REGRESI LINIER NONPARAMETRIK DENGAN METODE THEIL

SKRIPSI

ALDILA SARTI
0810433067

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS ANDALAS
PADANG 2012

KATA PENGANTAR

Syukur alhamdulillah, segala puji Penulis haturkan atas kehadirat Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya sehingga Penulis dapat menyelesaikan skripsi ini dengan judul "Regresi Linier Nonparametrik dengan Metode Theil". Shalawat dan salam semoga selalu tercurahkan kepada Rasulullah SAW yang telah menebarkan ilmu dan iman dalam cahaya Islam yang beliau bawa. Penulis menyampaikan ungkapan terima kasih dan penghargaan yang tulus kepada yang terhormat:

1. Ibu Izzati Rahmi HG, M.Si dan Ibu Hazmira Yozza, M.Si, sebagai pembimbing yang telah bersedia meluangkan waktu untuk memberikan bimbingan, pengarahan, dan saran kepada Penulis sampai selesainya skripsi ini.
2. Bapak Dr.Dodi Devianto, Bapak Yudiantri Asdi, M.Sc dan Ibu Nova Noliza Bakar, M.Si, sebagai penguji yang telah memberikan pengarahan dan saran untuk perbaikan penulisan skripsi ini.
3. Bapak Yudiantri Asdi, M.Sc, selaku pembimbing akademis yang telah memberikan nasehat dan motivasi kepada Penulis.
4. Bapak Dr. Syafrizal Sy, sebagai ketua jurusan pada jurusan Matematika FMIPA Universitas Andalas yang telah banyak membantu Penulis selama studi di jurusan Matematika Universitas Andalas.
5. Seluruh staf pengajar jurusan Matematika Universitas Andalas yang telah banyak memberikan ilmu yang bermanfaat bagi penulis dan seluruh staf tata usaha jurusan Matematika yang telah banyak membantu selama Penulis melaksanakan studi di jurusan Matematika Universitas Andalas.

6. Seluruh teman-teman yang telah mendukung dan memberikan semangat kepada Penulis terutama teman-teman angkatan 2008 (O'Laplace). Buat Ute, Putri, Kak Cesa, Kak Wiwik, Oni, dan Rara sahabat yang selalu bersama dari awal hingga kuliah selesai, buat Lia dan Kak Fatrika yang sama-sama berjuang dari awal penulisan skripsi hingga selesai, serta kakak-kakak senior dan adik-adik junior yang tidak bisa disebutkan satu persatu di Himpunan Mahasiswa Matematika (HIMATIKA) jurusan Matematika Universitas Andalas.
7. Semua pihak yang telah membantu Penulis yang tidak dapat disebutkan satu persatu.

Selesainya skripsi ini tidak lepas dari do'a yang tulus, motivasi, dorongan semangat, dan bantuan yang senantiasa diberikan oleh kedua orang tua, ayahanda Suardi Onyot dan ibunda Elmi (almarhumah), empat orang abangda, adinda Mia Audina dan seluruh keluarga besar Penulis.

Penulis menyadari penulisan skripsi ini masih jauh dari kesempurnaan. Oleh karena itu, dengan kerendahan hati penulis mengharapkan kritik dan saran agar kedepannya diperoleh hasil yang lebih baik. Penulis berharap agar skripsi ini dapat bermanfaat bagi semua pihak yang membacanya. Amin.

Padang, Agustus 2012

Penulis

ABSTRAK

Analisis regresi linier adalah analisis terhadap hubungan satu variabel terikat (y) dengan satu atau lebih variabel bebas (x). Pendugaan parameter biasanya diselesaikan dengan Metode Kuadrat Terkecil (MKT) yang harus memenuhi asumsi-asumsi tertentu. Salah satu asumsi yang harus dipenuhi dalam MKT adalah kenormalan dari galat, yaitu galat berdistribusi normal dengan rata-rata nol dan simpangan baku tertentu. Jika asumsi kenormalan galat tidak dapat dipenuhi, maka MKT tidak dapat digunakan. Analisis alternatif yang dapat digunakan untuk mengatasi masalah tersebut adalah prosedur regresi nonparametrik, salah satunya dengan metode Theil. Metode Theil adalah metode nonparametrik yang digunakan untuk menduga parameter-parameter pada model regresi linier berdasarkan data sampel yang teramati, dengan kondisi galat tidak menyebar normal. Skripsi ini membahas metode Theil yang menduga koefisien kemiringan (slope) sebagai median kemiringan dari seluruh pasangan garis dari titik-titik variabel x dan y dengan syarat semua nilai x_i harus berbeda. Pengujian hipotesis parameternya didasarkan pada statistik Tau Kendall. Interval kepercayaannya hanya untuk koefisien kemiringan (*slope*). Contoh penerapan untuk data banyak hafalan Al-Quran siswa dengan variabel terikat y_i adalah banyak hafalan Al-Quran siswa dan variabel bebas x_i adalah lama menghafal dapat disimpulkan: (1) model regresi Theil yang diperoleh adalah $\hat{y}_i = 0{,}9990 + 0{,}1667x_i$, (2) lama menghafal berpengaruh terhadap banyak hafalan Al-Quran siswa, (3) selang kepercayaan koefisien *slope* yaitu (0.0833,0.3).

