ISSN 1978-3787 (Cetak)
ISSN 2615-3505 (Online)

ESTIMASI MODEL LINEAR PARSIAL UNTUK DATA RESPON HILANG MENGGUNAKAN PENDEKATAN NORMAL

Oleh
Nur Salam[1)], Fuad Muhajirin Farid[2)], Maisarah[3)] & Nur Haviva[4)]
[1,2,3,4]Program Studi Statistika, FMIPA Universitas Lambung Mangkurat.
Email : [1]nursalam2011@gmail.com, [2]fuad.farid@ulm.ac.id, [3]maisarah@ulm.ac.id & [4]nurhafifah471@gmail.com

Abstrak
Paper ini akan membahas estimasi model linear parsial (semiparametrik) dengan respon hilang menggunakan pendekatan normal. Suatu kelas estimator didefinisikan yang memuat kasus-kasus khusus yaitu estimator imputasi regresi semiparametrik, estimator rata-rata marginal dan estimator berbobot skor kecenderungan. Kelas estimator tersebut adalah normal secara asimtotik. Tiga estimator khusus tersebut mempunyai variansi asimtotik yang sama. Dimana estimator ini mencapai batas efisiensi dalam kasus normal homoskedastik. Diperlihatkan bahwa metode *jackknife* dapat digunakan untuk mengestimasi variansi asimtotik secara konsisten. Berdasarkan keadaan-keadaan di atas akan diestimasi mean Y, sebut θ. Ketiga estimator khusus di atas akan digunakan untuk mengestimasi mean Y yaitu θ berupa estimasi titik dan interval kepercayaan dengan beberapa respon hilang menggunakan metode pendekatan normal. Suatu studi simulasi dilakukan untuk membandingkan diantara 3 estimator khusus dengan metode berdasarkan pendekatan normal dalam hal nilai rata-rata Y dan lebar dari interval kepercayaan.
Kata Kunci : Estimasi & Model Linear Parsial

PENDAHULUAN

Model linear parsial (semiparametrik) merupakan model pendekatan baru dalam regresi diantara dua model regresi sudah populer yaitu regresi parametrik dan regresi nonparametrik. Model linear parsial merupakan model gabungan yang memuat keduanya yaitu komponen parametrik dan komponen nonparametrik. Pendekatan model linear parsial ini menarik karena hanya ada sebagian informasi tentang hubungan antara variabel bebas (X) dengan variabel respon (Y) yang diketahui, sehingga penggunaan model regresi nonparametrik lengkap tidak lagi efisien dan penggunaan model regresi parametrik lengkap juga mungkin keliru. Setidaknya, ini merupakan motivasi penulis untuk membahas model regresi semiparametrik. Adapun model linear parsial adalah :

$$Y_i = X_i^T \beta + g(T_i) + \varepsilon_i, \quad (1)$$

dengan Y_i adalah variabel-variabel respon skalar yang *independent and identically distributed* (i.i.d), X_i adalah vektor-vektor kovariat random d-variabel i.i.d, T_i adalah vektor-vektor kovariat random d*-variabel i.i.d, β adalah vektor dari parameter yang tidak diketahui, fungsi g(.) tidak diketahui, dan *error-error* model ε_i adalah independen dengan mean 0 dan variansi tetap σ^2.(Wang et al. 2004)

Dalam praktek keseharian, sering kita berhadapan dengan tidak semuanya variabel respon tersedia karena berbagai alasan seperti kehilangan informasi yang disebabkan oleh faktor-faktor yang luar kontrol, kegagalan pada pihak investigator untuk menghimpun informasi yang benar, ketidakinginan beberapa unit yang disampel untuk menyediakan informasi yang diinginkan dan seterusnya. Dalam kasus ini, prosedur-prosedur inferensi dalam hal ini estimasi dan uji hipotesis tidak dapat diterapkan secara langsung. Suatu metode umum untuk menangani data yang hilang dalam suatu data set besar adalah mengimputasi (memasukkan) suatu nilai layak untuk setiap data yang hilang dan

