

Penerapan Metode Regresi *Kernel Smoothing* untuk Imputasi Data Lama Waktu Terinfeksi Covid-19

Tawangki Sri Fadilah, Abdul Kudus*

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

ARTICLE INFO

Article history :

Received : 11/2/2023

Revised : 20/5/2023

Published : 14/7/2023



Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Volume : 3

No. : 1

Halaman : 51 - 60

Terbitan : **Juli 2023**

ABSTRAK

Penelitian ini bertujuan untuk menjelaskan mengenai kasus lama waktu pasien terinfeksi COVID-19 serta menerapkan metode regresi kernel smoothing untuk imputasi data tersensor. Metode analisis data yang digunakan dalam penelitian ini adalah menggunakan metode kernel smoothing. Metode kernel smoothing merupakan metode pemulusan (smoothing) dengan menggunakan fungsi kernel. Metode ini dapat digunakan sebagai alternatif untuk menyelesaikan permasalahan data yang fluktuatif. Data yang digunakan yaitu lama waktu pasien terinfeksi covid-19 sampai sembuh (Y) dan usia pasien (X) dari hasil penelitian metode kernel smoothing, maka hasil scatterplot imputasi berdasarkan case dan usia menunjukkan bahwa titik-titik sebuah pola menyebar dari data yang ada dan tidak membentuk pola tertentu yang jelas. Lalu dapat disimpulkan pula dari hasil scatterplot diatas merupakan hasil imputasi berdasarkan usia menggunakan metode kernel smoothing dengan cenderung menghasilkan nilai dugaan yang hampir sama. Hal ini mungkin karena pemilihan nilai h yang relatif besar atau dapat disimpulkan bahwa nilai dugaan menghasilkan nilai yang relatif stabil.

Kata Kunci : COVID-19; Kernel Smoothing; Data Tersensor.

ABSTRACT

This study aims to explain cases of length of time patients are infected with COVID-19 and apply the kernel smoothing regression method for censored data imputation. The data analysis method used in this study is the kernel smoothing method. The kernel smoothing method is a smoothing method using the kernel function. This method can be used as an alternative to solving fluctuating data problems. The data used is the length of time the patient is infected with covid-19 until recovered (Y) and the patient's age (X) from the results of the kernel smoothing method, the results of the imputation scatterplot based on case and age show that the dots are a pattern spread from the existing data and does not form a clear pattern. Then it can also be concluded from the results of the scatterplot above that the results of imputation based on age using the kernel smoothing method tend to produce almost the same estimated values. This may be due to the selection of a relatively large h value or it can be concluded that the estimated value produces a relatively stable value.

Keywords : COVID-19; Kernel Smoothing; Censored.

© 2023 Jurnal Riset Statistika Unisba Press. All rights reserved.

A. Pendahuluan

Analisis regresi merupakan salah satu metode statistika yang digunakan untuk menggambarkan hubungan antara dua atau lebih variabel, sehingga suatu variabel dapat diprediksikan dari variabel yang lain [1]. Variabel yang ditaksir nilainya disebut variabel respon, sedangkan variabel penaksir disebut sebagai variabel prediktor [2]. Dalam analisis regresi terdapat dua jenis pendekatan dalam menentukan kurva regresi, yaitu pendekatan parametrik dan nonparametrik. Jika bentuk kurva regresi diketahui, maka digunakan pendekatan parametrik. Pendekatan nonparametrik merupakan pendekatan regresi yang sesuai untuk pola data yang tidak diketahui bentuknya, atau tidak terdapat informasi masa lalu tentang pola data [3].

Dalam model regresi parametrik terdapat asumsi yang harus terpenuhi yaitu bentuk kurva regresinya (pola hubungan antara variabel respons dan variabel prediktor) diketahui mengikuti pola rumus tertentu [4]. Sedangkan regresi nonparametrik merupakan pendekatan metode regresi dimana bentuk kurva dari fungsi regresinya tidak diketahui pola rumusnya. Dalam regresi nonparametrik kurva regresi hanya diasumsikan mulus (*smooth*) dalam arti termuat dalam suatu ruang fungsi tertentu sehingga mempunyai fleksibilitas yang tinggi [5]. Dalam hal ini, teknik pendekatan regresi nonparametrik bias menjadi alternatif karena penggunaannya tidak terikat pada asumsi-asumsi seperti dalam regresi parametrik.

Pendekatan nonparametrik digunakan ketika informasi mengenai bentuk kurva regresi terbatas atau tidak ada asumsi tentang bentuk kurva regresi. Pendekatan regresi nonparametrik telah banyak dikembangkan antara lain menggunakan *spline*, *kernel*, *polynomial local*, *wavelet*, dan *fourier* [6]. Salah satu model regresi dengan pendekatan nonparametrik yang sering digunakan untuk melakukan estimasi terhadap kurva regresi adalah regresi *kernel smoothing* [7]. Salah satu hal yang sangat penting dalam peramalan menggunakan *smoothing kernel* adalah pemilihan kernel yang tepat [8] [9]. Tujuan dari *smoothing* adalah untuk membuang variabilitas dari data yang tidak memiliki efek sehingga ciri-ciri dari data akan tampak lebih jelas [10].

