***MAXIMUM LIKELIHOOD ESTIMATOR* UNTUK MENGESTIMASI MODEL REGRESI ISOTONIK DENGAN PENDEKATAN POLINOMIAL BERNSTEIN PADA KASUS SATU VARIABEL INDEPENDEN**

**TUGAS AKHIR**


**Disusun Oleh:**

**Panji Satrio Kurniawan**

**14 611 155**

**PROGRAM STUDI STATISTIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS ISLAM INDONESIA**

**YOGYAKARTA**


**2018**

***MAXIMUM LIKELIHOOD ESTIMATOR* UNTUK MENGESTIMASI MODEL REGRESI ISOTONIK DENGAN PENDEKATAN POLINOMIAL BERNSTEIN PADA KASUS SATU VARIABEL INDEPENDEN**

**TUGAS AKHIR**

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Jurusan Statistika

**Disusun Oleh:**

**Panji Satrio Kurniawan**

**14 611 155**

**PROGRAM STUDI STATISTIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS ISLAM INDONESIA**

**YOGYAKARTA**

**2018**

**HALAMAN PERSETUJUAN PEMBIMBING**

**TUGAS AKHIR**

Judul : *Maximum Likelihood Estimator* untuk Mengetimasi Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein pada Kasus Satu Variabel Independen

Nama Mahasiswa : Panji Satrio Kurniawan

Nomor Mahasiswa: 14 611 155

**TUGAS AKHIR INI TELAH DIPERIKSA DAN DISETUJUI UNTUK DIUJIKAN**

Yogyakarta, Maret 2018

**Pembimbing**

**(Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc.)**

**HALAMAN PENGESAHAN**

**TUGAS AKHIR**

***MAXIMUM LIKELIHOOD ESTIMATOR* UNTUK MENGESTIMASI MODEL REGRESI ISOTONIK DENGAN PENDEKATAN POLINOMIAL BERNSTEIN PADA KASUS SATU VARIABEL INDEPENDEN**

**Nama Mahasiswa : Panji Satrio Kurniawan**

**Nomor Mahasiswa : 14 611 155**

**TUGAS AKHIR INI TELAH DIUJIKAN**

**MARET 2018**

| Nama Penguji | Tanda Tangan |
| --- | --- |
| 1. Annisa Uswatun Khasanah, ST., M.Sc | ................ |
| 2. Dr. Edy Widodo, M.Si | ................ |
| 3. M. Hasan Sidiq Kurniawan, S.Si., M.Sc. | ................ |

Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

UNIVERSITAS ISLAM INDONESIA YOGYAKARTA FAKULTAS MIPA

Drs. Allwar M.Sc., Ph.D

# KATA PENGANTAR

***Assalamualaikum Wr.Wb.***

Alhamdulillahirobbil alamin, Puji Syukur kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya selama melaksanakan Tugas Akhir sehingga dapat terselesaikan. Shalawat serta salam tercurah kepada Nabi Muhammad SAW beserta keluarga dan para pengikut-pengikutnya. Tugas Akhir yang berjudul "*Maximum Likelihood Estimator* untuk Mengestimasi Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein pada Kasus Satu Variabel Independen" ini sebagai salah satu syarat untuk memperoleh gelar sarjana Program Studi Statistika di Universitas Islam Indonesia. Selama menulis Tugas Akhir, penulis telah banyak mendapat bimbingan dan bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis bermaksud menyampaikan ucapan terima kasih kepada :

1. Bapak Drs. Allwar, M.Sc, Ph.D selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia, Yogyakarta beserta seluruh jajarannya.
2. Bapak Dr. RB. Fajriya Hakim, S.Si., M.Si, selaku Ketua Program Studi Statistika beserta seluruh jajarannya.
3. Bapak Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc, selaku Dosen Pembimbing penulis yang sangat berjasa dan sangat sabar membimbing penulis dalam menyelesaikan Tugas Akhir ini.
4. Dosen-dosen Statistika Universitas Islam Indonesia yang telah mendidik dan memberikan ilmunya kepada penulis serta selalu menginspirasi.
5. Kedua orang tua tersayang Bapak Firmansyah, S.Pd dan Ibu Lusi Haryati, A.Md.Keb., S.K.M yang selalu memberikan semangat, dukungan dan selalu mengiringi lewat untaian doa disetiap langkah penulis.
6. Kakak penulis, Rahayu Nofriyanti, A.Md.Keb yang dalam kata-kata terukir motivasi yang selalu membiru. Kakak Ipar penulis Ervin Zaenal, adik penulis, Septian Hadi Nugraha dan Keponakan penulis, Ibrahim Bahy Elfatih yang selalu membakar semangat dan selalu hadir disetiap perjalanan penulis.

7. Sahabat penulis, Salsabila, Leni dan Aulia yang selama ini menemani pengerjaan tugas akhir ini.
8. Teman-teman satu bimbingan Tugas Akhir (bimbingan Bapak Hasan) Ulin, Roni, Rima, Rati, Ajeng, Dhea, Indah, Marisa, Mia, Inayatus, Ellysa, Yusi, Tista dan Nilam yang selalu berbagi ilmu dan berbagi informasi.
9. Teman-teman KKN unit 416 Desa Sekaran, Dusun Jerukan, Kecamatan Wonosari, Kabupaten Klaten, Provinsi Jawa Tengah, Jody, Amri, Rengga, Zulfa, Dhila, Novita, Canda dan Ajeng, suka dan duka yang telah dilalui bersama tidak akan pernah terlupakan.
10. Sahabat Statistika 2014 (XIX) yang sudah banyak memberikan semangat dan motivasi dalam penyelesaian Tugas Akhir ini.
11. Mbak Ajeng, Mas Ardhian, Mas Doni dan Mas Bagus, terima kasih telah membuka tempat untuk berkeluh kesah dan memberi saran jika penulis mengalami kesulitan. Terima kasih juga telah menjadi kakak yang hebat, selalu memberikan informasi yang sangat berguna untuk perkembangan penulis.
12. Sahabatku Suryo, Pebri, Kiki, Atika, Riski dan Deyan, terima kasih perhatian kalian walau dari jauh sekalipun. Semoga kesuksesan dan mimpi kita di kabulkan Allah swt. Aamiin.
13. Semua pihak yang tidak dapat penulis sebutkan satu per satu, terima kasih.

Penulis menyadari bahwa Tugas Akhir ini masih jauh dari sempurna, oleh karena itu segala kritik dan saran yang bersifat membangun selalu penulis harapkan. Semoga Tugas Akhir ini dapat bermanfaat bagi penulis khususnya dan bagi semua yang membutuhkan. Akhir kata, semoga Allah SWT selalu melimpahkan rahmat serta hidayah-Nya kepada kita semua, Aamiin aamiin ya robbal’alamin.

***Wassalamu’alaikum, Wr.Wb***

Yogyakarta,   Maret 2018

Penulis

# DAFTAR ISI

## DAFTAR TABEL

## DAFTAR GAMBAR

# DAFTAR LAMPIRAN

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang sebelumnya pernah diajukan untuk memperoleh gelar kesarjanaan di sutau perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang diacu dalam naskah ini dan disebutkan daftar pustaka.

Yogyakarta, Maret 2018

[illegible]o Kurniawan

***MAXIMUM LIKELIHOOD ESTIMATOR*** **UNTUK MENGESTIMASI MODEL REGRESI ISOTONIK DENGAN PENDEKATAN POLINOMIAL BERNSTEIN PADA KASUS SATU VARIABEL INDEPENDEN**

Panji Satrio Kurniawan

Program Studi Statistika Fakultas MIPA

Universitas Islam Indonesia

## INTISARI

Salah satu fungsi Analisis statistika yaitu digunakan menentukan hubungan antara variabel dependen (*y*) dari variabel independen (*x*). Salah satu metode untuk menyelesaikan masalah tersebut yaitu analisis regresi. Analisis regresi nonparametrik yang didapat digunakan untuk menentukan model yang tepat dari variabel dependen (*y*) dan variabel independen (*x*) yaitu Regresi Isotonik. Regresi Isotonik digunakan jika hubungan antara variabel dependen (*y*) dan varibel independen (*x*) adalah monoton naik. Pengestimasian model Regresi Isotonik menggunakan *Maximum Likelihood Estimator* (MLE) karena marupakan estimasi teoritis yang lebih kuat dibandingkan dengan metode *Ordinary Least Square* (OLS). Model regresi isotonik dapat didekati menggunakan Polinomial Berstein. Polinomial Berstein dapat digunakan untuk mengestimasi persamaan regresi yang tidak diketahui. Banyak sifat-sifat baik dari Polinomial Bernstein yang sangat cocok digunakan sebagai aproksimasi model regresi. Analisis metode ini akan ditunjukkan melalui simulasi dan analisis menggunakan data nyata yaitu makanan jadi, minuman, rokok, dan tembakau (*y*) dan pendidikan, rekreasi, dan olahraga (*x*) dengan membandingkan *Maximum Likelihood Estimator* untuk Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein dengan Regresi Linear Sederhana. Berdasarkan analisis data tersebut didapatkan estimasi model Regresi Isotonik dengan pendekatan Polinomial Bernstein dan Regresi Linear Sederhana mempunyai selisih yang sangat kecil. Maka dapat disimpulkan bahwa estimasi model Isotonik dengan pendekatan Polinomial Bernstein dan Regresi Linear Sederhana mempunyai kebaikan yang sama. Oleh karena itu, estimasi model Regresi Isotonik dengan Polinomial Bernstein baik digunakan sebagai alternatif penentuan model pada regresi dengan plot data monoton naik.

**Kata Kunci :** Regresi Nonparamterik, Regresi Isotonik, *Maximum Likelihood Estimator*, Polinomial Bernstein.

***MAXIMUM LIKELIHOOD ESTIMATOR FOR REGRESSING ISOTONIC REGRESSION MODEL WITH BERNSTEIN POLINOMIAL APPROACH IN THE CASE ONE INDEPENDENT VARIABLE***

*Panji Satrio Kurniawan*

*Departement of Statistics, Faculty of Mathematics and Natural Sciences*

*Islamic University of Indonesia*

***ABSTRACT***

*One function of statistical analysis is the relationship between the dependent variable (y) of the independent variable (x). One method to overcome the problem is regression analysis. Nonparametric regression analysis was performed to determine the exact model of the dependent variable (y) and the independent variable (x) ie Isotonic Regression. Isotonic regression if the relationship between the dependent variable (y) and the independent variable (x) is a rising monotone. Estimating the Isotonic Regression model using Maximum Likelihood Estimator (MLE) is a much stronger theoretical estimate than the Ordinary Least Square (OLS) method. An isotonic regression model can be approximated using Polynomial Berstein. Polynomial Berstein can be used to estimate unknown regression equations. Many good properties of the Bernstein Polynomial are very easy to use as models of regression approximation. This analytical method will produce analysis and analysis using real data ie finished food, beverage, cigarette and tobacco (y) and education, recreation, and sport (x) by comparing probability of Maximum Possibility for Isotonic Regression Model with Bernstein Polynomial approach with Linear Regression Simple. Data analysis using Isotonic Regression model with Bernstein Polynomial approach and Simple Linear Regression have very small difference. It can be concluded that the estimation of the Isotonic model with the Bernstein Polynomial approach and Simple Linear Regression has the same good. Therefore, the estimation of the Isotonic Regression model with Bernstein Polynomial as a model of determination on the regression with the monotonous plot data increases.*

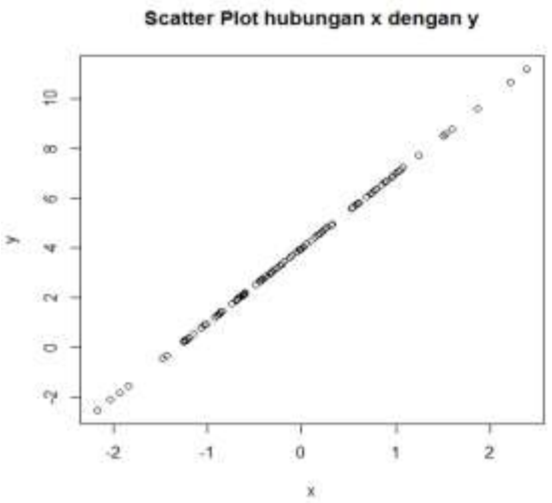
**Keywords:** *Nonparamteric Regression, Isotonic Regression, Maximum Likelihood Estimator, Bernstein Polynomial.*