ISSN No. 1978-3787 (Cetak)
ISSN 2615-3505 (Online)

kemudian menganalisis hasil itu seolah-olah mereka lengkap. Jika diandaikan X adalah suatu vektor dimensi-d dari faktor-faktor dan Y adalah suatu variabel respons yang dipengaruhi oleh X. Diperoleh suatu sampel random data yang tidak lengkap :

$$(X_i, Y_i, \delta_i), \quad i = 1, 2, ... n,$$

dengan X_i adalah semua data yang terobservasi dan $\delta_i = 0$ jika Y_i hilang dan kalau tidak $\delta_i = 1$. Secara spesifik, pada penelitian ini akan dibahas suatu kasus dimana beberapa nilai Y dalam suatu sampel berukuran n mungkin hilang, namun X dan T terobservasi secara lengkap. Yakni, diperoleh observasi tidak lengkap berikut :

$$(Y_i, \delta_i, X_i, T_i), \quad i = 1, 2, ..., n,$$

dari model (1), dimana semua X_i dan T_i terobservasi secara lengkap dan $\delta_i = 0$ jika Y_i hilang dan kalau tidak $\delta_i = 1$. Dalam paper ini kita tertarik dalam inferensi terhadap mean Y, katakanlah θ, bila ada respon yang hilang dalam model linear parsial (1).

Oleh karena itu suatu hal yang cukup menarik adalah bagaimana mengestimasi model linear parsial yakni mengestimasi parameter dan fungsinya menggunakan metode *least square* jika beberapa data hilang pada variabel yang diteliti, dalam hal ini pada variabel respon Y ada beberapa yang hilang. Adapun judul penelitian yang akan kami angkat pada kesempatan ini adalah estimasi model linear parsial untuk data hilang menggunakan pendekatan normal dan simulasinya dengan S-Plus.

LANDASAN TEORI

Sebelum membahas konsep estimasi dan normal asimtotik terlebih dahulu dibicarakan beberapa pengertian dasar yang merupakan konsep awal yang harus dipahami agar mudah mengikuti pembahasan yang dibicarakan.

Definisi 1.2.1 Data Hilang (Rubin,1987)

Data hilang (*missing data*) merupakan sebuah nilai yang mengindikasikan bahwa tidak ada data apapun yang tersimpan pada variabel pengamatan saat ini. Kasus Data hilang ini akan mengakibatkan ketidaklengkapan data (*incomplete data*) dalam suatu model, sehingga akan menghambat analisis statistik yang akan dilakukan.

Definisi 1.2.2. Estimasi (Bain, J.L & Engeilhardt, M, 1992)

Suatu statistik, $T = l(X_1, X_2,\ldots,X_n)$ yang digunakan untuk mengestimasi nilai $\tau(\theta)$ disebut estimator dari $\tau(\theta)$ dan suatu nilai observasi dari suatu statistik, $l(X_1, X_2,\ldots,X_n)$ disebut hasil estimasi.

Definisi 1.2.3. Konvergen dalam probabilitas (Casela & Berger, 1990)

Barisan variabel random $X_1, X_2, X_3, \ldots$ konvergen dalam probabilitas ke suatu variabel random X jika untuk setiap $\varepsilon > 0$, $\lim_{n\to\infty} P(|X_n - X| \geq \varepsilon) = 0$ atau $\lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$ atau bisa juga ditulis $X_n \xrightarrow{P} X$.

Definisi 1.2.4. Konvergen dalam distribusi (Casella & Berger, 1990)

Barisan variabel random $X_1, X_2, X_3, \ldots$ konvergen dalam distribusi ke suatu variabel random X, jika $\lim_{n\to\infty} F_{xn}(x) = F_x(x)$ pada setiap titik X, dimana $F_x(x)$ kontinu atau bisa ditulis $X_n \xrightarrow{d} X$.

METODE PENELITIAN

Adapun prosedur-prosedur yang dilakukan dalam penelitian ini adalah sebagai berikut :

1. Mengumpulkan bahan-bahan penelitian yang berhubungan dengan model linear parsial (regresi semiparametrik) dan cara dan proses estiamsinya serta metode *least square.*
2. Mempelajari bahan-bahan yang telah dikumpulkan pada point (1) di atas.
3. Menjelaskan model regresi parametrik dan bentuk modelnya dalam bentuk matriks.
4. Menjelaskan model regresi nonparametrik dan bentuk modelnya dalam bentuk matriks.
5. Mengkonstruksi model linear parsial untuk data hilang dan bentuk modelnya dalam bentuk matriks.