Kata kunci: *Analisis Regresi Linier, Metode Kuadrat Terkecil, Regresi Nonparametrik, Metode Theil, Tau Kendall*

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

No Halaman

BAB I

PENDAHULUAN

1.1. Latar Belakang

Analisis regresi didefinisikan sebagai kajian terhadap hubungan satu variabel yang diterangkan atau yang disebut variabel terikat dengan satu atau lebih variabel yang menerangkan atau yang disebut variabel bebas. Analisis regresi merupakan salah satu teknik statistika yang digunakan secara luas dalam ilmu pengetahuan terapan. Disamping digunakan untuk mengetahui bentuk hubungan antara variabel regresi, analisis regresi juga dapat digunakan untuk peramalan. [4]

Dalam kasus parametrik, Peneliti biasanya menggunakan Metode Kuadrat Terkecil (MKT) untuk menduga parameter-parameter regresi dengan data sampel yang teramati dan melandaskan kesimpulan-kesimpulan yang menyangkut parameter-parameter populasi pada asumsi-asumsi yang harus dipenuhi. Salah satu asumsi yang harus dipenuhi adalah kenormalan distribusi galat, yaitu bahwa galat berdistribusi normal dengan rata-rata nol dan simpangan baku tertentu.

Apabila asumsi kenormalan tidak dipenuhi, analisis alternatif yang dapat digunakan adalah dengan metode regresi nonparametrik, karena statistik nonparametrik tidak menuntut terpenuhi banyak asumsi, misalnya data yang dianalisis tidak harus berdistribusi normal. Oleh karena itu statistik nonparametrik sering disebut sebagai uji bebas distribusi (*distribution free test*).

Beberapa metode nonparametrik yang dapat digunakan untuk mencocokkan garis regresi linier dengan data sampel yang teramati adalah metode Theil, metode *Iterative Brown-Mood* dan metode *Weighted Medians*. Dari ketiga metode di atas, metode Theil adalah yang paling baik, karena penelitian yang dilakukan [10] berpendapat bahwa Metode Theil hampir seefisien Metode Kuadrat Terkecil jika kenormalan galat terpenuhi. [2]

Dalam tugas akhir ini akan dijelaskan prosedur regresi linier nonparametrik dengan metode Theil dan contoh penerapannya dengan data yang tidak berdistribusi normal. Metode Theil menduga koefisien kemiringan (*slope*) garis regresi dengan median kemiringan dari seluruh pasangan garis dari titik-titik variabel x dan y, dengan nilai x_i harus berbeda.

1.2. Perumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan dalam tugas akhir ini adalah bagaimana menentukan model regresi linier nonparametrik dengan metode Theil dan bagaimana menerapkan metode tersebut pada data?

1.3. Pembatasan Masalah

Tugas akhir ini dibatasi pada model regresi linier sederhana yaitu model dengan satu variabel bebas saja.

1.4. Tujuan Penulisan

Berdasarkan perumusan masalah di atas, maka tujuan tugas akhir ini adalah untuk mengkaji prosedur regresi linier nonparametrik dengan metode Theil serta menerapkan penggunaan metode Theil pada data.

1.5. Sistematika Penulisan

Penulisan tugas akhir ini terdiri atas:

Bab I : Pendahuluan

Bab ini berisi tentang latar belakang, perumusan masalah, pembatasan masalah, tujuan penelitian, dan sistematika penulisan.

Bab II : Landasan Teori

Bab ini merupakan teori-teori dasar yang mendukung penyelesaian tugas akhir ini.

Bab III : Pembahasan

Bab ini berisi pembahasan tentang prosedur regresi linier nonparametrik dengan metode Theil dan contoh penerapan.

Bab IV : Kesimpulan

Bab ini merupakan kesimpulan dari bab-bab sebelumnya.

BAB II

LANDASAN TEORI

2.1. Skala Pengukuran

Skala pengukuran adalah penempatan angka atau lambang untuk menyatakan suatu hasil pengamatan atau pengukuran yang dilakukan terhadap suatu objek. Secara umum, skala pengukuran dibedakan menjadi empat [8]:

1. Skala Nominal

   Skala nominal adalah skala pengukuran data yang digunakan hanya untuk membedakan suatu objek, orang, atau sifat. Jadi skala ini adalah hanya mengklasifikasi objek. Contoh skala ini adalah jenis kelamin, jenis warna dan merk motor.

2. Skala Ordinal

   Skala ordinal adalah skala yang digunakan untuk membedakan suatu objek dari objek lain dengan memberi lambang lebih besar atau lebih kecil tetapi angka yang digunakan untuk melambangkan kategori tidak memiliki nilai absolut, hanya menunjukkan posisi sebuah kategori relatif terhadap kategori lainnya. Jadi sifat skala ini adalah mengklasifikasi dan mengurutkan. Contoh skala ini adalah nilai mata kuliah (A, B, C, D, E) dan kerusakan (parah, sedang, ringan).

3. Skala Interval

Skala interval memiliki semua karakteristik skala ordinal, kecuali bahwa pada skala ini mempunyai satuan skala. Angka yang digunakan pada skala ini sudah memiliki arti. Selain itu, pada skala ini juga dapat dihitung jarak antara dua titik skala. Contoh skala ini adalah pengukuran suhu.

4. Skala Rasio

Skala rasio adalah skala yang memiliki ciri-ciri sama dengan skala interval, tetapi pada skala rasio perbandingan antar nilai sudah dapat diinterpretasikan. Contoh skala ini adalah massa, tinggi badan dan berat badan.

2.2. Koefisien Korelasi

Koefisien korelasi yang dinotasikan dengan r merupakan ukuran yang digunakan untuk mengukur derajat hubungan linier antara suatu variabel dengan variabel lainnya. Koefisien korelasi antara variabel x dan variabel y dirumuskan sebagai:

$$r = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{\left(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2\right)\left(n\sum_{i=1}^{n} y_i^2 - (\sum_{i=1}^{n} y_i)^2\right)}} \quad (2.2.1)$$

dengan $n =$ banyak data.