# BAB I

# PENDAHULUAN

## 1.1 Latar Belakang Masalah

Statistika merupakan ilmu pengetahuan yang berhubungan dengan cara-cara pengumpulan data, pengolahan, penganalisisannya, penarikan kesimpulan serta pembuatan keputusan yang cukup beralasan berdasarkan fakta yang ada (Somantri, 2006). Dalam rangka menentukan hubungan antara variabel dependen ($y$) dari variabel independen ($x$) dapat pula digunakan analisis statistika. Salah satu metode untuk menyelesaikan masalah tersebut yaitu analisis regresi. Analisis regresi yang menganalisis hubungan antara satu variabel $x$ dan variabel $y$ disebut regresi linier sederhana, sedangkan regresi linier yang melibatkan lebih dari satu variabel independen disebut Regresi Linier Berganda (Sembiring, 1995). Regresi Linier Sederhana akan menghasilkan pola data dengan kurva berbentuk garis lurus. Salah satu pola hubungan antara variabel $x$ dan variabel $y$ adalah pola berbentuk monoton. Pola hubungan antar variabel dikatakan monoton jika hubungan antara variabel $x$ dan variabel $y$ selalu konstan sepanjang waktu. Pola monoton dibagi menjadi dua yaitu monoton naik dan monoton turun. *Scatter Plot* monoton naik dapat dilihat pada Gambar 1.1 sebagai berikut




**Gambar 1.1** *Scatter Plot* Monoton Naik

dengan:

$x$ = data distribusi normal yang dibangkitkan dari *sofware* R

$\hat{y}_i = 3 + 5\ x_i$

Pola data berbentuk monoton dapat ditemukan dibeberapa kasus, seperti kasus pada penelitian yang dilakukan oleh Adawiyah (2011) tentang hubungan antara indeks harga konsumen pada makanan jadi, minuman, rokok dan tembakau ($y$) dengan indeks harga konsumen pada pendidikan, rekreasi dan olahraga ($x$) pada 66 kota di Indonesia mulai bulan Mei 2006 sampai Mei 2008, dimana datanya memiliki pola yang terus naik kemudian turun pada bulan Juni 2008 tetapi kembali naik sampai Mei 2009. Untuk menganalisis hubungan antara variabel $x$ dan variabel $y$ yang memiliki pola monoton naik, dapat digunakan Regresi Isotonik. Regresi Isotonik merupakan regresi nonparametrik yang baik dipakai untuk menganalisis pola hubungan monoton naik karena galat yang dihasilkan akan lebih kecil dibandingkan Regresi Linier sederhana (Setyawan, 2017). Asumsi pada Regresi Isotonik adalah hubungan antara variabel $x$ dan variabel $y$ yang monoton naik. Pada Tugas Akhir ini akan dibahas tentang analisis data yang memiliki bentuk pola hubungan monoton naik menggunakan Regresi Isotonik Sederhana. Dalam analisis regresi, estimasi parameter regresi perlu dilakukan karena salah satu tujuan analisis tersebut adalah untuk menduga variabel $y$. Oleh karena itu, dalam tugas ini juga akan dibahas tentang estimasi parameter Regresi Isotonik tersebut.

Estimasi (pendugaan) merupakan bahasan statistika yang berhubungan dengan pendugaan nilai-nilai parameter berdasarkan data yang diukur/data empiris yang berasal dari sampel acak. Parameter merupakan suatu konstanta yang mencirikan karakteristik populasi. Kegiatan estimasi berupaya untuk mengaproksimasi parameter populasi tersebut menggunakan suatu ukuran yang berasal dari sampel (Hede, 2016). Estimasi parameter pada Regresi Linier Sederhana dapat diestimasi menggunakan estimasi *Ordinary Least Square* (OLS) dan *Maximum Likelihood Estimator* (MLE). Pada kasus kali ini peneliti akan mengestimasi persamaan Regresi Isotonik menggunakan MLE karena metode MLE sangat berhubungan dengan kemampuan numerik, terutama dalam menghasilkan titik penyelesaian dalam suatu persamaan (Yendra, 2015).

Model Regresi Isotonik dapat didekati menggunakan Polinomial Berstein. Polinomial Berstein dapat digunakan untuk mengestimasi persamaan regresi yang tidak diketahui. Banyak sifat-sifat baik dari Polinomial Bernstein yang sangat cocok digunakan sebagai aproksimasi model regresi. Sifat-sifat baik tersebut adalah estimatornya memenuhi *pointwise consistency, asymptotic normality*, dan *uniform consistency* (Curtis, 2010). Berdasarkan uraian di atas maka penulisan Tugas Akhir ini akan membahas mengenai estmasi parameter Regresi Isotonik menggunakan MLE dan pendekatan Polinomial Bernstein dari data yang dibangkitkan dari *software* R. Selanjutnya akan dilakukan analisis data menggunakan Regresi Linier Sederhana dan Regresi Isotonik supaya diperoleh perbandingan antara kedua model regresi tersebut. Model terbaik dipilih berdasarkan nilai *Mean Square Error* (MSE) dan R-*Square*.

## 1.2 Rumusan Masalah

Bedasarkan latar belakang masalah di atas, dapat dirumuskan beberapa hal sebagai berikut:

1. Apakah *Estimasi Likelihood Estimator* dapat digunakan untuk estimasi model Regresi Isotonik?
2. Bagaimana aplikasi model Regresi Isotonik pada data studi kasus?
3. Bagaimana kebaikan model Regresi Isotonik dibandingkan Regresi Linear Sederhana untuk data monoton naik?

## 1.3 Batasan Masalah

Untuk memudahkan penulis dalam menulis Tugas Akhir ini, maka penulis membatasi pembahasan yang akan dibahas yaitu :

1. Analisis ini hanya akan melakukan pemodelan Regresi Isotonik dengan satu variabel dependen dan satu variabel independen.
2. Perhitungan dalam analisis ini menggunakan *package* stats4 yang diolah menggunakan *software* R versi 3.3.3.
3. Data yang digunakan dalam studi kasus adalah data yang memiliki pola hubungan monoton naik.

4. Estimasi model Regresi Isotonik dilakukan menggunakan metode *Maximum Likelihood Estimator* dan pendekatan Polinomial Bernstein.

## 1.4 Jenis Penelitian dan Metode Analisis

Jenis penelitian dalam tugas akhir ini adalah penelitian teoritis, yang mengacu pada penelitian yang berjudul Estimasi Parameter Model Regresi Isotonik Bayesian dengan Polinomial Bernstein yang dilakukan oleh Setyawan (2017). Dalam penelitian tersebut, estimasi parameter Regresi Isotonik dilakukan menggunakan metode Bayes. Sedangkan pada skripsi ini, estimasi parameter akan dilakukan menggunakan MLE.

## 1.5 Tujuan Penelitian

Tujuan penelitian dalam Tugas Akhir ini adalah:

1. Mengimplementasikan *Estimasi Likelihood Estimator* untuk estimasi model Regresi Isotonik.
2. Mempelajari aplikasi model Regresi Isotonik pada data studi kasus.
3. Mengetahui kebaikan model Regresi Isotonik dibandingkan Regresi Linear Sederhana untuk data monoton naik.

## 1.6 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini adalah:

1. Dapat memperdalam pemahaman tentang Regresi Isotonik.
2. Memodelkan suatu kasus dengan pola data monoton naik menggunakan Regresi Isotonik dengan pendekatan Polinomial Berstein berdasarkan MLE.
3. Mengetahui perbedaan antara Regresi Linier Sederhana dan Regresi Isotonik.

# BAB II
# KAJIAN TEORI

## 2.1 Penelitian Terdahulu

Pada penelitian Misbahussurur (2009) dengan judul "Estimasi Parameter Distribusi Gamma dengan Metode *Maksimum Likelihood*" membahas tentang estimasi parameter distribusi gamma yakni $\alpha$ dan $\beta$ menggunakan Maksimum *Likelihood*. Didapatkan *E(x)* dari distribusi gamma adalah $\alpha\beta$ dan var($x$) dari distribusi gamma adalah $\alpha\beta^2$. Oleh karena itu, $\widehat{\alpha\beta}$ dan $\hat{\alpha}\hat{\beta}^2$ merupakan estimasi yang baik dan $\alpha\beta$ dan $\alpha\beta^2$ karena telah memenuhi sifat-sifat estimasi yakni takbias, konsisten dan efisien.

Selanjutnya, penelitian yang dilakukan oleh Nuraini (2012) dengan judul "Estimasi Fungsi Penghalus pada Regresi Isotonik Aditif dengan Metode Kuadrat Terkecil". Dapat disimpulkan dari penelitian tersebut bahwa estimasi fungsi penghalus pada Regresi Isotonik Aditif menggunakan Metode Kuadrat Terkecil diselesaikan melalui algoritma *Backfitting* yaitu $s_j x_j = s_j^r x_j$, dengan $s_j^r x_j$ dihitung menggunakan *Pooled Adjacent Violators Algorithm (PAVA)*. Secara geometri, estimasi fungsi penghalus dengan menggunakan Metode Kuadrat Terkecil dapat dipandang sebagai masalah proyeksi dalam suatu ruang vektor $R^n$.

Penelitian Fajaroh (2012) dengan judul "Regresi Isotonik untuk Menentukan Nilai Ambang (*Threshold Value*)". Dari penelitian tersebut disimpulkan estimasi fungsi Regresi Isotonik dengan menggunakan Metode Maksimum *Likelihood* yaitu $\hat{\mu}(x) = \bar{y}^*(x)$ atau rata-rata sampel dengan pembobot $w(x)$. Nilai ambang diperoleh apabila uji statistik pada fungsi regresi isotonik melebihi daerah kritis. Jika terdapat lebih dari satu observasi yang melebihi daerah kritis, maka nilai ambang diambil dari observasi terkecil. Kelebihan regresi isotonik dibandingkan dengan metode lain dalam menentukan nilai ambang adalah tidak diperlukan asumsi linearitas pada data. Pada skripsi Lulu Atul Fajaroh hanya digunakan estimasi *Maximum Likelihood Estimator* sedangkan skripsi ini menggunakan estimasi *Maximum Likelihood Estimator* dan pendekatan Polinomial Bernstein.

Penelitian Hede (2016) dengan judul “Perbandingan Metode Kuadrat Terkecil dan Metode Kemungkinan Maksimum dalam Pendugaan Parameter Distribusi Weibul dengan Dua Parameter”. Dari penelitian tersebut didapatkan perbandingan metode terbaik dalam menduga parameter distribusi Weibull dengan dua parameter digunakan perbandingan Rata-rata Kuadrat Galat (*Mean Square Error*). Metode yang terbaik adalah metode yang memiliki rata-rata kuadrat galat minimum. Metode terbaik dalam pendugaan menggunakan data rata-rata kecepatan angin di Enugu adalah Metode Kemungkinan Maksimum. Sedangkan pada data rata-rata kecepatan angin per bulan di Sumenep, metode terbaik adalah Metode Kuadrat Terkecil.

Penelitian yang dilakukan oleh Setyawan (2017) dengan judul “Estimasi Parameter Model Regresi Isotonik Bayesian dengan Polinomial Bernstein”. Dari hasil penelitian ini disimpulkan bahwa semakin besar orde dari Polinomial Bernstein maka semakin baik digunakan dalam menentukan hubungan antara variabel dependen dan independen karena menghasilkan MSE yang lebih kecil dan $R^2$ yang lebih besar. Namun, semakin besar orde dari Polinomial Bernstein akan menyebabkan simulasi menjadi semakin lama.

Pada Tabel 2.1 menjadi perbandingan penelitian sebelumnya dan penelitian yang akan dilakukan oleh penulis.