6. Menentukan metode estimasi model linear parsial untuk data hilang yang tepat.
7. Mengestimasi model linear parsial untuk data hilang dengan cara mengestimasi bagian parametrik dan bagian bagian nonparametriknya menggunakan metode *least square* dengan pendekatan normal.
8. Membuat simulasi model regresi linear parsial untuk data hilang dengan cara mengestimasi bagian parametrik dan bagian bagian nonparametriknya dengan metode *least square* dengan pendekatan normal.
9. Menarik kesimpulan dari hasil pembahasan.

HASIL DAN PEMBAHASAN

Dalam bab ini didefinisikan estimator-estimator dan sifat-sifat asimtotiknya yang akan dibahas dalam paper ini.

Estimasi Model Linear Parsial

Terlebih dahulu dideskripsikan bagaimana mengestimasi fungsi regresi. Melalui pramultiplikasi (1) dengan indikator observasi diperoleh :

$$\delta_i Y_i = \delta_i X_i^T\beta + \delta_i g(T_i) + \delta_i\varepsilon_i,$$

dan dengan mengambil ekspektasi bersyarat yang diberikan T, didapat :

$$E[\delta_i Y_i | T_i = t] = E[\delta_i X_i^T | T_i = t]\beta + E[\delta_i | T_i = t]g(t)$$

dari hal di atas diperoleh :

$$g(t) = g_2(t) - g_1(t)^T\beta, \quad (2)$$

dengan :

$$g_1(t) = \frac{E[\delta X | T = t]}{E[\delta | T = t]} \quad \text{dan} \quad g_2(t) = \frac{E[\delta Y | T = t]}{E[\delta | T = t]}.$$

Sehingga dihasilkan:

$$\delta_i[Y_i - g_2(T_i)] = \delta_i[X_i - g_1(T_i)]^T\beta + \delta_i\varepsilon_i, \quad (3)$$

yang mengisyaratkan bahwa suatu estimator β dapat didasarkan pada suatu regresi kuadrat terkecil dengan menggunakan observasi $\delta_i = 1$ dan estimasi $g_j(.)$, $j = 1,2$.

Andaikan K(.) adalah suatu fungsi kernel dan andaikan h_n adalah suatu sekuens *bandwidth* yang cenderung ke 0 bila $n \to \infty$, dan didefinisikan bobot-bobot :

$$W_{nj}(t) = \frac{K\left((t-T_j)/h_n\right)}{\sum_{j=1}^{n}\delta_j K\left((t-T_j)/h_n\right)}$$

kemudian $\tilde{g}_{1n}(t) = \sum_{j=1}^{n}\delta_j W_{nj}(t) X_j$ dan $\tilde{g}_{2n}(t) = \sum_{j=1}^{n}\delta_j W_{nj}(t) Y_j$ adalah estimator-estimator konsisten dari $g_1(t)$ dan $g_2(t)$, secara berturut-turut. Dari (3), estimator β kemudian didefinisikan sebagai estimator yang memenuhi :

$$\min_{\beta} \sum_{i=1}^{n} \delta_i \left\{\left(Y_i - \tilde{g}_{2n}(T_i)\right) - \left(X_i - \tilde{g}_{1n}(T_i)\right)\beta\right\}^2 \quad (4)$$

Dari (4), dapat diperoleh bahwa estimator β diberikan oleh :

$$\hat{\beta}_n = \left[\sum_{i=1}^{n}\delta_i \left\{\left(X_i - \tilde{g}_{1n}(T_i)\right)\left(X_i - \tilde{g}_{1n}(T_i)\right)^T\right\}\right]^{-1} \times \sum_{i=1}^{n}\delta_i \left\{\left(X_i - \tilde{g}_{1n}(T_i)\right)\left(Y_i - \tilde{g}_{2n}(T_i)\right)\right\}$$

berdasarkan pada tripel yang diobservasi (X_i, T_i, Y_i) untuk itu $i \in \{i : \delta_i = 1\}$. Persamaan (2) mengisyaratkan bahwa suatu estimator g(t) dapat didefinisikan sebagai :

$$\hat{g}_n(t) = \tilde{g}_{2n}(t) - \tilde{g}_{1n}^T(t)\hat{\beta}_n$$

dengan mengganti β, $g_1(t)$ dan $g_2(t)$ dalam (2) dengan $\hat{\beta}_n$, $\tilde{g}_{1n}(t)$ dan $\tilde{g}_{2n}(t)$.