Koefisien korelasi (r) dapat digunakan untuk [1]:

1. Mengetahui derajat (keeratan) hubungan (korelasi linier) antara dua variabel.
2. Mengetahui arah hubungan antara dua variabel.

Nilai koefisien korelasi bernilai antara -1 sampai 1, atau dapat dituliskan $-1 \leq r \leq 1$.

Untuk melakukan interpretasi kekuatan hubungan antara dua variabel dilakukan dengan melihat angka koefisien korelasi hasil perhitungan dengan menggunakan kriteria sebagai berikut [7]:

- Jika angka koefisien korelasi menunjukkan 0, maka kedua variabel tidak mempunyai hubungan linier.
- Semakin dekat nilai koefisien korelasi dengan 1 atau -1, maka semakin kuat hubungan linier kedua variabel.
- Semakin dekat nilai koefisien korelasi dengan 0, maka semakin lemah hubungan linier kedua variabel.
- Jika angka koefisien korelasi sama dengan 1, maka kedua variabel mempunyai hubungan linier positif sempurna.
- Jika angka koefisien korelasi sama dengan -1, maka kedua variabel mempunyai hubungan linier negatif sempurna.

Tanda dari koefisien korelasi dapat digunakan untuk mengetahui arah hubungan antara dua variabel, tanda positif (+) pada nilai koefisien korelasi menunjukkan hubungan yang searah, artinya apabila variabel x naik, maka variabel y juga naik. Tanda negatif (-) pada koefisien korelasi menunjukkan hubungan yang berlawanan arah, artinya apabila nilai variabel x naik, maka nilai variabel y turun, sebaliknya apabila nilai variabel x turun, maka nilai variabel y naik. [7]

2.3. Analisis Regresi Parametrik

Analisis regresi merupakan salah satu analisis statistika yang digunakan untuk memodelkan hubungan antara dua variabel atau lebih dalam suatu sistem. Dalam analisis regresi, dikenal dua jenis variabel yaitu [7]:

1. Variabel tak bebas/terikat (*dependent variable*) yaitu variabel yang nilainya dipengaruhi oleh nilai variabel lain dalam suatu sistem.
2. Variabel bebas (*independent variable*) yaitu variabel yang nilainya tidak dipengaruhi oleh nilai variabel lain dalam suatu sistem.

Hubungan yang didapat pada umumnya dinyatakan dalam bentuk persamaan matematika yang menyatakan hubungan antara variabel terikat (*dependent variable*) y dengan satu atau lebih variabel bebas (*independent variable*) $x_1, x_2, \dots, x_n$. [6]

Metode statistika yang digunakan untuk membentuk model hubungan antara variabel terikat y dengan satu atau lebih variabel bebas x, dimana hubungan antara variabel tersebut linier dalam parameter disebut dengan regresi linier. Persamaan model untuk regresi linier dapat ditulis sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i \quad , i = 1,2,\dots,n \tag{2.3.1}$$

dengan:

y_i : variabel terikat,

x_i : variabel bebas,

$\beta_0, \beta_1, \dots, \beta_p$: parameter-parameter regresi,

ε_i : galat.

Jika suatu analisis regresi linier mengkaji hubungan linier antara satu variabel bebas dengan satu variabel terikat, maka analisis regresi linier tersebut dinamakan **analisis regresi linier sederhana**. Model regresi linier sederhana secara umum adalah sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad ; i = 1,2,\ldots,n, \tag{2.3.2}$$

dengan y_i adalah nilai variabel terikat dari data ke-i, x_i adalah nilai variabel bebas dari data ke-i, β_0 adalah koefisien intersep (*intercept*) dari garis regresi, β_1 adalah koefisien kemiringan (*slope*) garis regresi, dan ε_i merupakan galat data ke-i. [5]

Pada model regresi Persamaan (2.3.2) terdapat koefisien. Koefisien pada model regresi adalah nilai dugaan parameter dalam model regresi. Koefisien regresi dapat dibedakan menjadi dua macam, yaitu [7]:

1. Koefisien Intersep (*Intercept*)

   Secara statistika defenisi intersep adalah nilai rata-rata pada variabel terikat y jika nilai variabel bebas x adalah nol. Dengan kata lain, apabila variabel x tidak memberikan kontribusi maka secara rata-rata variabel y akan bernilai sebesar intersep. Intersep tidak selalu dapat atau perlu diinterpretasikan, karena apabila data pengamatan pada variabel x tidak mencakup nilai nol, maka intersep tidak memiliki makna yang berarti sehingga tidak perlu diinterpretasikan.

2. Koefisien Kemiringan (*Slope*)

   Slope adalah koefisien regresi untuk variabel bebas x. Dalam konsep statistika, *slope* merupakan suatu nilai yang menunjukkan seberapa besar kontribusi atau pengaruh variabel bebas x terhadap variabel terikat y. Dengan kata lain, *slope* diartikan sebagai rata-rata pertambahan yang terjadi pada variabel terikat y untuk setiap peningkatan satu satuan variabel bebas x.

2.3.1. Regresi Parametrik dengan Metode Kuadrat Terkecil

Regresi parametrik merupakan regresi yang digunakan untuk mengetahui bentuk hubungan antara variabel bebas x dan variabel terikat y. Setelah variabel bebas x diamati, nilai dugaan variabel terikat y diberikan oleh fungsi regresi. Dugaan kurva regresi umumnya dilakukan dengan pendekatan parametrik yang mulai diperkenalkan oleh Laplace sejak abad XVIII dan Boscovich pada tahun 1757.

