Multivariate Errors in Variable Model with Estimated Error Covariance Matrix*. Annals of Statistics. Annals of Statistics, 12, 497-509.
- Berry, S. A., Carroll, R. J. and Ruppert, D.2002. *Bayesian Smoothing And Regression Splines For Measurement Error Problems*. Journal of the American Statistical Association, 9, 160–169.
- Besag, J. 1986. *On the Statistical Analysis of Dirty Pictures” (with Discussion)*. Journal of the Royal Statistical Society, Series B, 48, 259-279.
- Box, G. E. P. and Tiao, G. .1973. *Bayesian Inference in Statistical Analysis*. Addison–Wesley, London.
- Carroll, R. J., Maca, J. D. and Ruppert, D. 1999. *Nonparametric Regression With Errors In Covariates*. Biometrika, 86, 541–554.
- Carroll, R. J., Ruppert, D., and Stefanski, L. A. 1995. *Measurement Error in Nonlinear Models*, Chapman and Hall, New York.
- Cook, J. R. and Stefanski, L. A. 1994. *Simulation–Extrapolation Estimation In Parametric Measurement Error Models*. Journal of the American Statistical Association, **89**, 1314–1328.
- Fan, J. and Truong, Y. K. 1993. *Nonparametric regression with errors in variables*. Annals of Statistics, 21, 1900–25.
- Fuller, W. A and Hidiroglou, M. A .1976. *Regression Estimation After Correcting for Attenuation*. Journal of the American Statistical Association, **73**, 99-169.
- Green, P. J. and Silverman, B. W. 1994. *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*, Chapman and Hall, London.
- Hardle, W. 1990. *Applied Nonparametrics Regression*, Australia, : Cambridge University Press.
- Hartatik.2012. *Aplikasi S-Plus Dalam Analisa Regresi Guna Estimasi Model Regresi Untuk Data Dengan Kesalahan Pengukuran Sebagai Dasar Dalam Pengambilan Keputusan*. Proceeding SNATI 2011, Yogyakarta.