Di dalam membahas mengenai estimasi θ, ditentukan kelas umum estimator-estimator θ yaitu :

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i Y_i}{P_n^*(X_i, T_i)} + \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\delta_i}{P_n^*(X_i, T_i)}\right)\left\{X_i^T\hat{\beta}_n + \hat{g}_n(T_i)\right\}$$

dengan $P_n^*(x,t)$ adalah suatu sekuens kuantitas dengan limit-limit probabilitas $P^*(x,t)$.

Kita khususnya berkepentingan dalam beberapa hal khusus. Pertama, bila $P_n^*(x,t) = 1$, diperoleh estimator imputasi regresi θ :

$$\hat{\theta}_I = \frac{1}{n}\sum_{i=1}^{n}\left\{\delta_i Y_i + (1-\delta_i)(X_i^T\hat{\beta}_n + \hat{g}_n(T_i))\right\}$$

Bila $P_n^*(x,t) = \infty$, diperoleh estimator rata-rata marginal :

$$\hat{\theta}_{MA} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i^T \hat{\beta}_n + \hat{g}_n(T_i)\right)$$

yang hanya rata-rata atas fungsi regresi yang diestimasi. Didefinisikan skore kecenderungan marginal $P_1(t) = P(\delta=1 \mid T=t)$. Bila : $P_n*(x,t) =$

$$\hat{P}_1(t) = \sum_{j=1}^{n}\delta_j K\left(\frac{t - T_j}{h_n}\right) \Big/ \sum_{j=1}^{n} K\left(\frac{t - T_j}{h_n}\right)$$

Diperoleh estimator berbobot skor kecenderungan (marginal) :

$$\hat{\theta}_{p1} = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{\delta_i Y_i}{\hat{P}_1(T_i)} + \left(1 - \frac{\delta_i}{\hat{P}_1(T_i)}\right)\left(X_i \hat{T} \hat{\beta}_n + \hat{g}_n(T_i)\right)\right].$$

Estimator $\hat{\theta}_{P1}$ berbeda dari metode pemberi bobot skor kecenderungan biasa yang menggunakan suatu estimator skore kecenderungan penuh. Andaikan $\hat{\theta}*$ menunjukkan $\hat{\theta}_I, \hat{\theta}_{MA}$ atau $\hat{\theta}_{P1}$. Estimator-estimator ini hanya berlandaskan pada operasi-operasi *smoothing* satu dimensi dan didefinisikan secara explisit. Dua sifat ini penting dari suatu sudut pandang komputasional dan statistika.

Normal Asimtotik

Selanjutnya diberikan beberapa sifat estimator $\hat{\theta}$ dan estimator-estimator variansi konsistennya. Andaikan $P_1(t) = p(\delta = 1 \mid T = t)$, $P(x,t) = p(\delta = 1 \mid X = x, T = t)$, $m(x,t) = x^T\beta + g(t)$ dan $\sigma^2(x,t) = E[(Y - X^T\beta - g(T))^2 \mid X = x, T = t]$. Kemudian didefinisikan $u(x,t) = x - g_1(t)$, $\Sigma = E[P(X,T)\, u(X,T)\, u(X,T)^T]$. $g_{1r}(.)$ menunjukkan komponen ke-r dari $g_1(.)$. Diberikan $\|.\|$ adalah *norm Euclidean.* Asumsi-asumsi berikut diperlukan untuk normal asimtotik dari $\hat{\theta}$, yakni :

1. $\text{Sup}_t\, E\left[\|X\| \mid T = t\right] < \infty$.
2. Fungsi densitas T, katakanlah r(t), ada dan memenuhi $0 < \inf_{t\in[0,1]} r(t) \le \sup_{t\in[0,1]} r(t) < \infty$.
3. $\text{Sup}_{x,t}\, E[Y^2 \mid X = x, T = t] < \infty$.
4. $g(.)$, $g_{1r}(.)$ dan $g_2(.)$ memenuhi syarat Lipschitz order 1.
5. (a) $P_1(t)$ memiliki derivatif-derivatif parsial terbatas hingga order 2 hampir pasti (*almost surely*).
   (b) inf x,t $P(x,t) > 0$.
6. $\Sigma = E[P(X,T)\ u(X,T)\ u(X,T)^T]$ adalah suatu matrik definit positif.
7. (a) Ada konstanta $M_1 > 0$, $M_2 > 0$ dan $P > 0$ sedemikian rupa sehingga :

   $$M_1 I[|u| \le \rho] \le K(u) \le M_2 I[|u| \le \rho].$$

   (b) K(.) adalah suatu fungsi kernel order 2.
   (c) K(.) mempunyai derivatif-derivatif parsial terbatas hingga order 2 hampir pasti (*almost surely*).
8. (a) Fungsi kernel W(.) adalah suatu fungsi kernel terbatas dengan dukungan (*support*) terbatas dan variasi terbatas.
   (b) W(.) adalah suatu kernel order k (> d+1).