Pada akhir tahun 2019, sejenis virus baru diidentifikasi sebagai penyebab dari sekumpulan kasus pneumonia di kota Wuhan, China, yang kemudian menyebar cepat ke seluruh dunia [11]. *World Health Organization* (WHO) memberi nama *Severe Acute Respiratory Syndrome CoronaVirus-2* (SARS-CoV-2) terhadap virus tersebut dan nama penyakitnya sebagai *Coronavirus disease 2019* (Covid-19) [12]. Wabah *Corona Virus Disease 19* (COVID-19) ditetapkan sebagai pandemi global setelah diumumkan oleh WHO atau Badan Kesehatan Dunia dan penyebarannya yang sangat cepat menjadikan *COVID-19* sebagai topik utama di seluruh penjuru dunia [13]. Salah satunya adalah di Negara Singapura karena jumlah masyarakat yang terinfeksi *COVID-19* mengalami peningkatan setiap harinya.

COVID-19 adalah virus yang menginfeksi saluran pernapasan yang dapat bergerak antar manusia melalui kontak langsung [14]. Virus ini sangat mirip dengan wabah *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS) [15]. *COVID-19* ini sangat berbahaya karena menginfeksi saluran pernapasan dan menular dengan cepat [16]. Penularan virus dengan cepat akan berdampak terhadap prevalensi kasus dan akan berkembang cepat. *Coronavirus* merupakan virus RNA strain tunggal positif, berkapsul, dan tidak bersegmen. *Coronavirus* tergolong ordo *Nidovirales*, keluarga *Coronaviridae* [17].

Prevalensi wabah *Coronavirus* hampir di semua negara mengalami peningkatan. Hingga 10 Desember 2020 prevalensi kasus tersebut bertambah di seluruh dunia melebihi 60 juta orang [18]. Menurut *European Centre for Disease Prevention and Control* (2020) [19] kasus konfirmasi terbanyak yaitu Amerika dengan 14 juta, India 9 juta, Brazil 6 juta, dan Rusia lebih dari 2 juta kasus. Sedangkan di Singapura jumlah kasus konfirmasi *COVID-19* adalah 3.000 ribu orang per tanggal 21 April. Virus ini menyebar dengan cepat di seluruh dunia termasuk Singapura.

Pandemi *COVID-19* mengakibatkan kehilangan banyak jiwa. Menurut *Johns Hopkins University* (2020) melaporkan kematian akibat wabah ini sejak Januari hingga Desember 2020 lebih dari 1,5 juta jiwa diseluruh dunia. Tingginya tingkat kematian di Singapura dipengaruhi oleh keberadaan penyakit penyerta yang dimiliki oleh pasien positif *COVID-19* yaitu usia rentan, dan fasilitas kesehatan yang kurang memadai. Besarnya tingkat kematian hari demi hari tidak hanya menimbulkan gejala dan penyakit fisik saja akan tetapi, berpengaruh besar terhadap kesejahteraan masyarakat Singapura yang didalamnya mencakup kesehatan mental.

B. Metode Penelitian

Pada bagian ini akan dipaparkan mengenai tahapan analisis dengan menggunakan metode kernell smoothing, sebagai berikut: 1) Melakukan analisis deskriptif dari variabel penelitian. 2) Susun matriks pembobot **W**. 3) Buat matriks **W*** yang merupakan anak matriks **W** yang berisi pengamatan data lengkap. 4) Melakukan imputasi data tersensor data yang diperoleh sebelum semua data teramati waktu hidupnya [20]) dengan menggunakan metode *kernel smoothing* sebagai berikut: Lakukan pendugaan *kernel smoothing* untuk respon data lengkap, Lakukan imputasi menggunakan metode *kernel smoothing* untuk respon data tersensor sebagai berikut: a) Untuk setiap respon data tersensor, yakni untuk setiap *i* dilmana $\delta_i = 0$ ambillah nilai varilabel penjelasnya x_i . b) Nilai imputasinya diperoleh dengan metode *kernel smoothing* terhadap pasangan data (x_i, y_i) dimana untuk individu *i* yang merupakan data lengkap, nilai y_i -nya diganti oleh \hat{z}_i^{KS} . 5) Sajikan data tersensor dengan hasil imputasinya. 6) Buatlah plot hasil imputasi.