Model regresi linier parametrik harus memenuhi asumsi-asumsi sebagai berikut [1]:

1. Tidak terjadi multikolinieritas antar variabel artinya antara variabel bebas yang satu dengan variabel bebas yang lain dalam model regresi tidak terdapat hubungan linier.
2. Homoskedastisitas yaitu varians dari semua nilai galat konstan (sama).
3. Galat ε_i menyebar menurut sebaran normal dengan nilai tengah nol dan ragam tertentu, dengan kata lain $\varepsilon_i \sim N(0, \sigma^2)$.
4. Tidak terdapat autokorelasi antara galat ε_i.

Metode Kuadrat Terkecil (*Ordinary Least Square*) merupakan suatu metode yang digunakan untuk menduga parameter dari model regresi parametrik dengan cara meminimumkan Jumlah Kuadrat Sisaan (JKS). Misalkan model regresi linier sederhana $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1,2, \ldots, n$. Dugaan dari model tersebut adalah $y_i = b_0 + b_1 x_i + e_i$ dengan e_i adalah sisaan data ke-i.

Untuk objek pengamatan ke-i, nilai sisaan e_i adalah selisih antara nilai y_i yang sesungguhnya dengan dugaan dari y_i yang diperoleh dari persamaan:

$$\hat{y}_i = b_0 + b_1 x_i \qquad ; i = 1,2, \dots, n. \qquad (2.3.1.1)$$

Dengan demikian Jumlah Kuadrat Sisaan (JKS) untuk persamaan regresi linier sederhana tersebut adalah:

$$J = \textstyle\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2. \qquad (2.3.1.2)$$

Pada Metode Kuadrat Terkecil (MKT), akan ditentukan b_0 dan b_1 yang dapat meminimumkan nilai JKS. JKS minimum dapat ditentukan dengan cara mencari turunan JKS terhadap b_0 dan b_1, yaitu:

$$\frac{\partial}{\partial b_0}\left(\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2\right) = -2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)$$

$$\frac{\partial}{\partial b_1}\left(\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2\right) = -2\sum_{i=1}^{n}((y_i - b_0 - b_1 x_i)(x_i)).$$

Dengan menyamakan kedua turunan tersebut dengan 0, akan diperoleh:

$$-2\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i) = 0 \Leftrightarrow \sum_{i=1}^{n}(y_i - b_0 - b_1 x_i) = 0 \Leftrightarrow n b_0 + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \qquad (2.3.1.3)$$

$$-2\sum_{i=1}^{n}((y_i - b_0 - b_1 x_i)(x_i)) = 0 \Leftrightarrow \sum_{i=1}^{n}((y_i - b_0 - b_1 x_i)(x_i)) = 0 \Leftrightarrow b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i.$$

$$(2.3.1.4)$$

Berdasarkan Persamaan (2.3.1.3) didapatkan b_0, yaitu:

$$n b_0 + b_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$b_0 = \frac{1}{n}\left(\sum_{i=1}^{n} y_i - b_1 \sum_{i=1}^{n} x_i\right) = \bar{y} - b_1 \bar{x}. \qquad (2.3.1.5)$$

kemudian Persamaan (2.3.1.5) disubstitusikan ke Persamaan (2.3.1.4), diperoleh:

$$b_0 \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i-1}^{n} x_i y_i$$

$$(\bar{y} - b_1 \bar{x}) \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i-1}^{n} x_i y_i$$

$$\bar{y} \sum_{i=1}^{n} x_i - b_1 \bar{x} \sum_{i=1}^{n} x_i + b_1 \sum_{i=1}^{n} x_i^2 = \sum_{i-1}^{n} x_i y_i$$

$$b_1 \left(\sum_{i=1}^{n} x_i^2 - \bar{x} \sum_{i=1}^{n} x_i \right) = \sum_{i-1}^{n} x_i y_i - \bar{y} \sum_{i=1}^{n} x_i$$

$$b_1 = \frac{\sum_{i-1}^{n} x_i y_i - \bar{y} \sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \bar{x} \sum_{i=1}^{n} x_i} = \frac{\sum_{i-1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

$$b_1 = \frac{\sum_{i-1}^{n} x_i y_i - [(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)]/n}{\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n}. \quad (2.3.1.6)$$

Jadi, dugaan b_1 dan b_0 dari persamaan regresi linier sederhana dengan menggunakan Metode Kuadrat Terkecil dinotasikan dengan $\hat{\beta}_1$ yaitu nilai yang diduga dari b_1 dan $\hat{\beta}_0$ yaitu nilai yang diduga dari b_0, dengan rumus sebagai berikut:

$$\hat{\beta}_1 = \frac{\sum_{i-1}^{n} x_i y_i - [(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)]/n}{\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

dengan n adalah banyak data, $\bar{y}$ adalah nilai rata-rata y.

Setelah $\hat{\beta}_1$ dan $\hat{\beta}_0$ diperoleh, maka dugaan model regresi linier sederhana adalah sebagai berikut:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i, \qquad i = 1,2,\dots,n. \tag{2.3.1.7}$$

2.3.2. Koefisien Determinasi

Koefisien determinasi yang diberi simbol R^2 merupakan ukuran yang dapat digunakan untuk mengetahui besarnya kontribusi varibel bebas terhadap variabel terikat dari suatu persamaan regresi. Apabila nilai koefisien determinasi R^2 dikalikan 100%, maka hal ini menunjukkan presentase total keragaman nilai variabel terikat yang dapat diterangkan oleh model regresi yang dihasilkan.