Teorema 3.1

Berdasarkan semua asumsi tersebut kecuali untuk 7 (c) diperoleh :

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N(0, V)$$

dengan :

$V = E[(\pi_0(X_1T)+\pi_1(X,T))^2 P(X,T)\ \sigma^2(X,T)] + Var[(X,T)]$ dengan $\pi_0(x,t) = 1/P_1(t)$ dan $\pi_1(x,t) = E[u(X,T)^T]\Sigma^{-1} u(x,t)$ bila $P_n*(x,t) \in \{1, \infty, \hat{P}_1(t)\}$ dan $\pi_0(x,t) = 1/P(x,t)$ dan $\pi_1(x,t) = 0$ bila $P_n*(x,t)$ diambil sebagai $\hat{P}(x,t)$.

Untuk mendefinisikan suatu estimator konsisten dari V, mungkin lebih dahulu didefinisikan estimator-estimator dari $P(x,t)$, $P_1(t)$, $\sigma^2(x,t)$ dan $g_1(t)$ oleh metode regresi kernel dan kemudian mendefinisikan suatu estimator konsisten dari V dengan suatu metode *plug in*. Namun demikian, metode ini mungkin tidak mengestimasi V dengan baik bila dimensi x tinggi. Ini bisa dihindari karena baik $P(x,t)$ dan $\sigma^2(x,t)$ hanya masuk dalam pembilang dan bisa diganti dengan residu-residu kuadrat atau fungsi indikator bila tepat.

Suatu alternatif adalah estimator variansi *jackknife*. Andaikan $\hat{\theta}(-)$ menjadi $\hat{\theta}$ didasarkan pada $\{(Y_j,\delta_j,X_j,T_j)\}_{j \neq i}$ untuk i = 1,2,...,n. Andaikan ini adalah nilai-nilai pseudo *jackknife*. Yakni $J_{ni} = n\hat{\theta} - (n-1)\hat{\theta}^{(-i)}$, i = 1,2,…,n.

Maka estimator variansi *jackknife* bisa didefinisikan sebagai :

$$\hat{V}_{nj} = \frac{1}{n}\sum_{i=1}^{n}(J_{ni} - \bar{J}_n)^2$$

dengan $\bar{J}_n = n^{-1}\sum_{i=1}^{n} J_{ni}$.

Teorema 3.2

Berdasarkan asumsi-asumsi dari Teorema 3.1, diperoleh $\hat{V}_{nj} \xrightarrow{p} V$.

Dengan Teorema 3.1 dan 3.2, interval kepercayaan berdasarkan perkiraan normal dengan level kepercayaan 1- α adalah $\hat{\theta} \pm \sqrt{\hat{V}_{nJ}/n}\ x\ u_{1-\alpha/2}$, dengan $u_{1-\alpha/2}$ adalah kuantil 1-α/2 dari distribusi normal standar.

Studi Simulasi

Pada bab ini akan dilakukan simulasi untuk metode pendekatan normal untuk model linear parsial dengan data respon hilang menggunakan paket program S-Plus. Simulasi ini menggunakan model regresi semiparametrik :

$$Y = bX + g(T) + \varepsilon \qquad (7.1)$$

dengan : $X \sim N(0,1)$ $\qquad b = 0,75$

$$\varepsilon \sim N(0, 0.015^2)$$

$$g(T) = T^3(1-T)^3$$ dengan

$$T = \text{sort}(N(0.5, 0.25^2))$$

Berdasarkan model dapat disusun program dan setelah dijalankan dengan ukuran sampel (n) 50, 100 dan 150 diperoeh *output* sebagai berikut :

Tabel 1. Nilai estimasi dan batas interval kepercayaan dari θ dengan beberapa respon hilang menggunakan metode pendekatan normal dan ukuran sampel n dengan level kepercayaan 0.95.