C. Hasil dan Pembahasan

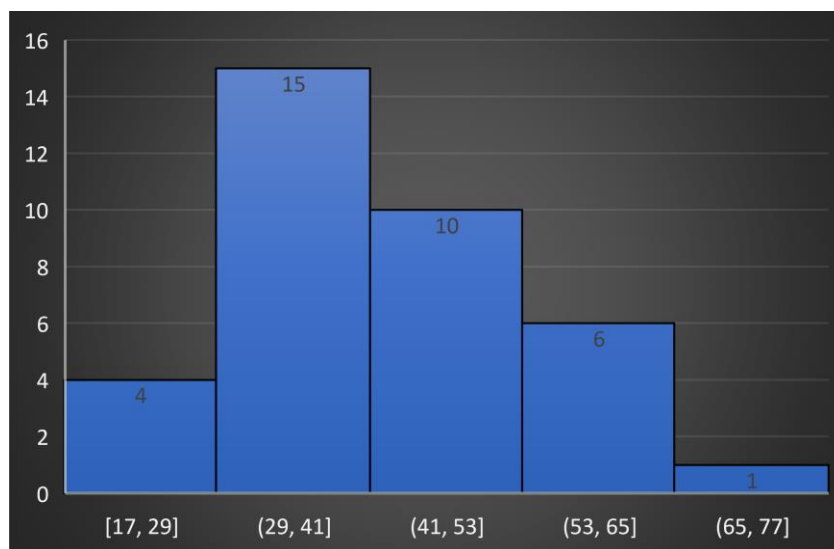
Dalam pembahasan ini akan menjelaskan mengenai hasil dari penelitian yang sudah dianalisis atau dilakukan, yaitu penerapan metode regresi *kernel smoothing* untuk imputasi data lama waktu terinfeksi covid-19. Selanjutnya pada bab ini akan membahas deskripsi statistik mengenai variabel-variabel yang ada pada data tersebut berdasarkan langkah-langkah yang telah dijelaskan pada metode penelitian.

Deskripsi Data Usia

Tabel 1. Analisis Deskriptif Berdasarkan Usia Pasien Terinfeksi Covid-19.

		Descriptive Statistics			
	N	Minimum	Maximum	Mean	Std. Deviation
Usia	36	17	66	42.22	11.623

Dapat disimpulkan dari hasil diatas menggunakan analisis deskriptif bahwa banyaknya individu 36 pasien. Dengan nilai minimumnya adalah 17 tahun, nilai maksimumnya adalah 66 tahun, nilai rata-ratanya adalah 42.22 tahun dan nilai simpangan bakunya adalah 11.623 tahun.



Gambar 1. Usia Pasien Terinfeksi Covid-19.

Pada gambar diatas dapat di deskripsikan bahwa pada usia 17-29 ada sebanyak 4 pasien yang terinfeksi covid-19. Dan pada usia 29-41 ada sebanyak 15 pasien yang terinfeksi covid-19. Pada usia 41-53 ada sebanyak 10 pasien yang terinfeksi covid-19. Pada usia 53-65 ada sebanyak 6 pasien yang terinfeksi covid-19. Dan yang terakhir pada usia 65-77 ada sebanyak 1 pasien yang terinfeksi covid-19. Maka total pasien yang terinfeksi

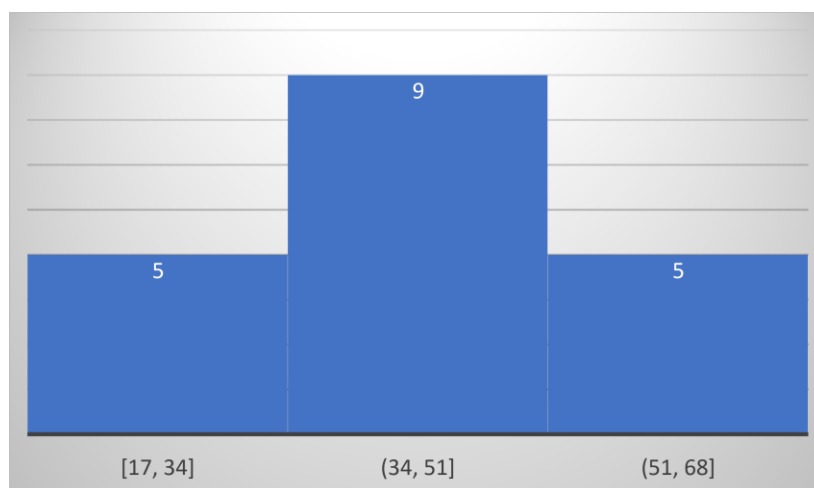
covid-19 ada sebanyak 36 pasien yang terinfeksi covid-19. Dengan jumlah minimum ada pada usia 65-77 dengan jumlah pasien 1 orang pasien yang terinfeksi covid-19 dan jumlah maksimumnya ada pada usila 29-41 dengan jumlah pasien 15 pasien yang terinfeksi covid-19. Dapat disimpulkan bahwa usia yang rentan terinfeksi covid-19 ada pada usia 29-41 dengan jumlah 15 pasien.

Deskripsi Data Usia dengan Data Lengkap

Tabel 2. Analisis Deskripsi Berdasarkan Usia Dengan Data Lengkap

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
usia_Data_Lengkap	19	17	66	41.58	12.773

Dapat disimpulkan dari hasil diatas menggunakan analisis deskripsi bahwa hasil diatas merupakan data usia dengan data lengkap atau yang dimaksud sudah sembuh dengan jumlah data usia ada 19 pasien. Dengan nilai minimumnya adalah 17 tahun, nilai maksimumnya adalah 66 tahun, nilai rata-rata nya adalah 41.58 tahun dan nilai simpangan bakunya adalah 12.773 tahun.



Gambar 2. Usia Pasien Terinfeksi Covid-19 Dengan Data Lengkap.