**Tabel 2.1** Perbandingan penelitian sebelumnya dengan penelitian yang penulis lakukan

| No | Nama/ Tahun | Metode | Judul Penelitian | Hasil Penelitian |
|---|---|---|---|---|
| 1 | Misbahussurur (2009) | Distribusi Gamma dan Estimasi *Maximum Likelihood* | Estimasi Parameter Distribusi Gamma dengan | Dari estimasi distribusi Gamma dengan menggunakan MLE didapatkan $\widehat{\alpha\beta}$ dan |

| No | Nama/ Tahun | Metode | Judul Penelitian | Hasil Penelitian |
|---|---|---|---|---|
| | | | Metode *Maksimum Likelihood* | $\hat{\alpha}\hat{\beta}^2$ merupakan estimasi yang baik untuk $\alpha\beta$ dan $\alpha\beta^2$ karena telah memenuhi sifat-sifat estimasi yakni takbias, konsisten dan efisien. |
| 2 | Nuraini (2012) | Regresi Isotonik Aditif, *Pooled Adjacent Violators Algorithm (PAVA)* dan Estimasi Metode Kuadrat Terkecil | Estimasi Fungsi Penghalus pada Regresi Isotonik Aditif dengan Metode Kuadrat Terkecil | Secara geometri, estimasi fungsi penghalus dengan menggunakan Metode Kuadrat Terkecil dapat dipandang sebagai masalah proyeksi dalam suatu ruang vektor $R^n$. |
| 3 | Fajaroh (2012) | Regresi Isotonik, *Pooled Adjacent Violators Algorithm (PAVA)*, uji rasio *Likelihood threshold* dan *Maximum* | Regresi Isotonik untuk Menentukan Nilai Ambang (*Threshold Value*) | Estimasi fungsi Regresi Isotonik dengan menggunakan Metode Maksimum *Likelihood* yaitu $\hat{\mu}(x) = \bar{y}^*(x)$ atau rata-rata sampel dengan pembobot $w(x)$. |

| No | Nama/ Tahun | Metode | Judul Penelitian | Hasil Penelitian |
|---|---|---|---|---|
| | | *Likelihood Estimator* | | |
| 4 | Hede (2016) | Metode Kuadrat Terkecil dan Metode Kemungkinan Maksimum | Perbandingan Metode Kuadrat Terkecil dan Metode Kemungkinan Maksimum dalam Pendugaan Parameter Distribusi Weibul dengan Dua Parameter | Berdasarkan perbandingan Rata-rata Kuadrat Galat (*Mean Square Error*) didapatkan metode terbaik dalam pendugaan menggunakan data rata-rata kecepatan angin di Enugu adalah Metode Kemungkinan Maksimum. Sedangkan pada data rata-rata kecepatan angin per bulan di Sumenep, metode terbaik adalah Metode Kuadrat Terkecil. |
| 5 | Setyawan (2017) | Regresi Isotonik, Bayesian, Polinomial Bernstein dan Metode *Markov* | Estimasi Parameter Model Regresi Isotonik | semakin besar orde dari Polinomial Bernstein maka semakin baik digunakan dalam menentukan |

| No | Nama/ Tahun | Metode | Judul Penelitian | Hasil Penelitian |
|---|---|---|---|---|
| | | *Chain Monte Carlo* (MCMC) | Bayesian dengan Polinomial Bernstein | hubungan antara variabel dependen dan independen karena menghasilkan MSE yang lebih kecil dan $R^2$ yang lebih besar. |

Penelitian yang akan dilakukan adalah menentukan estimasi Regresi Isotonik dengan menggunakan *Maximum Likelihood Estimator* dan pedekatan Polinomial Bernstein.

## 2.2 Landasan Teori

### 2.2.1 Estimasi Parameter

Estimasi merupakan proses yang digunakan untuk menghasilkan suatu nilai tertentu terhadap suatu parameter. Data yang digunakan untuk melakukan estimasi parameter ini merupakan suatu sampel, yang pada perkembangannya akan digunakan oleh suatu estimator untuk menghasilkan suatu nilai parameter. Estimasi parameter pada mulanya akan digunakan untuk menduga suatu populasi dari sampel. Estimasi digolongkan menjadi dua yaitu estimasi titik dan estimasi parameter. Estimasi merupakan suatu tahapan yang terpenting dalam menentukan model peluang yang tepat dari sekumpulan data (Yendra, 2015).

Adapun sifat-sifat estimasi adalah sebagai berikut:

a. Tak Bias

Menurut Wibisono (2005) dalam Mubtadiah (2011) estimator tak biasa bagi parameter tak bias bagi parameter $\theta$, jika

$$E(\hat{\theta}) = \theta$$

dan dikatakan estimator biasa bagi parameter $\theta$, jika

$$E(\hat{\theta}) \neq \theta$$

Namun penaksir bias dapat diubah menjadi penaksir tak bias jika ruas kanan dibalikan atau ditambahkan dengan konstanta tertentu.

b. Konsisten

Gujarati (2007) dalam Mubtadiah (2011) menerangkan penaksir parameter $\hat{\theta}$ dikatakan konsisten bila nilai-nilainya mendekati nilai parameter yang sebenarnya meskipun ukuran sampel semakin besar.

Suatu statistik $\hat{\theta}$ disebut penaksir yang konsisten untuk parameter $\theta$ jika dan hanya jika $\hat{\theta}$ konvergen dalam probabilitas ke parameter $\theta$ atau

$$plim\ \hat{\theta} = \theta$$

jika $\hat{\theta}_n$ adalah penaksiran untuk $\theta$ yang didasarkan pada sampel acak berukuran *n*, maka $\hat{\theta}_n$ dikatakan konsisten bagi parameter $\theta$, jika $\lim_{n\to\infty} P(\left|\hat{\theta}_n - \theta\right| < \varepsilon = 1)$.

Penentuan penaksiran konsisten ini dapat dilakukan dengan menggunakan ketidaksamaan Chebyshev's, $\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < k\sigma\right) \geq 1 - \frac{1}{k^2}$.

c. Efisien

Jika distribusi sampling dari dua statistik memiliki mean atau ekspektasi yang sama, maka statistik dengan varians yang lebih kecil disebut sebagai estimator efisien dari mean, sementara statistik yang lain disebut estimator tak efisien. Adapun nilai-nilai yang berkorespondensi dengan statistik-statistik ini masing-masing disebut sebagai estimasi efisien dan estimasi tak efisien (Mubtadiah, 2011).

### 2.2.2 Metode *Maximum Likelihood*

Misalnya $x_1, x_2, \ldots, x_n$ disebut variabel random berukuran *n* dari suatu distribusi dengan pdf $f(x;\theta)$, yang bergantung pada $\theta \in \Omega, \Omega$ disebut semesta dari parameter. Karena $x_1, x_2, \ldots, x_n$ merupakan sampel random, pdf bersama dari $x_1$, $x_2$, …, $x_n$ dapat dinyatakan sebagai:

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1;\theta)\, f(x_2;\theta) \ldots f(x_n;\theta) \qquad (2.1)$$

(Hogg, 1995 dalam Wulandari, 2006).

Pdf bersama dari $x_1, x_2, \ldots, x_n$ mengandung parameter $\theta$, sehingga Persamaan (2.1) dapat dituliskan sebagai suatu fungsi dari $\theta$, disebut $L(\theta)$.

$$L(\theta) = f(x_1, x_2, \ldots, x_n; \theta)$$

$$= f(x_1;\theta)f(x_2;\theta)\ldots f(x_n;\theta)$$

$$= \prod_{i=1}^{n} f(x_i;\theta) \tag{2.2}$$

$L(\theta)$ disebut fungsi *likelihood*.

Akan dicari $\theta$ yang memaksimumkan $L(\theta)$ . Untuk mempermudah perhitungan dalam mencari nilai $\theta$, $L(\theta)$ dapat dimodifikasi ke dalam bentuk log, karena nilai $\theta$ yang memaksimumkan log $L(\theta)$ sama dengan nilai $\theta$ yang memaksimumkan log $L(\theta)$. Sehingga persamaan (2.2) dimodifikasi menjadi:

$$\log L(\theta) = \log[\prod_{i=1}^{n} f(x_i;\theta)]$$

$$= \sum_{i=1}^{n} \log f(x_i;\theta) \tag{2.3}$$

Nilai $\theta$ yang memaksimumkan $\log L(\theta)$ , diperoleh dengan mendifferensialkan $\log L(\theta)$ terhadap $\theta$ dan menyamakannya dengan 0, dan memastikan bahwa differensial keduanya kurang dari 0.

$$\frac{d \log L(\theta)}{d\theta} = 0 \tag{2.4}$$

$$\frac{d^2 \log L(\theta)}{d\theta^2} < 0 \tag{2.5}$$

nilai $\theta = u(x_1, x_2, \ldots, x_n)$ yang memaksimumkan $\log L(\theta)$ disebut sebagai taksiran *maximum likelihood* dari $\theta$ dan dinotasikan dengan $\hat{\theta}$ (Ilmma, 2009).

### 2.2.2 Fungsi Polinomial

Sebuah fungsi *p* merupakan sebuah Polinomial, dimana *n* adalah bilangan bulat non-negatif dan nilai $a_0, a_1, a_2, \ldots, a_n$ adalah konstanta yang disebut sebagai koefisien dari polinomial.

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0 \tag{2.6}$$

Domain dari setiap polinomial adalah bilangan rill $R = (-\infty, \infty)$. Jika koefisien $a_n \neq 0$ maka derajat atau pangkat tertinggi dari polinomial adalah n (Anton, 2012 dalam Wowor, 2017).

### 2.2.3 Polinomial Bernstein

Diberikan Polinomial Bernstein dengan derajat *M* untuk fungsi kontinu *f(.)* adalah

$$y_i = w_M(x) = \sum_{k=0}^{M} f\left(\frac{k}{M}\right)\binom{M}{k} {x_i}^k (1 - x_i)^{M-k} \tag{2.7}$$

dengan

$y_i$ = variabel dependen unit ke-i

$x_i$ = variabel independen unit ke-i

$M$ = jumlah variabel

$k$ = 0, 1, ..., $M$

Jika $f\left(\frac{k}{M}\right) = \beta_k$ , maka dapat dilihat bahwa membatasi koefisien $\beta_k$ ekuivalen dengan membatasi nilai dari fungsi $f$ pada titik $\left(\frac{k}{M}\right)$ untuk $x$ dan $M$ merupakan bilangan bulat positif. Jika fungsi konstan $f\left(\frac{k}{M}\right) = 1$, maka $w_M(x) = 1$. Untuk membuktikan identitas ini dapat digunakan ekspansi binomial

$$w_M(x) = \sum_{k=0}^{M} {x_i}^k (1 - x_i)^{M-k}$$
$$= [x_i + (1 - x_i)]^M = 1^M = 1$$

Dalam penelitian Setyawan (2017) dilakukan reparameterisasi, fungsi regresi Polinomial Bernstein disederhanakan menjadi kombinasi linear dari $u_0, u_1$ dan kovariat $w_M(x_i, 0), w_M(x_i, 1)$, dimana $w_M(x_i, 0), w_M(x_i, 1)$ dapat dihitung dengan fungsi *incomplete* beta

$$y_i = w_M(x) = \sum_{k=0}^{M} \binom{M}{k} {x_i}^k (1 - x_i)^{M-k} \qquad (2.8)$$

(Setyawan, 2017).

### 2.2.3 Distribusi Normal

Distribusi Normal (Gaussian) yang dinotasikan dengan $N(\mu, \sigma^2)$ mempunyai mean $\mu$ dan variansi $\sigma^2$ mempunyai densitas

$$f(x_i;\, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right] \qquad (2.9)$$

dengan $-\infty < x_i < \infty, -\infty < \mu < \infty, \sigma > 0$

(Nugraha, 2007).

### 2.2.4 Regresi Linier Sederhana

Regresi Linier Sederhana merupakan hubungan antara satu peubah bebas *x* dan satu peubah terikat *y*, datanya dapat disajikan sebagai pasangan pengamatan $\{(x_i,\ y_i); i = 1,2, \ldots, n\}$. Regresi Linier Sederhana dapat ditulis dengan persamaan berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i; i = 1,2, \dots, n \quad (2.10)$$

dengan:

$y_i$ = variabel dependen observasi dari unit ke-i

$x_i$ = variabel independen observasi dari unit ke-i

$\beta_k$ = koefisien regresi, *k = 0,1*

$\varepsilon_i$ = *error* dari model

(Walpole, 1995).