Ukuran Sampel n= 50	Estimasi Titik	Batas Bawah Interval Kepercayaan	Batas Atas Interval Kepercayaan	Lebar Interval Kepercayaan
θ_I	0.4845758	0.3205440	0.6486076	0.328064
θ_{MA}	0.4845971	0.3205752	0.6486190	0.328064
θ_{P1}	0.4845754	0.3205434	0.6486074	0.328064
Ukuran Sampel n= 100	**Estimasi Titik**	**Batas Bawah Interval Kepercayaan**	**Batas Atas Interval Kepercayaan**	**Lebar Interval Kepercayaan**
θ_I	0.5321706	0.3880358	0.6763054	0.2882698
θ_{MA}	0.5321279	0.3879898	0.6762659	0.2882698
θ_{P1}	0.5321724	0.3880375	0.6763073	0.2882698
Ukuran Sampel n= 150	**Estimasi Titik**	**Batas Bawah Interval Kepercayaan**	**Batas Atas Interval Kepercayaan**	**Lebar Interval Kepercayaan**
θ_I	0.5277817	0.4252602	0.6303033	0.2050434
θ_{MA}	0.5277469	0.4252258	0.6302680	0.2050434
θ_{P1}	0.5277876	0.4252659	0.6303093	0.2050434

Dari tabel di atas terlihat bahwa hasil simulasi untuk ke tiga estimator yaitu θ_I, θ_{MA} dan θ_{P1} dapat mengestimasi nilai rata-rata variabel respon yang hampir sama dimana ada beberapa variabel respon yang hilang dan untuk setiap keadaan dimana ukuran sampel random yang berbeda n, yaitu n = 50, n = 100 dan n = 150. Berdasarkan hasil simulasi juga diperoleh lebar interval untuk ketiga estimator yakni θ_I, θ_{MA} dan θ_{P1} adalah semakin kecil untuk n yang semakin besar serta mempunyai lebar interval yang sama untuk setiap keadaan dimana beberapa variabel respon hilang dan untuk setiap keadaan dimana ukuran sampel yang berbeda n, yakni n = 50, n = 100 dan n = 150.

PENUTUP

Kesimpulan

Kesimpulan yang diperoleh berdasarkan hasil pembahasan adalah sebagai berikut:

1. Hasil estimasi titik model linear parsial dengan respon hilang dalam bentuk kelas estimator yaitu :

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n}\frac{\delta_i Y_i}{P_n*(X_i,T_i)} + \frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{\delta_i}{P_n*(X_i,T_i)}\right)\left(X_i^T\hat{\beta}_n + \hat{g}_n(T_i)\right).$$

2. Hasil estimasi interval kepercayaan model linear parsial dengan respon hilang menggunakan metode pendekatan normal dengan level kepercayaan $1-\alpha$ untuk θ adalah :

$$\hat{\theta} - \sqrt{\frac{\hat{V}_{nJ}}{n}} \times u_{1-\alpha/2} < \theta < \hat{\theta} + \sqrt{\frac{\hat{V}_{nJ}}{n}} \times u_{1-\alpha/2}.$$

3. Semakin besar nilai n ($n \to \infty$) maka nilai θ cenderung mempunyai nilai tertentu yang hampir sama, serta lebar interval kepercayaan θ cenderung semakin mengecil.

Ucapan Terima Kasih

Ucapan terima kasih penulis haturkan kepada Universitas Lambung Mangkurat yang telah mendanai penelitian ini melalui dana DIPA Universitas Lambung Mangkurat tahun anggaran 2020.

DAFTAR PUSTAKA

[1] Bain, L.J., & Engelhardt M.(1992). *Introduction to Probability and Mathematical Statistics*, Second Edition, Duxbury Press, California.
[2] Casella, G., & Berger,L.R. (2002). *Statistical Inference*, Duxbury, USA.
[3] Draper, R.N., & Smith, H. (1998) *Applied Regression Analysis*, Johan Wiley & Sons, INC.
[4] Hardle,W., (1994). *Applied Nonparametric Regression*, The Syndicate of the Universitas of Cambridge.
[5] Hardle, W, liang. H & Gao. J., **(**2000). *Partially Linear Model*, Physica- Verlag, Heidelberg.
[6] Owen, A.(1990). Empirical Likelihood Ratio Confidence Regions, *The Annals of Statistics*, 18, 90-120.
[7] Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys.* NY: John Wiley & Sons. New York.
[8] Wang. Q., Linton,O., & Hardle,W. (2004). Semiparametric Regression Analysis With Missing Response at Random, *Journal of the American Statistical Association* ; 99 ; 334-345.
[9] Walpole,E.R (1995). *Pengantar Statistik*, PT, Gramedia Pustaka Utama