Pada gambar diatas merupakan histogram dari data usia pasien terinfeksi covid-19 dengan data lengkap atau pasien yang sudah sembuh dari hasil diatas dapat disimpulkan bahwa pada usia 17-34 ada sebanyak 5 pasien. Dan pada usia 34-51 ada sebanyak 9 pasien. Dan yang terakhir pada usia 51-68 ada sebanyak 5 pasien. Maka total pasien yang terinfeksi covid-19 dengan data lengkap atau sudah sembuh ada sebanyak 19 pasien. Dengan nilai minimumnya ada pada usia 17-34 dan 51-68 dengan sebanyak 5 pasien dan nilai maksimumnya ada pada usia 34-51 dengan sebanyak 9 pasien. Dan dapat disimpulkan bahwa usia yang rentan terinfeksi covid-19 dengan data yang lengkap atau data yang sudah sembuh ada pada usia 34-51 dengan jumlah 9 pasien.

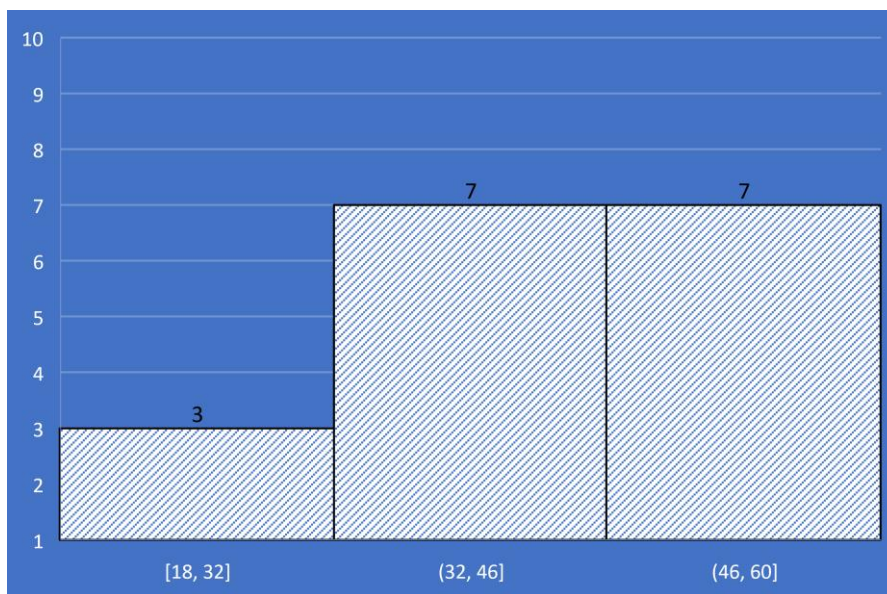
Deskripsi Data Usia dengan Data Tersensor

Tabel 3. Analisis Deskripsi Berdasarkan Usia Dengan Data Tersensor.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Usia_Data_tersensor	17	18	56	42.94	10.533

Dapat disimpulkan dari hasil diatas menggunakan analisis deskripsi bahwa hasil diatas merupakan data usia dengan data tersensor atau yang dimaksud pasien belum sembuh dengan jumlah data usia ada 17 pasien.

Dengan nilai minimumnya adalah 18 tahun, nilai maksimumnya adalah 56 tahun, nilai rata-rata nya adalah 42.94 tahun dan nilai simpangan bakunya adalah 10.533 tahun [21].



Gambar 3. Usia Pasien Terinfeksi Covid-19 Dengan Data Tersensor.

Pada gambar diatas merupakan histogram dari data usia pasien terinfeksi covid-19 dengan data tersensor atau pasien yang belum sembuh dari hasil diatas dapat disimpulkan bahwa pada usia 18-32 ada sebanyak 3 pasien. Dan pada usia 32-46 ada sebanyak 7 pasien. Dan yang terakhir pada usia 46-60 ada sebanyak 7 pasien. Maka total pasien yang terinfeksi covid-19 dengan data lengkap atau sudah sembuh ada sebanyak 17 pasien. Dengan nilai minimumnya ada pada usia 18-32 dengan sebanyak 3 pasien dan nilai maksimumnya ada pada usia 32-46 dan 46-60 dengan sebanyak 7 pasien. Dan dapat disimpulkan bahwa usia yang rentan terinfeksi covid-19 dengan data yang tersensor atau data yang belum sembuh ada pada usia 32-60 dengan jumlah 7 pasien.

Analisis Regresi Kernel Smoothing

Pada bab ini, dilakukan analisis regresi *kernel smoothing* untuk imputasi data lama waktu terinfeksi covid-19 di Singapura [22]. Data yang diinputkan merupakan matriks dari data lama waktu pasien terinfeksi covid-19 di Singapura dengan faktor-faktor yang mempengaruhi yaitu usia pasien atau Variabel X. Lalu, dalam data tersebut dengan variabel Y nya adalah lama waktu pasien terinfeksi covid-19 sampai sembuh atau tersensor. Berikut dibawah ini merupakan langkah-langkah dari perhitungan analisis regresi *kernel smoothing*: Pertama, Susun Matriks Pembobot **W**. Berdasarkan sesuai unsur-unsurnya. Maka susunan matriks tersebut sebagai berikut:

$$\mathbf{W}_{36 \times 36} = \begin{bmatrix}
 \mathbf{W1} & \mathbf{W2} & \mathbf{W3} & \dots & \mathbf{W34} & \mathbf{W35} & \mathbf{W36} \\
 0,10452 & 0,05952 & 0,04454 & \dots & 0,01099 & 0,00767 & 0,05951 \\
 0,02731 & 0,04797 & 0,04797 & \dots & 0,02731 & 0,02266 & 0,04796 \\
 0,00245 & 0,01727 & 0,04039 & \dots & 0,03933 & 0,04039 & 0,01726 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 0,00412 & 0,02233 & 0,03909 & \dots & 0,03922 & 0,03869 & 0,02233 \\
 0,00299 & 0,01932 & 0,01742 & \dots & 0,04034 & 0,04088 & 0,01931 \\
 0,02731 & 0,04797 & 0,04797 & \dots & 0,02731 & 0,02266 & 0,04796
 \end{bmatrix}$$