**2.2.4.1 Estimasi Parameter Regresi Linier Sederhana**

Tujuan dari Metode Kemungkinan Maksimum dari Regresi Linear Sederhana adalah untuk menduga vektor parameter

$$\theta = [\beta_0, \beta_1, \sigma]$$

Untuk mencari penduga Kemungkinan Maksimum dari $\beta_0, \beta_1$ , dan $\sigma$ dengan menggunakan asumsi bahwa galat ( $\varepsilon_i$ ) independen dan berditribusi Normal ($\varepsilon_i \sim N(0, \sigma^2)$). Misalnya $y_1, y_2, \dots, y_n$ variabel random dependen dan distribusi Normal $N(\beta_0 + \beta_1 x_i, \sigma^2)$ untuk $i = 1,2, \dots, n$.

Fungsi probabilitas dari distribusi Normal dengan mean $\beta_0 + \beta_1 x_i$ dan variansi $\sigma^2$ adalah

$$f(Y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 x_i)^2\right]$$

Selanjunya, diperoleh fungsi *likelihood* sebagai berikut

$$L(\beta_0 + \beta_1 X_i, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 x_i)^2\right]$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^{n} \exp\left[-\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 x_i)^2\right]$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2\right]$$

Maka diperoleh fungsi *log-likelihood* sebagai berikut:

$$\log[L(\beta_0 + \beta_1 x_i, \sigma^2)] = \log\left\{\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2\right]\right\}$$

$$= n \log 1 - n \log\left(\sqrt{2\pi\sigma^2}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

$$= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

$$= -\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Penduga Kemungkinan Maksimum dari $\beta_0, \beta_1$, dan $\sigma$ dapat diperoleh dengan mencari turunan parsial $\log[L(\beta_0 + \beta_1 X_i, \sigma^2)]$ terhadap $\beta_0, \beta_1$, dan $\sigma$ dan menyamakan dengan nol, maka diperoleh

$$\frac{\partial \log L}{\partial \beta_0} = \frac{\partial}{\partial \beta_0}\left[-\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right] = 0$$

$$= \frac{2}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0 \qquad (2.11)$$

$$\frac{\partial \log L}{\partial \beta_1} = \frac{\partial}{\partial \beta_1}\left[-\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right] = 0$$

$$= -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)\,(-x_i)\left(\frac{1}{2\sigma^2}\right) = 0$$

$$= \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i(y_i - \beta_0 - \beta_1 x_i) = 0 \qquad (2.12)$$

$$\frac{\partial \log L}{\partial \sigma} = \frac{\partial}{\partial \sigma^2}\left[-\frac{n}{2}\log 2\pi - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right] = 0$$

$$= -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = 0 \qquad (2.13)$$

Berdasarkan Persamaan (2.11) diperoleh

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1\sum_{i=1}^{n} x_i \qquad (2.14)$$

Berdasarkan Persamaan (2.12) diperoleh

$$\frac{1}{\sigma^2}\sum_{i=1}^{n} x_i(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\sum_{i=1}^{n} x_i\, y_i = \beta_0\sum_{i=1}^{n} x_i + \beta_1\sum_{i=1}^{n} {x_i}^2 \qquad (2.15)$$

Dengan menggunakan metode eliminasi pada Persamaan (2.14) dan Persamaan (2.15), maka diperoleh

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} {x_i}^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (2.16)$$

$$\beta_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad (2.17)$$

Berdasarkan Persamaan (2.13) diperoleh

$$-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1x_i)^2=0$$

$$-\frac{n\sigma^2}{\sigma^3}+\frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1x_i)^2=0$$

$$\frac{n\sigma^2}{\sigma^3}=\frac{1}{\sigma^3}\sum_{i=1}^{n}(y_i-\beta_0-\beta_1x_i)^2$$

$$\sigma^2=\frac{\sum_{i=1}^{n}(y_i-\beta_0-\beta_1x_i)^2}{n} \tag{2.18}$$

Permasamaan (2.16) dan Persamaan (2.17) menunjukkan bahwa penduga Kemungkinan Maksimum dari Regresi Linear Sederhana menghasilkan penduga (*estimator*) yang sama dengan penduga yang dihasilkan dengan Metode Kuadrat Terkecil. Penduga Kemungkinan Maksimum dari $\sigma^2$ yang ditulis dalam Persamaan (2.18) adalah rata-rata kuadrat galat sampel (Hede, 2016).

### 2.2.5 Regresi Isotonik dengan Pendekatan Polinomial Bernstein

Regresi Isotonik adalah suatu model regresi yang digunakan untuk fungsi regresi yang berbentuk monoton naik (Fajaroh, 2012).

**Defenisi 2.1** *Diberikan x=* $\{x_1, x_2, \dots, x_n\}$*, dimana* $x_1 < x_2 < \cdots < x_n$ *menunjukkan himpunan berhingga dari n buah nilai terobservasi dari variabel independen dan* $y_1, y_2, \dots, y_n$ *menunjukkan nilai terobservasi yang berkorespondensi dari variabel dependen. Fungsi f pada x disebut isotonik jika* $x_i \le x_j$*, i≠j, berakibat* $f(x_i) \le f(x_j), i = 1, \dots, n$ (Barlow, 1972 dalam Setyawan, 2017).

Regresi Isotonik dengan Pendekatan Polinomial Bernstein dapat ditulis dengan persamaan berikut:

$$\hat{y}_i = w_i{}' \, u_j$$

$$= w_M(x_i, 0)u_0 + w_M(x_i, 1)\, u_1 \tag{2.19}$$

dengan:

$\hat{y}_i$ = nilai yang diprediksikan

$w_M(x_i, k)$ = fungsi Polinomial Bernstein, $k = 0, 1,\ M = 1$

$u_j$ = parameter regresi, $j = 0, 1$

#### 2.2.5.1 *Maximum Likelihood Estimator* untuk Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein dalam Kasus Satu Variabel Independen

Dalam Regresi Isotonik dapat digunakan model dengan Polinomial Bernstein pada Persamaan (2.8). Salah satu sifat Polinomial Bernstein adalah bahwa semua derivatifnya mempunyai sifat konvergensi yang sama. Tenbusch (1997) telah membuktikan banyak sifat baik fungsi regresi nonparametrik dengan Polinomial Bernstein. Sifat-sifat baiknya meliputi *pointwise consistency*, yaitu $E(f_n(x) - f(x))^2 = 0$ , kemudian *asymptotic normality*, yaitu semakin besar sampel maka sampel akan mendekati distribusi normal, dan *uniform consistency*, yaitu $\underset{x}{\text{Sup}}\ MSE\big(f_n(x)\big)$ atau batas atas terkecil dari residualnya cenderung mendekati 0 (Tenbusch, 1997 dalam Setyawan, 2017).

Sebelum mencari penduga dengan menggunakan MLE, terlebih dahulu akan dicari fungsi densitas gabungan (*likelihood*) dari sampel random yang dinotasikan $\pi(y|u, \sigma^2)$. Diberikan fungsi densitas dan variabel random $y$ adalah berdistribusi normal dengan rata-rata $w'u$ dan variansi $\sigma^2$ yaitu

$$\pi(y|u, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(y_i - w_i{}'u_j\right)^2}{2\sigma^2}\right).$$

Fungsi densitas gabungan untuk vektor sampel $y = (y_1, y_2, \dots, y_n)'$ dengan $n$ buah observasi adalah

$$\pi(y|u, \sigma^2) = \prod_{i=1}^{n} \pi(y_i|u, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{\left(y_i - w_i' u_j\right)^2}{2\sigma^2}\right) \qquad (2.\ 20)$$

dengan $w = (w_M(x_i, 0), w_M(x_i, 1))'$.

Pada regresi isotonik pada kasus satu variabel independen diketahui matrik $w = \begin{bmatrix} w_M(x_i, 0) \\ w_M(x_i, 1) \end{bmatrix}$ dan matrik u = $\begin{bmatrix} u_0 \\ u_1 \end{bmatrix}$ . Maka didapatkan matrik $w' = [w_M(x_i, 0) \quad w_M(x_i, 1)]$. Dengan nilai *M = 1*. Selanjutnya, dari Persamaan (2.20) didapatkan persamaan sebagai berikut:

$$\pi(y|u_0, u_1, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y_i-(w_M(x_i,0)u_0+w_M(x_i,1)u_1)^2}{2\sigma^2}\right)$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\sum_{i=1}^{n}(y_i-(w_M(x_i,0)u_0+w_M(x_i,1)u_1))^2}{2\sigma^2}\right) \quad (2.21)$$

Fungsi log-*likelihood* dari persamaan diatas adalah

$$\log[\pi(y|u_0, u_1, \sigma^2)] = \log\left[\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{\sum_{i=1}^{n}(y_i-(w_M(x_i,0)u_0+w_M(x_i,1)u_1))^2}{2\sigma^2}\right)\right]$$

$$= -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - (w_M(x_i,0)u_0 + w_M(x_i,1)u_1))^2 \quad (2.\ 22)$$

Penduga kemungkinan maksimum dari $u_0, u_1$ dan $\sigma^2$ adalah penduga yang memaksimumkan $\log[\pi(y|u, \sigma^2)]$, dengan mencari nilai turunan parsial terhadap $u_0, u_1$ dan $\sigma^2$ dan menyamakan dengan nol, maka diperoleh:

$$\frac{\partial \log[\pi(y|u_0, u_1, \sigma^2)]}{\partial u_0} = 0$$

$$= \frac{\partial}{\partial u_0}\left[-\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - (w_M(x_i,0)u_0 + w_M(x_i,1)u_1))^2\right] = 0$$

$$= 0 - \frac{1}{2\sigma^2}(-2)\{(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0)) - (\sum_{i=1}^{n} w_M(x_i,0))^2 u_0 - \left(\sum_{i=1}^{n} w_M(x_i,0)\sum_{i=1}^{n} w_M(x_i,1)\right)u_1\} = 0$$

$$= (\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0)) - (\sum_{i=1}^{n} w_M(x_i,0))^2 u_0 - \left(\sum_{i=1}^{n} w_M(x_i,0)\sum_{i=1}^{n} w_M(x_i,1)\right)u_1 = 0 \quad (2.23)$$

$$\frac{\partial \log[\pi(y|u_0, u_1, \sigma^2)]}{\partial u_1} = 0$$

$$= 0 - \frac{1}{2\sigma^2}(2)\{\sum_{i=1}^{n}\left(y_i - w_M(x_i,0)u_0 - w_M(x_i,1)u_1\right)(-\sum_{i=1}^{n} w_M(x_i,1))\} = 0$$

$$= \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1) - (\sum_{i=1}^{n} w_M(x_i,1))^2 u_1 - \left(\sum_{i=1}^{n} w_M(x_i,0)\sum_{i=1}^{n} w_M(x_i,1)\right)u_0 = 0 \quad (2.24)$$

Eliminasi Persamaan (2.23) dengan Persamaan (2.24)

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - u_0 \left(\sum_{i=1}^{n} w_M(x_i,0)\right)^2 - u_1 \left(\sum_{i=1}^{n} w_M\,(x_i,0) \sum_{i=1}^{n} w_M\,(x_i,1)\right) = 0$$

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1) - u_0 \left(\sum_{i=1}^{n} w_M\,(x_i,0) \sum_{i=1}^{n} w_M\,(x_i,1)\right) - u_1 \left(\sum_{i=1}^{n} w_M(x_i,1)\right)^2 = 0$$

misalkan:

A = $\left(\sum_{i=1}^{n} w_M\,(x_i,0) \sum_{i=1}^{n} w_M\,(x_i,1)\right)$

B = $(\sum_{i=1}^{n} w_M(x_i,0))^2$

C = $(\sum_{i=1}^{n} w_M(x_i,1))^2$

maka menjadi persamaan eliminasi sebagai berikut:

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - u_0 B - u_1 A = 0 \qquad (2.25)$$

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1) - u_0 A - u_1 C = 0 \qquad (2.26)$$

kalikan Persamaan (2.25) dengan A dan Persamaan (2.26) dengan B sebagai berikut:

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - u_0 B - u_1 A = 0 \mid \times A$$