Besarnya koefisien determinasi dinyatakan dalam bilangan antara 0 sampai 1. Apabila koefisien determinasi mendekati 0, maka kecil kontribusi variabel bebas terhadap variabel terikat. Sebaliknya, jika koefisien determinasi mendekati 1, maka besar kontiribusi variabel bebas terhadap variabel terikat. Nilai koefisien determinasi dapat ditentukan menggunakan rumus berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{2.3.2.1}$$

dengan $0 \le R^2 \le 1$, y adalah nilai variabel terikat yang diperoleh dari data, $\hat{y}_i$ adalah nilai variabel terikat yang diduga dari model, $\bar{y}$ adalah rata-rata variabel terikat dengan rumus $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$. [1]

2.4. Uji Kenormalan

Uji statistika parametrik adalah suatu uji yang modelnya menetapkan syarat-syarat tertentu (asumsi-asumsi) dari sebaran (distribusi) data populasinya.

Statistika parametrik lebih banyak digunakan untuk menganalisis data yang berskala interval atau rasio dengan dilandasi asumsi tertentu seperti kenormalan. Selain itu, ukuran datanya biasanya besar, sekurang-kurangnya lebih besar atau sama dengan 30 data.

Kenormalan merupakan salah satu asumsi standar pada uji statistik parametrik. Uji kenormalan adalah uji untuk mengukur apakah data yang dimiliki berdistribusi normal sehingga dapat dipakai dalam statistik parametrik. Untuk melihat penyimpangan kenormalan secara umum, dapat dibuat histogram dari data. Data yang berdistribusi normal adalah data yang memencar mengikuti kurva normal yaitu kurva yang menunjukkan frekuensi tertinggi berada ditengah-tengah dan bagian sisi kiri dengan sisi kanannya simetris. Cara lain yang biasa digunakan untuk melihat kenormalan adalah dengan uji *Kolmogorov-Smirnov* dengan hipotesis:

H_0 : Data berdistribusi normal

H_1 : Data tidak berdistribusi normal.

Statistik Uji yang digunakan adalah:

$$D = maksimum\ \{|F_0(x_i) - S_n(x_i)|,\ i = 1,2,\dots,n\} \quad (2.4.1.1)$$

dengan $F_0(x)$: nilai fungsi kumulatif sebaran, yang merupakan $P(X \leq x)$ jika asumsi $X \sim N(\mu, \sigma^2)$ benar. Dalam hal ini, μ diduga dari $\bar{x}$ dan σ^2 diduga dari S^2.

$S_N(x)$: peluang kumulatif sampel, dimana $S_N(x) = {}^{k}/_{N}$, dengan k adalah banyak nilai data sampel yang kurang dari atau sama dengan x.

n : banyak data

Tolak H_0 jika $D \geq D_\alpha$ atau terima H_0 jika $D < D_\alpha$, dimana D_α adalah nilai-nilai kritis dalam tabel statistik uji *Kolmogorov-Smirnov* seperti yang disajikan pada Lampiran 6, atau dengan memperhatikan nilai signifikansi dengan pengambilan keputusan dilakukan sebagai berikut:

jika signifikansi < 0,05 maka tolak H_0 yang berarti data berdistribusi tidak normal,

jika signifikansi > 0,05 maka tidak tolak H_0 yang berarti data berdistribusi normal.

2.5. Regresi Nonparametrik

Analisis regresi nonparametrik dikenal sebagai alat analisis statistik alternatif saat analisis parametrik tidak dapat menggunakan. Analisis regresi nonparametrik adalah prosedur statistik yang tidak mengacu pada parameter tertentu.

Dalam banyak hal, data-data yang akan dikaji tidak selalu memenuhi asumsi-asumsi yang mendasari uji-uji parametrik sehingga sering sekali dibutuhkan teknik-teknik statistika dengan validitas yang tidak bergantung pada asumsi-asumsi yang kaku. Dalam hal ini, teknik-teknik dalam regresi nonparametrik memenuhi kebutuhan ini karena tetap valid walaupun tidak diperlukan pemenuhan asumsi kenormalan galat dan hanya berlandaskan asumsi-asumsi yang sangat umum. [2]

Penggunaan regresi nonparametrik dilandasi pada asumsi [1]:

1. Data yang diambil bersifat acak,
2. Data berskala nominal atau ordinal,
3. Regresi antara variabel y dengan variabel x bersifat linier,
4. Peubah x_i tidak berkolerasi.

2.6. Kombinasi

Beberapa metode nonparametrik membutuhkan pengetahuan mengenai jumlah subkumpulan yang terdiri dari r data yang dapat dibentuk dari sekumpulan n data. Hal ini disebut kombinasi. [10]

Definisi 2.6.1 [10]

Kombinasi r unsur yang diambil dari n unsur yang berbeda adalah suatu pilihan dari r unsur tanpa memperhatikan urutannya $(r \leq n)$. Banyaknya kombinasi r unsur yang diambil dari n unsur yang tersedia dilambangkan dengan notasi C_r^n.

Teorema 2.6.2 [10]

Untuk $r \leq n$, $C_r^n = \frac{n!}{r!(n-r)!}$.