Kedua, Matriks **W***. Matriks **W*** yang merupakan anak matriks **W** yang berisi pengamatan data lengkap. Maka susunan matriks **W*** sebagai berikut:

$$W^*_{19 \times 19} = \begin{bmatrix} W1 & W2 & W3 & \dots & W26 & W27 & W30 \\ 0,10452 & 0,05952 & 0,04454 & \dots & 0,01533 & 0,02404 & 0,00065 \\ 0,02731 & 0,04797 & 0,04797 & \dots & 0,03205 & 0,03875 & 0,00504 \\ 0,00245 & 0,01727 & 0,04039 & \dots & 0,03766 & 0,03274 & 0,02904 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0,00576 & 0,02627 & 0,03175 & \dots & 0,03931 & 0,03815 & 0,01857 \\ 0,00940 & 0,03302 & 0,01227 & \dots & 0,03965 & 0,04086 & 0,01388 \\ 0,00036 & 0,00612 & 0,02480 & \dots & 0,02750 & 0,01977 & 0,05820 \end{bmatrix}$$

Ketiga, Melakukan imputasi data tersensor dengan menggunakan metode *kernel smoothing* sebagai berikut: Langkah selanjutnya, lakukan perhitungan pendugaan *kernel smoothing* untuk respon data lengkap. Sebagai berikut:

$$\hat{z}^{KS}_{19 \times 1} = W^* z$$

$$\begin{bmatrix} W1 & W2 & W3 & \dots & W26 & W27 & W30 \\ 0,10452 & 0,05952 & 0,04454 & \dots & 0,01533 & 0,02404 & 0,00065 \\ 0,02731 & 0,04797 & 0,04797 & \dots & 0,03205 & 0,03875 & 0,00504 \\ 0,00245 & 0,01727 & 0,04039 & \dots & 0,03766 & 0,03274 & 0,02904 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0,00576 & 0,02627 & 0,03175 & \dots & 0,03931 & 0,03815 & 0,01857 \\ 0,00940 & 0,03302 & 0,01227 & \dots & 0,03965 & 0,04086 & 0,01388 \\ 0,00036 & 0,00612 & 0,02480 & \dots & 0,02750 & 0,01977 & 0,05820 \end{bmatrix} \times \begin{bmatrix} z \\ 1 \\ 4 \\ 2 \\ \vdots \\ 2 \\ 9 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \hat{z}^{KS} \\ 1,64 \\ 2,20 \\ 3,51 \\ \vdots \\ 3,06 \\ 2,55 \\ 4,53 \end{bmatrix}$$

Keempat, Lakukan imputasi menggunakan metode *kernel smoothing* untuk respon data tersensor sebagai berikut: Untuk setiap respon data tersensor, yakni untuk setiap i dimana $\delta_i = 0$ ambillah nilai variabel penjelasnya x_{ij} . Nilai imputasinya diperoleh dengan metode *kernel smoothing* terhadap pasangan data (x_{ij}, y_{ij}) dimana untuk individu i yang merupakan data lengkap, nilai y_{ij} -nya diganti oleh \hat{z}_i^{KS} .

$$\hat{g}_{36 \times 1} = W y$$

$$\begin{bmatrix} W1 & W2 & W3 & \dots & W34 & W35 & W36 \\ 0,10452 & 0,05952 & 0,04454 & \dots & 0,01099 & 0,00767 & 0,05951 \\ 0,02731 & 0,04797 & 0,04797 & \dots & 0,02731 & 0,02266 & 0,04796 \\ 0,00245 & 0,01727 & 0,04039 & \dots & 0,03933 & 0,04039 & 0,01726 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0,00412 & 0,02233 & 0,03909 & \dots & 0,03922 & 0,03869 & 0,02233 \\ 0,00299 & 0,01932 & 0,01742 & \dots & 0,04034 & 0,04088 & 0,01931 \\ 0,02731 & 0,04797 & 0,04797 & \dots & 0,02731 & 0,02266 & 0,04796 \end{bmatrix} \times \begin{bmatrix} y \\ 1,64 \\ 2,20 \\ 3,51 \\ \vdots \\ 9 \\ 9 \\ 19 \end{bmatrix}$$