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1) - u_0 A - u_1 C = 0 \mid \times B$$

didapatkan persamaan berikut

$$A \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - u_0 AB - u_1 A^2 = 0$$

$$B \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1) - u_0 AB - u_1 BC = 0 \quad -$$

hasil eliminasi

$$A \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - B \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1) + u_1 BC - u_1 A^2 = 0 \qquad (2.27)$$

misalkan,

D = $A \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - B \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,1)$

maka

$$D + u_1(BC - A^2) = 0$$

$$u_1(BC - A^2) = -D$$

$$u_1 = \frac{-D}{(BC - A^2)}$$

$$u_1 = \frac{D}{(A^2 - BC)} \tag{2.28}$$

berdasarkan nilai $u_1$ pada Persamaan (2.28), substitusikan ke Persamaan (2.25)

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i, 0) - u_0 B - \left(\frac{D}{(A^2 - BC)}\right) A = 0$$

$$\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i, 0) - u_0 B - \left(\frac{AD}{(A^2 - BC)}\right) = 0$$

$$u_0 B = \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i, 0) - \left(\frac{AD}{(A^2 - BC)}\right)$$

$$u_0 = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i,0) - \left(\frac{AD}{(A^2-BC)}\right)}{B} \tag{2.29}$$

$$\frac{\partial \log[\pi\left(y \middle| u_0, u_1, \sigma^2\right)]}{\partial \sigma^2} = 0$$

$$= \quad -\frac{n}{2}\left(\frac{1}{2\pi\sigma^2}\right) 2\pi - \frac{1}{2}\,(-2)\sigma^{-4} \sum_{i=1}^{n} (y_i - (\,w_M(x_i, 0)u_0 + w_M(x_i, 1)u_1)^2) = 0$$

$$= \quad -\left(\frac{n}{2\pi\sigma^2}\right) + \frac{1}{\sigma^4} \sum_{i=1}^{n} (y_i - (\,w_M(x_i, 0)u_0 + w_M(x_i, 1)u_1))^2 = 0$$

$$= \frac{-n\sigma^2 + 2\sum_{i=1}^{n} (y_i - (w_M(x_i,0)u_0 + w_M(x_i,1)u_1))^2}{2\sigma^4} = 0$$

$$= 2\sum_{i=1}^{n} (y_i - (\,w_M(x_i, 0)u_0 + w_M(x_i, 1)u_1)^2) = 0$$

$$-n\sigma^2 = 2\sum_{i=1}^{n} (y_i - w_M(x_i, 0)u_0 - w_M(x_i, 1)u_1)^2$$

$$\sigma^2 = \frac{2\sum_{i=1}^{n} (y_i - w_M(x_i, 0)u_0 - w_M(x_i, 1)u_1)^2}{n}$$

misalkan

E= $w_M(x_i, 0) \sum_{i=1}^{n} y_i \sum_{i=1}^{n} w_M(x_i, 0)$

F = $w_M(x_i, 0) \left(\frac{AD}{(A^2 - BC)}\right)$

G = $w_M(x_i, 1)\left(\frac{D}{(A^2 - BC)}\right)$

maka

$$\sigma^2 = \frac{2\sum_{i=1}^{n}\left(y_i - \left(\frac{E-F}{B}\right) + G\right)^2}{n} \quad (2.30)$$

### 2.2.5 Kriteria Pemilihan Model Terbaik

Kriteria yang digunakan pada penelitian ini yaitu *Mean Squared Error* (MSE) dan *R-Square* ($R^2$). Model terbaik adalah model yang memiliki nilai MSE terkecil dan nilai $R^2$ terbesar.

1) *Mean Squared Error* (MSE)

*Mean of Squares Error* (MSE) merupakan salah satu pengukuran kesalahan yang banyak digunakan. Nilai MSE dihitung dengan mengkuadratkan selisih antara nilai prediksi dengan nilai sebenarnya. Umumnya, semakin kecil nilai MSE maka semakin tinggi tingkat keakuratan nilai suatu prediksi. MSE dapat dihitung dengan menggunakan persamaan berikut (Ghozali, 2009).

$$MSE = \frac{\sum_{i=0}^{n}(y_i - \hat{y}_i)^2}{n} \; ; \; i = 1, 2, \ldots, n \quad (2.31)$$

2) *R-square* ($R^2$)

*R-square* merupakan proporsi variasi variabel dependen yang dapat diterangkan oleh variabel independen dalam model. *R-square* dijadikan sebagai ukuran seberapa baik hasil estimasi kurva regresi mendekati data aslinya sehingga semakin besar nilai *R-square* maka semakin baik pula model tersebut. Nilai *R-square* pada umumnya terletak di antara nol dan satu. Jika sama dengan nol, maka tidak ada variasi *y* diterangkan oleh perubahan-perubahan variabel-variabel penjelas. Jika sama dengan nol, maka tidak ada variasi *y* yang diterangkan oleh perubahan-perubahan variabel-variabel penjelas. Untuk mencari *R-Square* dengan rumus sebagai berikut :

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \; ; \; i = 1, 2, \ldots, n \quad (2.32)$$

(Setyawan, 2017 dan Sarwoko, 2007).

# BAB III

# PEMBAHASAN

## 3.1 Studi Kasus

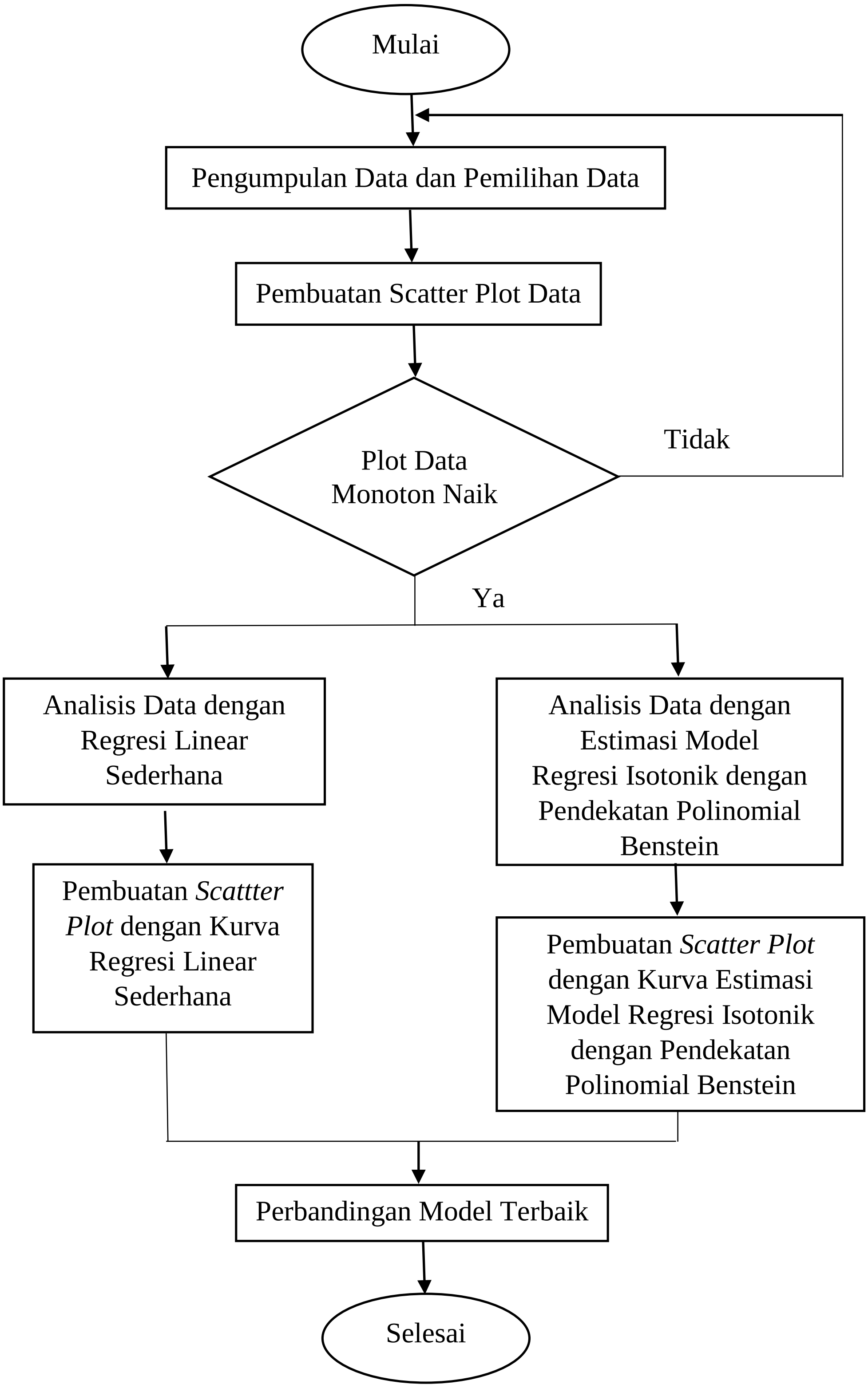
Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari penelitian Adawiyah (2011) menggunakan makanan jadi, minuman, rokok, dan tembakau sebagai variabel dependen. Sedangkan, pendidikan, rekreasi, dan olahraga sebagai variabel independen. Data yang diperoleh yaitu Mei 2006 sampai Mei 2008. Data tersebut digunakan karena makanan jadi, minuman, rokok, dan tembakau berpengaruh positif terhadap pendidikan, rekreasi, dan olahraga. Artinya semakin besar nilai makanan jadi, minuman, rokok, dan tembakau maka akan semakin besar pendidikan, rekreasi, dan olahraga.

## 3.2 Perangkat Lunak yang Digunakan

Dalam studi kasus ini, peneliti menganalisis data tersebut digunakan *software* R versi 3.3.3 dengan *package* stats4.

## 3.3 Metodologi Penelitian

Pengumpulan dan pemilihan data akan dilakukan pertama-tama dalam penelitian ini. Untuk melihat pola data diperlukan pembuatan *scatter plot*. Jika Plot data bukan monoton naik maka penulis akan mengumpulkan dan pemilihan data kembali tetapi jika pola data monoton naik dilanjutkan dengan analisis data dengan Regresi Linear Sederhana dan kemudian pembuatan kurva dari model Regresi Linear Sederhana. Kemudian, analisis data dengan estimasi model Regresi Isotonik dengan pendekatan Polinomial Benstein dilanjutkan dengan pembuatan kurva dari model Regresi Isotonik dengan pendekatan Polinomial Bernstein. Setelah itu, pemilihan model terbaik dari perbandingan metode keduanya.




**Gambar 3.1** Tahapan Penelitian

## 3.4 Hasil Analisis

### 3.4.1 Algoritma Regresi Linear Sederhana

*Input*:

Y = Data variabel dependen

X = Data variabel independen

*Output*: Model dari Estimasi Parameter Regresi Linear Sederhana dan plot data dengan Kurva Regresi Linear Sederhana

Metode:

Langkah 1: Penginputan Data

Langkah 2 : Pembentukan *Scatter Plot*

Langkah 3 : Analisis Regresi Linear Sederhana dan Pembentukan Model Regresi Linear Sederhana, Serta Penentuan *R-Square*

Langkah 4 : Perhitungan Prediksi dari model Regresi Linear Sederhana

Langkah 5 : Perhitungan MSE

Langkah 6 : Pembentukan *Scatter Plot* dengan Kurva Regresi Linear Sederhana

Tahapan perhitungan dan komputasi (menggunakan *software* R) algoritma Regresi Linear Sederhana pada Tabel 3.1.