Bukti

$$C_r^n = \frac{P_r^n}{r!}$$

$$= \frac{n!/(n-r)!}{r!}$$

$$= \frac{(n!)}{r!\,(n-r)!}$$

dengan $P_r^n = n(n-1)(n-2)\ldots(n-r+1) = \frac{n(n-1)(n-2)\ldots 2.1}{(n-r)(n-r-1)\ldots 2.1}$

$$= n!/(n-r)!.$$

2.7. Median

Menurut [11], median adalah nilai yang terdapat persis di tengah-tengah jika sekumpulan data diurutkan dari yang terkecil sampai terbesar. Terdapat 2 kasus dalam menentukan median (M_e), yaitu:

- Kasus 1: Jika banyak data (n) ganjil, maka mediannya adalah data yang paling tengah setelah data disusun menurut urutan nilainya. Dapat ditulis sebagai berikut:

$$M_e = x_{\left(\frac{1}{2}(n+1)\right)} \qquad (2.7.1)$$

- Kasus 2: Jika banyak data (n) genap, maka mediannya adalah rata-rata hitung dua data tengah setelah data disusun menurut urutan nilainya. Dapat ditulis sebagai berikut:

$$M_e = \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n}{2}+1\right)}\right). \qquad (2.7.2)$$

2.8. Koefisien Korelasi Tau Kendall

Koefisien Korelasi Tau Kendall merupakan salah satu koefisien korelasi nonparametrik yang dapat digunakan untuk data minimal berskala ordinal, yang berarti datanya dapat berbentuk peringkat. Korelasi Tau Kendall pertama kali diperkenalkan oleh G.T.Fechner pada tahun 1900 dan disempurnakan kembali oleh M.G.Kendall pada tahun 1938. Jika ada kecendrungan untuk nilai x yang tinggi, maka akan dihubungkan dengan nilai y yang tinggi pula. [2]

Misalkan diperoleh data-data (x_i, y_i) dan (x_j, y_j) maka jika $x_i > x_j$ sangat mungkin bahwa $y_i > y_j$. Pasangan-pasangan data yang mempunyai sifat ini digambarkan sebagai pasangan yang serasi (*concordant*) yang banyaknya disimbolkan dengan N_c. Contoh data (2,3) dan (4,5) adalah pasangan serasi (*concordant*) karena 4 > 2 dan 5 > 3 sehingga mempunyai tanda yang sama. Jika tanda dari pembilang dan penyebut adalah berlawanan atau tidak sama sehingga $(y_j - y_i)/(x_j - x_i)$ adalah negatif maka pasangan dikatakan tidak serasi (*discordant*) yang banyaknya disimbolkan dengan N_d. Contoh data (3,4) dan (5,2) adalah pasangan tidak serasi karena 5 > 3 tetapi 2 < 4. [10]

Koefisien korelasi Tau Kendall disimbolkan dengan τ untuk menyatakan parameter populasi dan $\hat{\tau}$ untuk menyatakan statistik sampelnya. Tau Kendall didasarkan pada peringkat-peringkat dari data. Tau Kendall memiliki nilai dari -1 sampai 1. Jika tidak ada angka x maupun y yang sama rumus yang digunakan untuk mendapatkan nilai koefisien korelasi τ Kendall adalah [10]:

$$\hat{\tau} = \frac{N_c - N_d}{n(n-1)/2} = \frac{S}{n(n-1)/2} \tag{2.8.1}$$

dengan n adalah banyak data,

N_c adalah jumlah pasangan data yang serasi,

N_d adalah jumlah pasangan data yang tidak serasi,

S adalah jumlah *concordant* (N_c) dikurangi *discordant* (N_d).

Jika dua data atau lebih pada varibel x maupun variabel y mempunyai nilai yang sama, pembagi pada Persamaan (2.8.1) untuk $\hat{\tau}$ menjadi [8]:

$$\hat{\tau} = \frac{S}{\sqrt{\frac{1}{2}n(n-1) - T_x}\sqrt{\frac{1}{2}n(n-1) - T_y}} \tag{2.8.2}$$

dimana $$T_x = \frac{1}{2}\sum_{i=1}^{m}(t_x)_i((t_x)_i - 1) \tag{2.8.3}$$

$$T_y = \frac{1}{2}\sum_{i=1}^{m}(t_y)_i((t_y)_i - 1) \tag{2.8.4}$$

dengan m : jumlah kelompok yang mempunyai angka yang sama pada variabel x atau y,

$(t_x)_i$: banyak nilai x yang sama untuk suatu data ke-i,

$(t_y)_i$: banyak nilai y yang sama untuk suatu data ke-i.

BAB III

REGRESI LINIER NONPARAMETRIK DENGAN METODE THEIL

Metode Theil adalah metode nonparametrik yang digunakan untuk menduga parameter-parameter pada model regresi linier berdasarkan data sampel yang teramati, dengan kondisi galat tidak menyebar normal. Untuk mengetahui galat menyebar normal atau tidak, maka terlebih dahulu dilakukan uji kenormalan terhadap residual, yaitu sisaan atau perbedaan antara nilai hasil pengamatan variabel terikat terhadap nilai hasil dugaan variabel terikat.

Metode Theil menduga koefisien kemiringan (*slope*) garis regresi dengan cara mencari median kemiringan seluruh pasangan garis dari titik-titik data (x_i, y_i), dengan syarat nilai x_i harus berbeda. Pengujian koefisien kemiringan (*slope*) disusun berdasarkan statistik τ Kendall yang digunakan untuk mengetahui bentuk hubungan variabel-variabel dalam persamaan regresi.