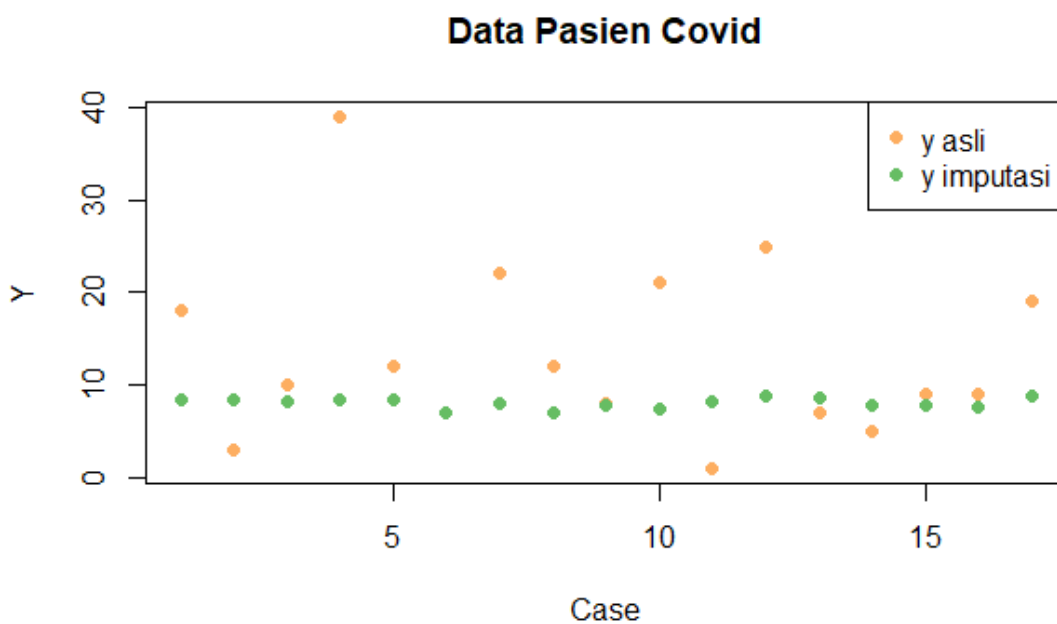
$$= \begin{bmatrix} y \\ 7,81 \\ 8,82 \\ 7,57 \\ \vdots \\ 7,73 \\ 7,62 \\ 8,82 \end{bmatrix}$$

Kelima, Sajikan data tersensor dan hasil imputasinya. Keenam, Setelah melakukan perhitungan imputasi disajikannya data pada tabel 4 merupakan data tersensor dan data hasil imputasinya dengan jumlah *case* data tersensor ada 17 *case*. Berikut dibawah ini data tersensor dan hasil imputasi:

Tabel 4. Data tersensor dan hasil imputasi.

i	usia	y asli	y imputasi
Case 08	56	18	84,791
Case 09	56	3	84,791
Case 10	56	10	83,046
Case 15	47	39	84,277
Case 17	47	12	84,277
Case 19	28	7	70,168
Case 21	44	22	80,294
Case 24	32	12	70,894
Case 25	40	8	77,369
Case 28	18	21	74,099
Case 29	41	1	82,118
Case 31	53	25	88,408
Case 32	42	7	86,090
Case 33	39	5	78,069
Case 34	40	9	77,369
Case 35	38	9	76,228
Case 36	53	19	88,259

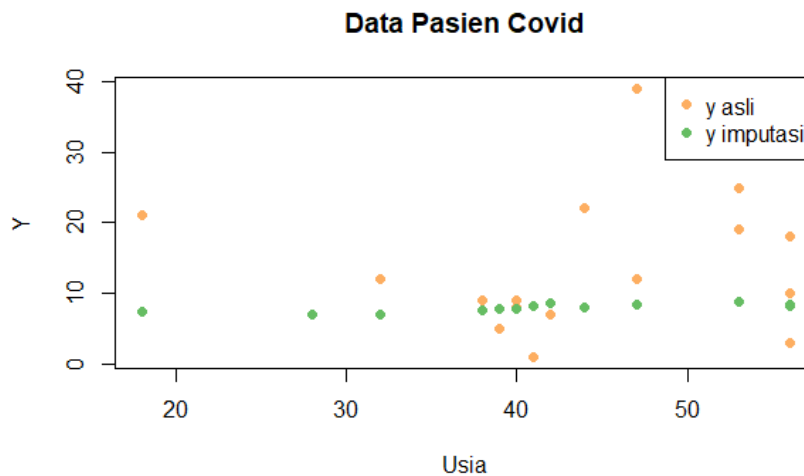
Ketujuh, Plot hasil imputasi. Membuat plot kesesuaian untuk membandingkan antara nilai respon asli dengan hasil imputasi menggunakan metode *kernel smoothing*. Plot hasil imputasi berdasarkan *case*:



Gambar 4. Plot Hasil Imputasi Berdasarkan *Case*

Dengan hasil output *scatterplot* diatas, dapat dilihat bahwa dari hasil plot imputasi berdasarkan *case* dengan titik-titik sebuah pola menyebar dari data yang ada dan tidak membentuk pola tertentu yang jelas. Pola penyebaran data yang berupa titik-titik pada *scatterplot* menyebar diatas dan dibawah. Lalu dapat disimpulkan pula dari hasil *scatterplot* diatas merupakan hasil imputasi menggunakan metode *kernel smoothing* dengan

cenderung menghasilkan nilai imputasi yang hampir sama. Hal ini mungkin karena pemilihan nilai h yang relatif besar atau dapat disimpulkan bahwa nilai imputasi menghasilkan nilai yang relatif stabil. Plot hasil imputasi berdasarkan usia:



Gambar 5. Plot Hasil Imputasi Berdasarkan Usia

Dengan hasil output *scatterplot* diatas, dapat dilihat bahwa dari hasil plot imputasi berdasarkan usia dengan titik-titik sebuah pola menyebar dari data yang ada dan tidak membentuk pola tertentu yang jelas. Pola penyebaran data yang berupa titik-titik pada *scatterplot* menyebar diatas dan dibawah. Lalu dapat disimpulkan pula dari hasil *scatterplot* diatas merupakan hasil imputasi berdasarkan usia menggunakan metode *kernel smoothing* dengan cenderung menghasilkan nilai imputasi yang hampir sama. Hal ini mungkin karena pemilihan nilai h yang relatif besar atau dapat disimpulkan bahwa nilai imputasi menghasilkan nilai yang relatif stabil.

D. Kesimpulan

Dari hasil penelitian diatas dapat disimpulkan sebagai berikut: 1) Hasil analisis deskriptif berdasarkan usia dengan jumlah data 36 pasien. Memiliki nilai minimum yaitu 17 tahun dan nilai maksimum yaitu 66 tahun. Dengan nilai rata-rata yaitu 42.22 tahun dan nilai simpangan baku yaitu 11.623 tahun. Untuk kelompok usia 17-29 tahun ada sebanyak 4 pasien, kelompok usia 29-41 tahun ada sebanyak 15 pasien, kelompok usia 41-53 tahun ada sebanyak 9 pasien, kelompok usia 53-65 tahun ada sebanyak 6 pasien, kelompok usia 65-77 tahun ada sebanyak 1 pasien. Hal ini menunjukkan bahwa usia yang rentan terinfeksi covid-19 ada pada usia 29-41 tahun dengan jumlah 15 pasien. 2) Hasil analisis deskriptif berdasarkan usila dengan data lengkap atau yang dimaksud sudah sembuh memiliki 19 pasien. Dengan nilai minimum yaitu 17 tahun dan nilai maksimum yaitu 66 tahun. Memiliki nilai rata-rata yaitu 41.58 tahun dan nilai simpangan baku yaitu 12.773 tahun. Untuk kelompok usia 17-34 tahun ada 5 pasien, kelompok usia 34-51 tahun ada 9 pasien, kelompok usia 51-68 tahun ada 5 pasien. Hal ini menunjukkan bahwa usia yang rentan terinfeksi covid-19 ada pada usia 34-51 tahun dengan jumlah 9 pasien. 3) Hasil analisis deskriptif berdasarkan usia dengan data tersensor atau yang dimaksud belum sembuh memiliki 17 pasien. Dengan nilai minimum yaitu 18 tahun dan nilai maksimum yaitu 56 tahun. Memiliki nilai rata-rata yaitu 42.94 tahun dan nilai simpangan baku yaitu 10.533 tahun. Untuk kelompok usia 18-32 tahun ada sebanyak 3 pasien, kelompok usia 32-46 tahun ada sebanyak 7 pasien, kelompok usia 46-60 tahun ada sebanyak 7 pasien. Hal ini menunjukkan bahwa usia yang rentan terinfeksi covid-19 ada pada usia 32-60 tahun dengan jumlah 7 pasien. 4) Hasil *scatterplot* imputasi berdasarkan *case* dan usia menunjukkan bahwa titik-titik sebuah pola menyebar dari data yang ada dan tidak membentuk pola tertentu yang jelas. Pola penyebaran data yang berupa titik-titik pada *scatterplot* menyebar diatas dan dibawah. Lalu dapat disimpulkan pula dari hasil *scatterplot* diatas merupakan hasil imputasi berdasarkan usia dan *case* menggunakan metode

kernel smoothing dengan cenderung menghasilkan nilai imputasi yang hampir sama. Hal ini mungkin karena pemilihan nilai h yang relatif besar atau dapat disimpulkan bahwa nilai imputasi menghasilkan nilai yang relatif stabil.