**Tabel 3.1** Algoritma Perhitungan dan Komputasi Regresi Linear Sedehana

| **Alur Analisis Regresi Linear Sederhana** | **Program R Analisis Regresi Linear Sederhana** |
|---|---|
| Mulai<br>Input data | ##memuat data yang digunakan##<br>Y= scan()<br>X= scan() |
| Pembentukan *Scatter Plot* | ##Plot data##<br>#variabel Y yaitu makanan jadi, minuman, rokok, dan tembakau#<br>#variabel X yaitu pendidikan, rekreasi, dan olahraga#<br>plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan jadi, minuman, rokok, dan tembakau", main="Scatter Plot pendidikan, rekreasi, dan olahraga dengan makanan jadi, minuman, rokok, dan tembakau") |
| Pembentukan Model dan penentuan *R-Square*<br>Perhitungan Prediksi | ##Regresi Liniear Sederhana##<br>reg.lin.sederhana=lm(Y~X)<br>summary(reg.lin.sederhana)<br>##Prediksi Regresi Linear Sederhana##<br>Ytopilinear=32.47091+0.73886*(X) |
| Menghitungan MSE | ##Sum Square Error Regresi Linear Sederhana##<br>SSE_RLS=sum((Y-ytopilinear)^2)<br>SSE_RLS<br><br>##Mean Square Error Regresi Linear Sederhana##<br>MSE=SSE_RLS/n<br>MSE |
| Pembentukan *Scatter Plot* dengan Kurva Regresi Linear Sederhana<br>Selesai | ##Plot Data Regresi Linear##<br>plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan jadi, minuman, rokok, dan tembakau", main="Scatter Plot pendidikan, rekreasi, dan olahraga dengan makanan jadi, minuman, rokok, dan tembakau")<br>lines(sort(X),sort(ytopilinear), col="red") |

| Alur Analisis Regresi Linear Sederhana | Program R Analisis Regresi Linear Sederhana |
|---|---|
| | `legend("topleft", c("Linear Fit"), bty="n", lwd=2, col=c("red"))` |

### 3.4.2 Algoritma *Maximum Likelihood Estimator* untuk Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein

*Input*:

Y = Data variabel dependen

X = Data variabel independen

*Output*: Model dari Estimasi Regresi Isotonik dengan pendekatan Polinomial Bernstein dan Plot data dengan Kurva Estimasi Regresi Isotonik dengan pendekatan Polinomial Bernstein

Metode:

Langkah 1 : Pemanggilan *Package* yang Digunakan

Langkah 2 : Penginputan Data

Langkah 3 : Pembentukan *Scatter Plot*

Langkah 4 : Pembentukan Model MLE untuk Regresi Isotonik dengan Polinomial Bernstein

Langkah 5 : Perhitungan Prediksi Estimasi Regresi Isotonik dengan Polinomial Bernstein

Langkah 6 : Perhitungan MSE dan *R-square*

Langkah 7 : Pembentukan *Scatter Plot* dengan Kurva  Estimasi Regresi Isotonik dengan Polinomial Bernstein

Tahapan perhitungan dan komputasi (menggunakan *software* R) algoritma *Maximum Likelihood Estimator* untuk model Regresi Isotonik dengan pendekatan Polinomial Bernstein pada Tabel 3.2.

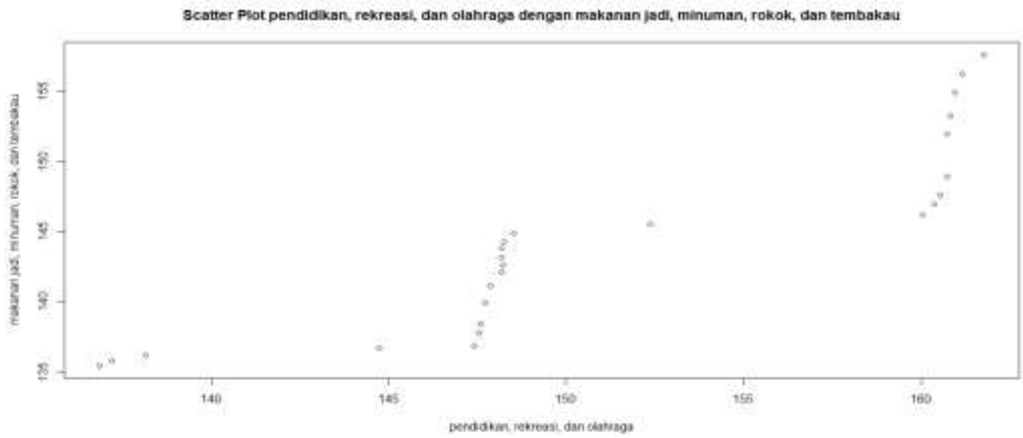
**Tabel 3.2** Algoritma Perhitungan dan Komputasi *Maximum Likelihood Estimator* untuk model Regresi Isotonik dengan pendekatan Polinomial Bernstein

| **Alur Analisis Regresi Isotonik** | **Program R Analisis Regresi Isotonik** |
|---|---|
| Mulai<br>Pemanggilan *Package* yang digunakan<br>*Input* data | `##memuat package yang diperlukan##`<br>`library(stats4)`<br><br>`##memuat data yang digunakan##`<br>`Y= scan()`<br>`X= scan()` |
| Pembentukan *Scatter Plot* | `##Plot data##`<br>`#variabel Y yaitu makanan jadi, minuman, rokok, dan tembakau#`<br>`#variabel X yaitu pendidikan, rekreasi, dan olahraga#`<br>`plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan jadi, minuman, rokok, dan tembakau", main="Scatter Plot pendidikan, rekreasi, dan olahraga dengan makanan jadi, minuman, rokok, dan tembakau")` |
| Pembentukan Model dan Perhitungan Estimasi | `L=function(U0, U1, sigma)`<br>`{`<br>`-1*(-n/2*log(2*pi*sigma^2)-(1/(2*sigma^2))*(sum((Y-W0*U0-W1*U1)^2)))`<br>`}`<br>`estimasi=mle(minuslog=L, start=list(U0=0.005,U1=0.007, sigma=0.006))`<br>`summary(estimasi)` |
| Menghitung Prediksi | `ytopi=array(dim=c(n))`<br>`for(i in 1:n)`<br>`{`<br>`ytopi[i]=(W0[i]*U0+W1[i]*U1)`<br>`}`<br>`Ytopi` |
| A | `##Menghitung Sum Square Error (SSE) dan Mean Square Error (MSE) Regresi Isotonik##`<br>`SSE_isotonik=sum((Y-ytopi)^2)`<br>`SSE_isotonik` |

| **Alur Analisis Regresi Isotonik** | **Program R Analisis Regresi Isotonik** |
|---|---|
| A<br>↓<br>Menghitung MSE dan *R-Square* | `MSE_isotonik=SSE_isotonik/n`<br>`MSE_isotonik`<br><br>`##Menghitung R-square Regresi isotonik##`<br>`mean=mean(Y)`<br>`mean`<br>`SST=sum((Y-mean)^2)`<br>`SST`<br>`Rsq=1-(SSE_isotonik/SST)`<br>`Rsq` |
| ↓<br>Pembentukan *Scatter Plot* dengan Estimasi Regresi Isotonik dengan Polinomial Benstein<br>↓<br>Selesai | `##Plot Data dengan kurva regresi isotonik##`<br>`plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan jadi, minuman, rokok, dan tembakau", main="Scatter Plot pendidikan, rekreasi, dan olahraga dengan makanan jadi, minuman, rokok, dan tembakau")`<br>`lines(X,ytopi, col= "blue")`<br>`legend("topleft", c("Isotonic Fit"), bty="n", lwd=2, col=c("blue"))` |

## 3.5 Analisis Deskriptif

Berdasarkan pola data pada Gambar 3.2, dapat dilihat hubungan antara pendidikan, rekreasi dan olahraga dengan makanan jadi, minuman, rorkok dan tembakau terus naik. Sehingga dapat diartikan bahwa pendidikan, rekreasi dan olahraga memiliki pengaruh positif terhadap makanan jadi, minuman, rokok dan temabakau artinya makanan jadi, minuman, rokok dan tembakau merupakan fungsi *non-descreasing* karena semakin besar nilai pendidikan, rekreasi dan olahraga menyebabkan pendidikan, rekreasi dan olahraga semakin naik dengan korelasi 0.897 artinya hubungan makanan jadi, minuman, rokok dan tembakau dengan pendidikan, rekreasi dan olahraga cukup kuat cukup kuat. Sehingga *Maximum Likelihood Estimator* untuk model Regresi Isotonik dengan pendekatan Polinomial Bernstein dapat digunakan untuk menganalisis data tersebut.




**Gambar 3.2** *Scatter Plot* hubungan antara pendidikan, rekreasi dan olahraga dengan makanan jadi, minuman, rokok dan tembakau

### 3.6 Analisis Regresi Linear Sederhana

Regresi Linear Sederhana dilakukan terhadap data hubungan antara makanan jadi, minuman, rokok dan tembakau dengan pendidikan, rekreasi dan olahraga diperoleh nilai $\beta_0$ sebesar 32.471 dan nilai $\beta_1$ sebesar 0.739 maka model regresi diperoleh

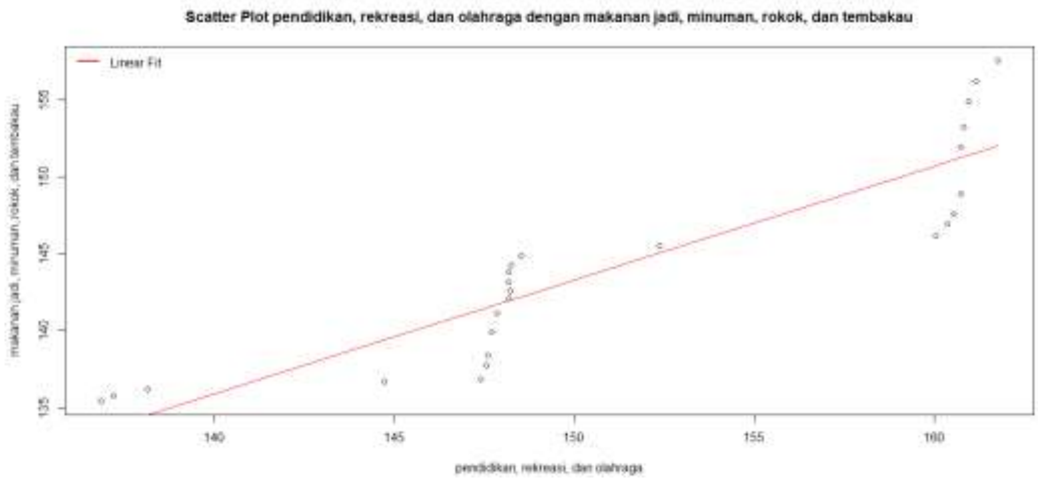
$$\hat{y}_i = 32.471 + 0.793\ x_i \quad (4.1)$$

dengan:

$\hat{y}_i$ = makanan jadi, minuman, rokok dan tembakau unit ke-i

$x_i$ = pendidikan, rekreasi dan olahraga unit ke-i

Berdasarkan model diatas artinya jika pendidikan, rekreasi dan olahraga mendekati nol, maka nilai makanan jadi, minuman, rokok dan tembakau akan mendekati 32.471. Sedangkan jika pendidikan, rekreasi dan olahraga mengalami peningkatan satu satuan harga maka makanan jadi, minuman, rokok dan tembakau akan mengalami peningkatan sebesar 0.793. Berdasarkan model didapatkan nilai MSE sebesar 8.246 dan dapat dihasilkan kurva garis lurus sebagai berikut.




**Gambar 3.3** *Scatter Plot* dengan kurva Regresi Linier Sederhana

Berdasarkan Gambar 3.3 kurva Regresi Linear Sederhana akan konstan dengan nilai *R-square* 0.7969 menjelaskan bahwa sebesar 79.69% proporsi pengaruh pendidikan, rekreasi dan olahraga mempengaruhi makanan jadi, minuman, rokok dan tembakau.

Berdasarkan *scatter plot* pada Gambar 3.3 monoton naik maka penulis juga menggunakan metode estimasi model Regresi Isotonik dengan pendekatan Polinomial Bernstein.