Asumsi-asumsi yang melandasi analisis regresi linier sederhana dengan metode Theil sebagai berikut [2]:

a. Persamaan regresinya adalah

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad i = 1,2,\dots,n \tag{3.1}$$

dengan y_i adalah variabel terikat, x_i adalah variabel bebas, β_0 dan β_1 adalah parameter-parameter yang tidak diketahui, ε_i adalah galat.

b. Semua nilai x_i berbeda dan dapat ditetapkan $x_1 < x_2 < \cdots < x_n$.

c. Nilai-nilai ε_i saling bebas.

3.1. Pendugaan Parameter-parameter Regresi Linier Sederhana

Misalkan terdapat n pasangan data $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ dan dari data tersebut akan dibentuk persamaan regresi linier sederhana sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \qquad i = 1,2, \dots, n$$

dengan y_i adalah nilai variabel terikat dari data ke-i, x_i adalah nilai variabel bebas dari data ke-i, β_0 adalah koefisien intersep, β_1 adalah koefisien kemiringan (*slope*) garis regresi, dan ε_i merupakan galat data ke-i. Selanjutnya akan dijelaskan penentuan dugaan untuk β_0 dan β_1 dengan notasi masing-masing $\tilde{\beta}_0$ dan $\tilde{\beta}_1$ yang akan dihitung dengan metode Theil.

3.1.1. Pendugaan Koefisien Kemiringan (*slope*)

Pada metode Theil, perkiraan kemiringan (*slope*) garis regresi merupakan median kemiringan *(slope)* dari seluruh pasangan garis yang menghubungkan pasangan titik-titik dengan nilai x yang berbeda. Metode ini dapat digunakan jika tidak ada nilai x_i yang bernilai sama, sehingga dapat ditetapkan $x_1 < x_2 < \cdots < x_n$. [12]

Untuk setiap pasangan (x_i, y_i) dan (x_j, y_j) nilai kemiringannya dinotasikan dengan b_{ij} dan dirumuskan sebagai:

$$b_{ij} = \frac{y_j - y_i}{x_j - x_i} \tag{3.2}$$

untuk $i < j$.

Untuk mendapatkan penduga $\tilde{\beta}_1$, pertama-tama hitung semua kemungkinan nilai b_{ij}. Dapat dibuktikan bahwa $b_{ij} = b_{ji}$ untuk seluruh i dan j.

Perhatikan bahwa

$$b_{ij} = \frac{y_j - y_i}{x_j - x_i} = \frac{-(y_j - y_i)}{-(x_j - x_i)} = \frac{(-y_j + y_i)}{(-x_j + x_i)} = \frac{y_i - y_j}{x_i - x_j} = b_{ji},$$

dengan demikian dapat dibuktikan bahwa $b_{ij} = b_{ji}$. Jika ada n banyaknya data, maka akan terdapat $N = \binom{n}{2} = \frac{n(n-1)}{2}$ kemungkinan nilai b_{ij} yang berbeda.

Penduga bagi β_1 yang dinotasikan dengan $\tilde{\beta}_1$ dihitung berdasarkan median dari b_{ij} dengan mengurutkan nilai b_{ij} dari terkecil sampai terbesar yang berjumlah N. Jadi, penduga koefisien kemiringan (*slope*) dapat dinyatakan sebagai berikut:

$$\tilde{\beta}_1 = median\,(b_{ij}), \tag{3.3}$$

atau misalkan $b_{(1)}, b_{(2)}, \dots, b_{(N)}$ adalah nilai-nilai b_{ij} yang telah diurutkan dari terkecil sampai terbesar, maka penduga bagi β_1 dapat juga ditulis sebagai berikut:

$$\tilde{\beta}_1 = \begin{cases} b_{\left(\frac{1}{2}(N+1)\right)} & , N\ ganjil \\ \frac{1}{2}\left(b_{\left(\frac{N}{2}\right)} + b_{\left(\frac{N}{2}+1\right)}\right) & , N\ genap. \end{cases} \tag{3.4}$$

3.1.2. Pendugaan Koefisien Intersep (*intercept*)

Setelah penduga $\tilde{\beta}_1$ diperoleh, maka persamaan regresinya ditulis:

$$y_i = \beta_0 + \tilde{\beta}_1 x_i, \qquad i = 1,2,\dots,n. \tag{3.5}$$

Penduga dari intersep β_0 dinotasikan dengan $\tilde{\beta}_0$. Misalkan disubstitusikan a_i ke β_0, maka akan diperoleh

$$y_i = a_i + \tilde{\beta}_1 x_i, \qquad i = 1,2,\dots,n \tag{3.6}$$

$$a_i = y_i - \tilde{\beta}_1 x_i. \tag{3.7}$$

Berdasarkan Persamaan (3.7) dapat ditentukan nilai a_i untuk semua data. Penduga $\tilde{\beta}_0$ dihitung berdasarkan median dari seluruh nilai a_i, dengan mengurutkan nilai a_i dari terkecil sampai terbesar yang berjumlah n. Maka penduga koefisien intersep dinyatakan sebagai berikut:

$$\tilde{\beta}_0 = median\ (a_i), \tag{3.8}$$

atau misalkan $a_{(1)}, a_{(2)}, \ldots, a_{(n)}$ adalah nilai-nilai a_i yang telah diurutkan dari terkecil sampai terbesar, maka penduga bagi β_0 dapat juga ditulis sebagai berikut:

$$\tilde{\beta}_0 = \begin{cases} a_{\left(\frac{1}{2}(n+1)\right)} & ,n\ ganjil \\ \frac{1}{2}\left(a_{\left(\frac{n}{2}\right)} + a_{\left(\frac{n}{2}+1\right)}\right) & ,n\ genap. \end{cases} \tag{3.9}$$

Berdasarkan Persamaan (3.3) dan (3.8), maka diperoleh dugaan model regresinya berbentuk $\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_i$, $i = 1,2, \ldots, n$.