Daftar Pustaka

- [1] R. Syamsidah and A. Kudus, "Penerapan Metode Restricted Cubic Spline pada Kasus Data Riwayat Pasien Terinfeksi Virus Corona," doi: 10.29313/v7i1.26272.
- [2] M. Yahya Matdoan, M. Wiston Talakua, and R. J. Djami, "PEMODELAN REGRESI QUANTIL DENGAN KERNEL SMOOTHING PADA FAKTOR-FAKTOR YANG MEMPENGARUHI PENYEBARAN API MALARIA DI INDONESIA," *Uniqbu Journal of Exact Sciences*, vol. 1, no. 2, pp. 1–9, 2020.
- [3] I. N. Budiantara, *Estimator spline terbobot dan spline parsial terbobot dalam regresi nonparametrik dan semiparametrik heteroskedastik untuk data longitudinal*. 2010.
- [4] D. A. Pratiwi, I. N. Budiantara, and W. Wibowo, "Pendekatan Regresi Semiparametrik Spline untuk Memodelkan Rata-Rata Umur Kawin Pertama (UKP) di Provinsi Jawa Timur," *Jurnal Sains dan Seni ITS*, vol. 6, no. 1, pp. 129–136, 2017.
- [5] Winarti and S. Sony, "Pendekatan Regresi Semiparametrik Spline (Pada data nilai Ujian Nasional siswa SMKN 1 Nguling Pasuruan)," *Jurnal Sains dan Seni Pomits*, vol. 3, no. 2, pp. 194–199, 2010.
- [6] R. L. Eubank, "Nonparametric Regression and Spline Smoothing," *Marcel Dekker Inc.*, 1999.
- [7] A. T. Laia, "Model Estimasi Regresi Nonparametrik Dengan Metode Kernel," Universitas Sumatera Utara, Medan, 2015.
- [8] A. Faizal Arifin and E. Tri Astuti, "Pemodelan Nilai Ekspor Kelapa Sawit Di Indonesia Menggunakan Smoothing Kernel," 2020.
- [9] D. Aydin and Er. Yilmaz, "KERNEL SMOOTHING AS AN IMPUTATION TECHNIQUE FOR RIGHT-CENSORED DATA," *Eskişehir Technical University Journal of Science and Technology A - Applied Sciences and Engineering*, vol. 21, pp. 1–6, Nov. 2020, doi: 10.18038/estubtda.817979.
- [10] S. Halim and I. Bisono, "Fungsi-Fungsi Kernel Pada Metode Regresi Nonparametrik Dan Aplikasinya Pada Priest River Experimental Forest's Data," *Jurnal Teknik Industri*, vol. 8, no. 1, pp. 73–81, 2006.
- [11] R. Elviani, C. Anwar, and R. Januar Sitorus, "GAMBARAN USIA PADA KEJADIAN COVID-19," *JAMBI MEDICAL JOURNAL "Jurnal Kedokteran dan Kesehatan"*, vol. 9, no. 1, pp. 204–209, May 2021, doi: 10.22437/jmj.v9i1.11263.
- [12] M. D. Simatupang and I. M. Arcana, "Risiko Kematian Pasien Covid-19 dan Faktor yang Memengaruhinya," *Seminar Nasional Official Statistics*, vol. 2021, no. 1, pp. 889–898, Nov. 2021, doi: 10.34123/semnasoffstat.v2021i1.1085.
- [13] R. Mantovani, "PENGARUH PANDEMI COVID-19 TERHADAP TINGKAT KEMISKINAN DI KOTA MAKASSAR," Universitas Muhammadiyah Makassar, Makassar, 2021.
- [14] W. Hariadi and S. Sulantari, "ANALISIS SURVIVAL LAMA WAKTU SEMBUH PASIEN COVID-19 DENGAN METODE KAPLAN-MEIER DAN LOG-RANK DI KABUPATEN JEMBER," *Transformasi : Jurnal Pendidikan Matematika dan Matematika*, vol. 5, no. 1, pp. 415–425, May 2021, doi: 10.36526/tr.v5i1.1000.
- [15] L. Guo *et al.*, "Profiling Early Humoral Response to Diagnose Novel Coronavirus Disease (COVID-19)," *Clinical Infectious Diseases*, vol. 71, no. 15, pp. 778–785, Jul. 2020, doi: 10.1093/cid/ciaa310.
- [16] S. Sulantari and W. Hariadi, "ANALISIS SURVIVAL WAKTU SEMBUH PASIEN COVID-19 DI KABUPATEN BANYUWANGI," *Transformasi : Jurnal Pendidikan Matematika dan Matematika*, vol. 4, no. 2, pp. 375–386, Dec. 2020, doi: 10.36526/tr.v4i2.1001.

- [17] B. Audina and M. Fatekurohman, “Analisis Survival pada Data Pasien Covid 19 di Kabupaten Jember (Survival Analysis on Covid 19 Patients Data in Jember Regency),” *BERKALA SAINSTEK*, vol. 4, pp. 118–121, 2020.
- [18] World Health Organization, “WHO Coronavirus Disease (COVID-19) Dashboard,” *World Health Organization*, 2020.
- [19] European Centre for Disease Prevention and Control, “Outbreak of Acute Respiratory Syndrome Associated with a Novel Coronavirus, Wuhan, China; first update,” *Rapid Risk Assessment*, 2020.
- [20] D. R. Sari, “ANALISIS SURVIVAL UNTUK DATA TERSENSOR TIPE II MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK,” Universitas Negeri Yogyakarta, Yogyakarta, 2011.
- [21] H. Hikmah and D. Ekawati, “Analisi Data Tersensor Berpasangan dengan Estimasi Kaplan Meier dan Nelson Aalen,” *SAINTIFIK*, vol. 7, no. 2, pp. 133–138, Jul. 2021, doi: 10.31605/saintifik.v7i2.334.
- [22] L. Anisa and N. A. K. Rifai, “Analisis Regresi Logistik Biner dengan Metode Penalized Maximum Likelihood pada Penyakit Covid-19 di RSUD Pringsewu,” *Jurnal Riset Statistika*, pp. 129–136, Dec. 2022, doi: 10.29313/jrs.v2i2.1425.