### 3.7 Analisis Regresi Isotonik dengan Pendekatan Polinomial Bernstein

Hasil Regresi Isotonik dengan pendekatan Polinomial Bernstein diperoleh estimasi parameter untuk $u_0$ adalah 47.9380 dan estimasi parameter untuk $u_1$ adalah 1.2895 maka model regresi yang diperoleh adalah sebagai berikut:

$$\hat{y}_i = w_i{}' \, u_j$$
$$= w_M(x_i, 0)u_0 + w_M(x_i, 1)\, u_1$$
$$= 31.816\, w_M(x_i, 0) + 0.743\, w_M(x_i, 1) \quad (4.2)$$

misalnya:

$w_M(x_i, k) = 1 \longrightarrow \hat{y}_1 = 31.816\ (1) + 0.743\ (1) = 32.559$

$w_M(x_i, k) = 2 \longrightarrow \hat{y}_2 = 31.816\ (2) + 0.743\ (2) = 65.118$

$w_M(x_i, k) = 3 \longrightarrow \hat{y}_3 = 31.816\ (3) + 0.743\ (3) = 97.677$

$w_M(x_i, k) = 4 \longrightarrow \hat{y}_4 = 31.816\ (4) + 0.743\ (4) = 130.236$

maka

$\hat{y}_1$ = 32.559 Selisih kenaikan $= \frac{(65.118-32.559)}{32.559} \times 100\% = 100\%$

$\hat{y}_2$ = 65.118 Selisih kenaikan $= \frac{(97.677-65.118)}{65.118} \times 100\% = 50\%$

$\hat{y}_3$ = 97.677

$\hat{y}_4$ = 130.236 Selisih kenaikan $= \frac{(130.236-97.677)}{97.677} \times 100\% = 33.333\%$

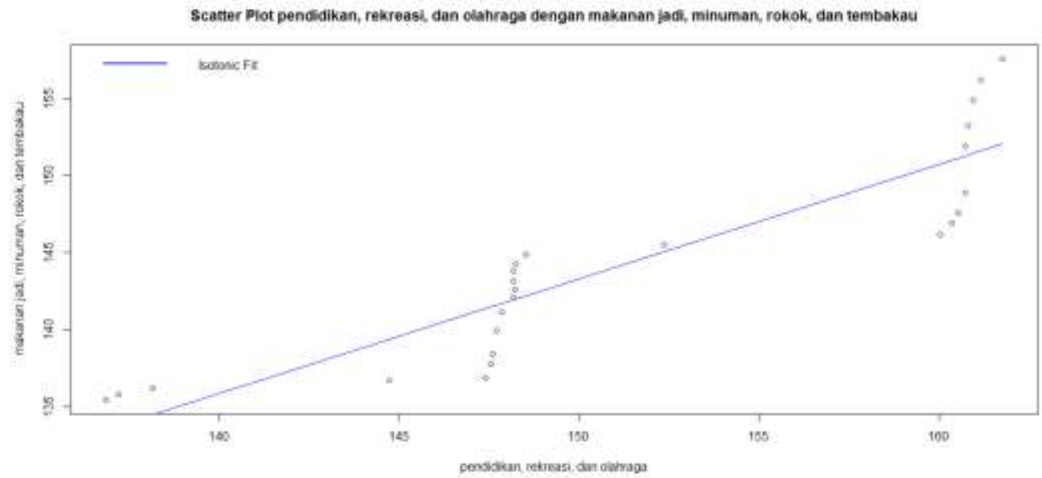
dengan:

$\hat{y}_i$ = makanan jadi, minuman, rokok dan tembakau unit ke-i

$w_M(x_i, k)$ = pendidikan, rekreasi dan olahraga unit ke-i, *k = 0,1*

Berdasarkan uraian di atas artinya kenaikan satu satuan nilai pendidikan, rekreasi dan olahraga akan kenaikan nilai makanan jadi, minuman, rokok dan tembakau yang berbeda-beda karena akan menyesuaikan bobot dari nilai pendidikan, rekreasi dan olahraga dengan MSE sebesar 8.248. Berbeda dengan Regresi Linear Sederhana kenaikan nilai prediksi akan selalu konstan.

Pada Regresi Isotonik dihasilkan kurva membentuk garis lurus seperti Regresu Linear Sederhana sebagai berikut

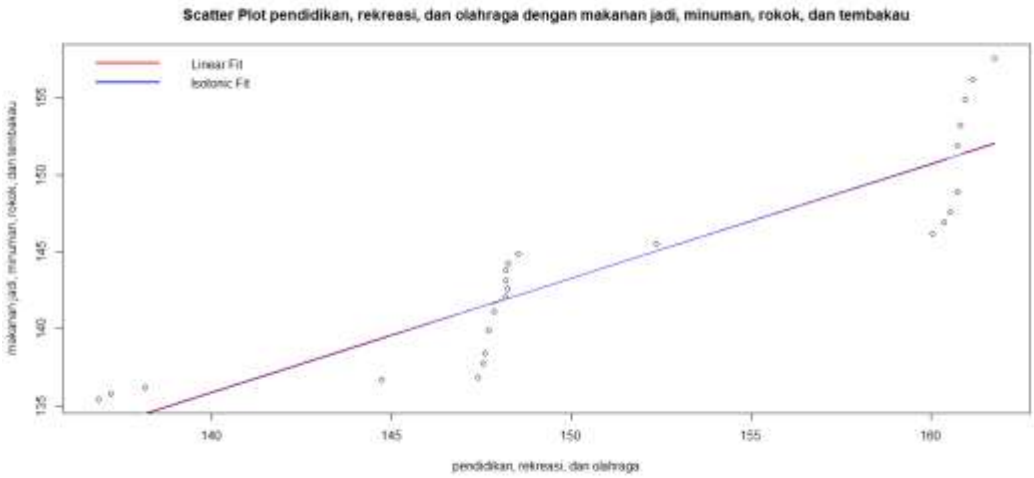



**Gambar 3.4** *Scatter Plot* dengan kurva Regresi Isotonik

berdasarkan Gambar 3.4 didapatkan nilai *R-square* 0.8053 menjelaskan bahwa sebesar 80.53% proporsi pengaruh pendidikan, rekreasi dan olahraga mempengaruhi makanan jadi, minuman, rokok dan tembakau.

### 3.6 Perbandingan Regresi Linear dengan Estimasi Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein

Perbandingan kebaikan model dapat dilihat dari dari plot seperti berikut:




**Gambar 3.5** Kurva Perbandingan Regresi Linear Sederhana dan Regresi Isotonik

berdasarkan Gambar 3.5 diketahui bahwa Regresi Isotonik dan Regresi Linear Sederhana mempunyai kurva yang sama yang membentuk garis lurus.

**Tabel 3.3** Tabel Perbandingan Regresi Linear Sederhana dengan Regresi Isotonik

| | **Regresi Linear Sederhana** | **Estimasi Regresi Isotonik dengan pendekatan Polinomal Bernstein** |
|---|---|---|
| *R-Square* | 79.69% | 80.52% |
| MSE | 8.246 | 8.247 |

Berdasarkan Tabel 3.3 estimasi model Regresi Isotonik dengan pendekatan Polinomial Bernstein dan Regresi Linear Sederhana mempunyai selisih yang sangat kecil. Maka dapat disimpulkan bahwa estimasi model Isotonik dengan pendekatan Polinomial Bernstein dan Regresi Linear Sederhana mempunyai kebaikan yang sama. Oleh karena itu, estimasi model Regresi Isotonik dengan Polinomial Bernstein baik digunakan sebagai alternatif penentuan model pada regresi dengan plot data monoton naik.

# BAB IV

# PENUTUP

## 4.1 Kesimpulan

Berdasarkan hasil pembahasan, diperoleh kesimpulan bahwa:

1. Selain menggunakan metode Bayes sebagaimana pada penelitian Bagus Setyawan (2017), estimasi model Regresi Isotonik juga dapat dilakukan menggunakan *Maximum Likelihood Estimator*.
2. Dari studi kasus pada Bab III, pada Estimasi Regresi Isotonik dengan pendekatan Polinomial Bernstein kenaikan satu satuan nilai pendidikan, rekreasi dan olahraga akan memiliki kenaikan nilai makanan jadi, minuman, rokok dan tembakau yang berbeda-beda karena akan menyesuaikan bobot dari pendidikan rekreasi dan olahraga.
3. Estimasi model Regresi Isotonik dengan pendekatan Polinomial Bernstein dan Regresi Linear Sederhana mempunyai selisih yang sangat kecil. Maka dapat disimpulkan bahwa estimasi model Isotonik dengan pendekatan Polinomial Bernstein dan Regresi Linear Sederhana mempunyai kebaikan yang sama. Oleh karena itu, estimasi model Regresi Isotonik dengan Polinomial Bernstein baik digunakan sebagai alternatif penentuan model pada regresi dengan plot data monoton naik.

## 4.2 Saran

Saran yang diajukan untuk penelitian selanjutnya yaitu:

1. Regresi Isotonik dengan pendekatan Polinomial Bernstein dalam skripsi ini menggunakan *Maximum Likelihood Estimator* untuk mengestimasi parameter. Untuk penelitian selanjutnya dapat menggunakan metode estimasi parameter lainnya, seperti *Ordinary Least Square* (OLS).
2. *Maximum Likelihood Estimator* untuk model Regresi Isotonik dengan pendekatan Polinomial Bernstein yang dibahas pada Tugas Akhir ini hanya

menggunakan satu vari’abel independen saja. Penelitian selanjutnya dapat menggunakan beberapa variabel independen.

3. Data yang digunakan dalam analisis dapat diambil dari berbagai bidang seperti kesehatan, keuangan, sosial dan sebagainya. Variabel dependen yang digunakan dapat bertipe diskrit maupun kontinu dan dapat bertipe *time series* maupun *cross-sectional*.

## DAFTAR PUSTAKA


Adawiyah, R. 2011. *Model Distribusi Lag dan Autoregressive Dengan Pendekatan Koyck*. Skripsi. Jurusan Matematika FMIPA Universitas Pendidikan Indonesia.

Curtis, S. M. dan Ghosh, S. K. 2010. *A variable selection approach to monotonicm regression with Bernstein Polynomials,* Journal of Applied Statistics, Volume 38, Issue 5.

Farizal., Amar R dan Hadi A. 2014. *Model Peramalan Konsumsi Bahan Bakar Jenis Premium di Indonesia dengan Regresi Linear Berganda*. Jurnal Ilmu teknik Industri, Volume 13, Nomor 2, Desmber 2014, Halaman 166-176, ISSN 1412-6869.

Fajaroh, L. A. 2012. *Regresi Isotonik untuk Menentukan Nilai Ambang (Threshold Value)*. Skripsi. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sebelas Maret Surakarta.

Ghozali, I. 2009. Aplikasi Analisis Multivariate Dengan Program SPSS (Edisi keepat). Semarang: Universitas Dipenegoro.

Hede, R. P. A. 2016. *Perbandingan Metode Kuadrat Terkecil Dan Metode Kemungkinan Maksimum Dalam Pendugaan Parameter Distribusi Weibull Dengan Dua Parameter*. Skripsi. Program Studi Matematika Jurusan Matematika Fakultas Sains dan Teknologi Universitas Sanata Dharma Yogyakarta.

Ilmma, A. 2009. *Taksiran Maksimum Likelihood Pada Model Persamaan Struktural Nonlinear*. Skripsi. Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia Depok.

Misbahussurur, A. 2009. *Estimasi Parameter Distribusi Gamma dengan Metode Maksimum Likelihood*. Skripsi. Jurusan Matematika Fakultas Sains dan teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Mubtadiah, L. 2011. Estimasi Parameter Model Regresi Pasial *Error* Dengan Metode *Maximum Likelihood Estimator*. Skripsi. Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri (UIN) Maulana Malik Ibrahim, Malang.

Neelon, B dan David B. D. 2003. *Bayesian Isotonic Regression and Trend Analysis.* Biometrics, 60, 398-406.

Nugraha, J. 2007. *Menghitung Nilai Probabilitas Pada Distribusi Normal Multivariate*. Jurnal Pythagoras, Volume 3, Nomor 2, Desember 2007, Halaman 1-14.

Nuraini, Y. S. 2012. *Estimasi Fungsi Penghalus Pada Regresi Isotonik Aditif Dengan Metode Kuadrat Terkecil*. Skripsi. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sebelas Maret Surakarta.

Sarwoko. 2007. *Statistika Inferensi untuk Ekonomi dan Bisnis*. Yogyakarta: Penerbit ANDI.

Sembiring, R.K. 1995. *Analisis Regresi*. Bandung: Penerbit ITB.

Somantri, A dan Sambas A.M. 2006. *Aplikasi statistika dalam Penelitian*. pustaka ceria : Bandung.

Setyawan, B. 2017. *Estimasi Parameter Model Regresi Isotonik Bayesian dengan Polinomial Bernstein*. Skripsi. Program Studi Statistika Depatemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Gadjah Mada Yogyakarta.

Tenbusch, A. 1997. *Nonparametric curve estimation with Bernstein estimates.* Metrika, 45, 1-30.

Walpole, R.E., 1995, *Pengantar Statistika*, diterjemahkan oleh Sumantri, B., Edisi 2, Gramedia Pustaka Utama, Jakarta.

Wowor, A. D. 2017. *Regrenerasi Fungsi Polinomial Dalam Perancangan Algoritma Berbasis CSPRNG CHAOS Sebagai Pembbangkit Kunci Pada Kriptografi Block Cipher*. Jurnal Math And Its Appl, Vol 14, No. 1, Mei 2017, Halaman 1-15.

Wulandari, E. 2006. *Penentuan Estimator Parameter Distribusi Cauchy dengan Metode Estimasi-M*. Skripsi. Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Airlangga Surabaya.

Yendra, R dan Elsa T. R. 2015. *Perbandingan Estimasi Parameter Pada Distribusi Eksponensial Dengan Menggunakan Metode Maksimum Likelihood Dan Metode Bayesian.* Jurnal Sains Matematika dan Statistika, Volume 1, Nomor 2, Juli 2015, Halaman 62-72, ISSN 2460-4542.

# LAMPIRAN

**Lampiran 1.** Data pendidikan, rekreasi dan olahraga (*y*) dan makanan jadi, minuman, rokok dan tembakau (*x*)

| No | Tahun dan Bulan | *y* | *x* |
|---|---|---|---|
| 1 | Mei-06 | 135.43 | 136.86 |
| 2 | Jun-06 | 135.78 | 137.2 |
| 3 | Jul-06 | 136.2 | 138.15 |
| 4 | Agu-06 | 136.68 | 144.74 |
| 5 | Sep-06 | 136.86 | 147.41 |
| 6 | Okt-06 | 137.74 | 147.56 |
| 7 | Nov-06 | 138.39 | 147.6 |
| 8 | Des-06 | 139.93 | 147.7 |
| 9 | Jan-07 | 141.15 | 147.85 |
| 10 | Feb-07 | 142.07 | 148.19 |
| 11 | Mar-07 | 142.58 | 148.23 |
| 12 | Apr-07 | 143.12 | 148.18 |
| 13 | Mei-07 | 143.79 | 148.19 |
| 14 | Jun-07 | 144.27 | 148.24 |
| 15 | Jul-07 | 144.84 | 148.52 |
| 16 | Agu-07 | 145.53 | 152.37 |
| 17 | Sep-07 | 146.19 | 160.05 |
| 18 | Okt-07 | 146.93 | 160.38 |
| 19 | Nov-07 | 147.56 | 160.55 |
| 20 | Des-07 | 148.9 | 160.74 |
| 21 | Jan-08 | 151.9 | 160.75 |
| 22 | Feb-08 | 153.24 | 160.82 |
| 23 | Mar-08 | 154.89 | 160.97 |
| 24 | Apr-08 | 156.22 | 161.18 |
| 25 | Mei-08 | 157.57 | 161.78 |

Sumber :

Data Skripsi Adawiyah (2011).

Keterangan:

$x$ = pendidikan, rekreasi dan olahraga

$y$ = makanan jadi, minuman, rokok dan tembakau

**Lampiran 2.** *Script* R Regresi Linear Sederhana

```
##memuat data yang digunakan##
Y= scan()
X= scan()

##Plot data##
#variabel Y yaitu makanan jadi, minuman, rokok, dan tembakau#
#variabel X yaitu pendidikan, rekreasi, dan olahraga#
plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan
            jadi, minuman, rokok, dan tembakau", main="Scatter Plot
            pendidikan, rekreasi, dan olahraga dengan makanan jadi,
            minuman, rokok, dan tembakau")

##Regresi Liniear Sederhana##
reg.lin.sederhana=lm(Y~X)
summary(reg.lin.sederhana)

##Prediksi Regresi Linear Sederhana##
ytopilinear=32.47091 + 0.73886*(X)
ytopilinear

##Sum Square Error Regresi Linear Sederhana##
SSE_RLS=sum((Y-ytopilinear)^2)
SSE_RLS

##Mean Square Error Regresi Linear Sederhana##
MSE=SSE_RLS/n
MSE

##Plot Data Regresi Linear##
plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan
            jadi, minuman, rokok, dan tembakau", main="Scatter Plot
            pendidikan, rekreasi, dan olahraga dengan makanan jadi,
            minuman, rokok, dan tembakau")
lines(sort(X),sort(ytopilinear), col="red")
legend("topleft", c("Linear Fit"), bty="n", lwd=2, col=c("red"))

##Plot Data Perbandingan##
plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan
            jadi, minuman, rokok, dan tembakau", main="Scatter Plot
            pendidikan, rekreasi, dan olahraga dengan makanan jadi,
            minuman, rokok, dan tembakau")
lines(sort(X),sort(ytopilinear), col = "red")
lines(X,ytopi, col= "blue")
legend("topleft", c("Linear Fit", "Isotonic Fit"), bty="n", lwd=2,
            col=c("red","blue"))
legend("topleft", c("Linear Fit"), bty="n", lwd=2, col=c("red"))
```

**Lampiran 3.** *Script* R *Maximum Likelihood Estimator* untuk Regresi Isotonik dengan pendekatan Polinomial Bernstein

```
##memuat package yang diperlukan##
library(stats4)

##memuat data yang digunakan##
Y= scan()
X= scan()

##Plot data##
#variabel Y yaitu makanan jadi, minuman, rokok, dan tembakau#
#variabel X yaitu pendidikan, rekreasi, dan olahraga#
plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan
            jadi, minuman, rokok, dan tembakau", main="Scatter Plot
            pendidikan, rekreasi, dan olahraga dengan makanan jadi,
            minuman, rokok, dan tembakau")

#Polinomial Beirstein#
n=length(X)
n

W0=array(dim=c(n))
for(i in 1:n)
{
W0[i]= (choose(1,0)*(1-X[i]))+(choose(1,1)*X[i])
}
W0

W1=array(dim=c(n))
for(i in 1:n)
{
W1[i]= (choose(1,1)*X[i])
}
W1

##Maximum  Likelihood  Estimator  untuk  Estimasi  Model  Regresi
            Isotonik dengan Pendekatan Polinomial Bernstein##
Y
W0
W1
L=function(U0, U1, sigma)
{
-1*(-n/2*log(2*pi*sigma^2)-(1/(2*sigma^2))*(sum((Y-W0*U0-
            W1*U1)^2)))
}
estimasi=mle(minuslog=L,start=list(U0=0.005,U1=0.007,sigma=0.006))
summary(estimasi)
```

```
##Parameter Regresi Isotonik dengan MLE##
U0= 31.8159778
U1=  0.7431715

##Prediksi Maximum Likelihood Estimator untuk Estimasi Model
            Regresi Isotonik dengan Pendekatan Polinomial
            Bernstein##
ytopi=array(dim=c(n))
for(i in 1:n)
{
ytopi=W0*U0+W1*U1
}
ytopi

##Menghitung Sum Square Error (SSE) dan Mean Square Error (MSE)
            Regresi Isotonik##
SSE_isotonik=sum((Y-ytopi)^2)
SSE_isotonik

MSE_isotonik=SSE_isotonik/n
MSE_isotonik

##Menghitung R-square Regresi isotonik##
mean=mean(Y)
mean
SST=sum((Y-mean)^2)
SST
Rsq=1-(SSE_isotonik/SST)
Rsq

##Plot Data dengan kurva regresi isotonik##
plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan
            jadi, minuman, rokok, dan tembakau", main="Scatter Plot
            pendidikan, rekreasi, dan olahraga dengan makanan jadi,
            minuman, rokok, dan tembakau")
lines(X,ytopi, col= "blue")
legend("topleft", c("Isotonic Fit"), bty="n", lwd=2, col=c("blue"))
```

**Lampiran 4.** *Script* R Plot Data Perbandingan dan Korelasi

```
##Plot Data Perbandingan##
plot(X,Y, xlab="pendidikan, rekreasi, dan olahraga", ylab="makanan
            jadi, minuman, rokok, dan tembakau", main="Scatter Plot
            pendidikan, rekreasi, dan olahraga dengan makanan jadi,
            minuman, rokok, dan tembakau")
lines(sort(X),sort(ytopilinear), col = "red")
lines(X,ytopi, col= "blue")
legend("topleft", c("Linear Fit", "Isotonic Fit"), bty="n", lwd=2,
            col=c("red","blue"))

##korelasi##
korelasi = cor(Y,X)
korelasi
```

**Lampiran 5.** *Output* Analisis Regresi Linear Sederhana

```
> ##Regresi Liniear Sederhana##
> reg.lin.sederhana=lm(Y~X)
> summary(reg.lin.sederhana)

Call:
lm(formula = Y ~ X)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.5349 -2.7330  0.5884  1.9380  5.5669 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 32.47091   11.48167   2.828  0.00953 ** 
X            0.73886    0.07575   9.754 1.22e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.994 on 23 degrees of freedom
Multiple R-squared:  0.8053,	Adjusted R-squared:  0.7969 
F-statistic: 95.14 on 1 and 23 DF,  p-value: 1.221e-09
```

```
> ytopilinear
 [1] 133.5913 133.8425 134.5444 139.4135 141.3863 141.4971 141.5266 141.6005
 [9] 141.7114 141.9626 141.9921 141.9552 141.9626 141.9995 142.2064 145.0510
[17] 150.7255 150.9693 151.0949 151.2353 151.2427 151.2944 151.4052 151.5604
[25] 152.0037
> 
> ##Sum Square Error Regresi Linear Sederhana##
> SSE_RLS=sum((Y-ytopilinear)^2)
> SSE_RLS
[1] 206.1611
> 
> ##Mean Square Error Regresi Linear Sederhana##
> MSE=SSE_RLS/n
> MSE
[1] 8.246446
```

**Lampiran 6.** *Output* Analisis *Estimasi Maximum Likelihood* untuk Regresi Istonik dengan Pendekatan Polinomial Bernstein

```
> L=function(U0, U1, sigma)
+ {
+ -1*(-n/2*log(2*pi*sigma^2)-(1/(2*sigma^2))*(sum((Y-W0*U0-W1*U1)^2)))
+ }
> estimasi=mle(minuslog=L,start=list(U0=0.005,U1=0.007,sigma=0.006))
> summary(estimasi)
Maximum likelihood estimation

Call:
mle(minuslogl = L, start = list(U0 = 0.005, U1 = 0.007, sigma = 0.006))

Coefficients:
        Estimate  Std. Error
U0    31.8159778 10.91255986
U1     0.7431715  0.07199459
sigma  2.8451169  0.39683314

-2 log L: 123.6994
```

```
> ##Prediksi Maximum Likelihood Estimator untuk Estimasi Model Regresi Isotonik dengan Pendekatan Polinomial Bernstein##
> ytopi=array(dim=c(n))
> for(i in 1:n)
+ {
+ ytopi=W0*U0+W1*U1
+ }
> ytopi
 [1] 133.5264 133.7791 134.4851 139.3826 141.3669 141.4784 141.5081 141.5824
 [9] 141.6939 141.9466 141.9763 141.9391 141.9466 141.9837 142.1918 145.0530
[17] 150.7606 151.0058 151.1322 151.2734 151.2808 151.3328 151.4443 151.6004
[25] 152.0463
>
> ##Menghitung Sum Square Error (SSE) dan Mean Square Error (MSE) Regresi Isotonik##
> SSE_isotonik=sum((Y-ytopi)^2)
> SSE_isotonik
[1] 206.1903
>
> MSE_isotonik=SSE_isotonik/n
> MSE_isotonik
[1] 8.247612
>
> ##Menghitung R-square Regresi isotonik##
> mean=mean(Y)
> mean
[1] 144.3104
> SST=sum((Y-mean)^2)
> SST
[1] 1058.95
> Rsq=1-(SSE_isotonik/SST)
> Rsq
[1] 0.8052881
> |
```

**Lampiran 7.** *Output* Korelasi

```
> ##korelasi##
> korelasi = cor(Y,X)
> korelasi
[1] 0.8973938
```