3.2. Pengujian Koefisien Kemiringan (*Slope*)

Pengujian koefisien kemiringan garis regresi dengan menggunakan metode Theil disusun berdasarkan statistik τ Kendall. Pada pengujian ini dinyatakan hipotesis:

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

H_0 bermakna tidak ada pengaruh variabel x terhadap variabel y, sedangkan H_1 bermakna terdapat pengaruh variabel x terhadap variabel y.

Statistik uji dalam pengujian ini adalah koefisien korelasi Tau Kendall yang dirumuskan sebagai:

a. Jika tidak ada nilai x dan y yang sama, maka statistik ujinya

$$\hat{\tau} = \frac{N_c - N_d}{n(n-1)/2} = \frac{S}{n(n-1)/2}$$

dengan $\hat{\tau}$ = statistik uji τ Kendall,

n = banyaknya data yang diamati,

N_c = banyak pasangan yang serasi (*concordant*),

N_d = banyak pasangan yang tidak serasi (*discordant*),

S = selisih antara N_c dan N_d.

b. Jika ada nilai y yang sama, maka statistik ujinya

$$\hat{\tau} = \frac{S}{\sqrt{\frac{1}{2}n(n-1)}\sqrt{\frac{1}{2}n(n-1) - T_y}} \qquad (3.10)$$

dengan $T_y = \frac{1}{2}\sum_{i=1}^{m}(t_y)_i((t_y)_i - 1)$

m : jumlah kelompok yang mempunyai angka yang sama pada variabel y,

t_y : banyak nilai y yang sama untuk suatu data.

Terlihat bahwa rumus $\hat{\tau}$ pada pengujian ini sedikit berbeda dengan Persamaan (2.8.2) yang dijelaskan pada Bab 2. Hal ini terjadi karena pada metode Theil, disyaratkan bahwa nilai x_i tidak boleh ada yang sama untuk semua data.

Untuk mendapatkan nilai S dan $\hat{\tau}$, dilakukan langkah-langkah sebagai berikut:

- Menyusun pasangan-pasangan (x_i, y_i) dalam kolom menurut besarnya nilai-nilai x, dari nilai x yang paling kecil ke nilai x paling besar.
- Membandingkan masing-masing y_i dengan setiap y_j yang ada di bawahnya. Dalam melakukan pembandingan ini dapat dikatakan, bahwa suatu pasangan

nilai-nilai y berada dalam urutan yang serasi bila y yang di bawah lebih besar dari y yang di atasnya. Sedangkan jika y yang di bawah lebih kecil dari y yang di atasnya, maka dapat dikatakan bahwa suatu pasangan nilai-nilai y berada pada urutan tidak serasi.

- Menetapkan N_c sebagai banyaknya (y_i, y_j) yang berurutan serasi (dari kecil ke besar) dan menetapkan N_d sebagai banyaknya (y_i, y_j) yang berurutan tidak serasi (dari besar ke kecil).
- Hitung S sebagai selisih antara N_c dan N_d, yaitu $S = N_c - N_d$.

Kriteria pengambilan keputusan:

untuk suatu taraf uji α tertentu, H_0 akan ditolak jika $|\hat{\tau}| > \tau^*_{(n,\alpha/2)}$ dan tidak tolak H_0 jika $|\hat{\tau}| \leq \tau^*_{(n,\alpha/2)}$. Titik kritis $\tau^*_{(n,\alpha/2)}$ disajikan pada tabel statistik uji τ Kendall pada Lampiran 4.

3.3. Selang Kepercayaan Koefisien Kemiringan (*Slope*)

Metode pembentukan selang kepercayaan terhadap koefisien kemiringan dilandaskan pada prosedur pengujian hipotesis Theil untuk β_1, asumsi-asumsi yang mendasari prosedur pengujian hipotesis ini juga berlaku pada pembentukan selang kepercayaan (1-α)100% bagi β_1.

Selanjutnya [2] menjelaskan bahwa untuk menentukan selang kepercayaan bagi β_1, perlu ditentukan terlebih dahulu suatu konstanta k yang dirumuskan sebagai:

$$k = \frac{\binom{n}{2} - S_{(n,\alpha/2)} + 2}{2} \tag{3.11}$$

dengan k = konstanta untuk selang kepercayaan,

$\binom{n}{2}$ = banyaknya nilai b_{ij} yang mungkin dari n data,

$S_{(n,\alpha/2)}$ = titik kritis S dalam tabel statistik Tau Kendall untuk n data pada taraf α.

Selang kepercayaan bagi β_1 adalah $\tilde{\beta}_L < \beta_1 < \tilde{\beta}_U$ dengan $\tilde{\beta}_L$ adalah batas bawah selang kepercayaan untuk β_1 dan $\tilde{\beta}_U$ adalah batas atas selang kepercayaan untuk β_1. Misalkan $b_{(1)}, b_{(2)}, \ldots, b_{\binom{n}{2}}$ adalah nilai-nilai b_{ij} yang telah diurutkan dari terkecil sampai terbesar, maka batas bawah kepercayaan $\tilde{\beta}_L$ adalah nilai b_{ij} dalam urutan ke-k jika nilai-nilai b_{ij} tersebut disusun dari yang terkecil dan batas atas kepercayaan $\tilde{\beta}_U$ adalah nilai b_{ij} dalam urutan ke-k jika nilai-nilai b_{ij} diurutkan dari yang terbesar.

3.4. Contoh Penerapan Metode Theil

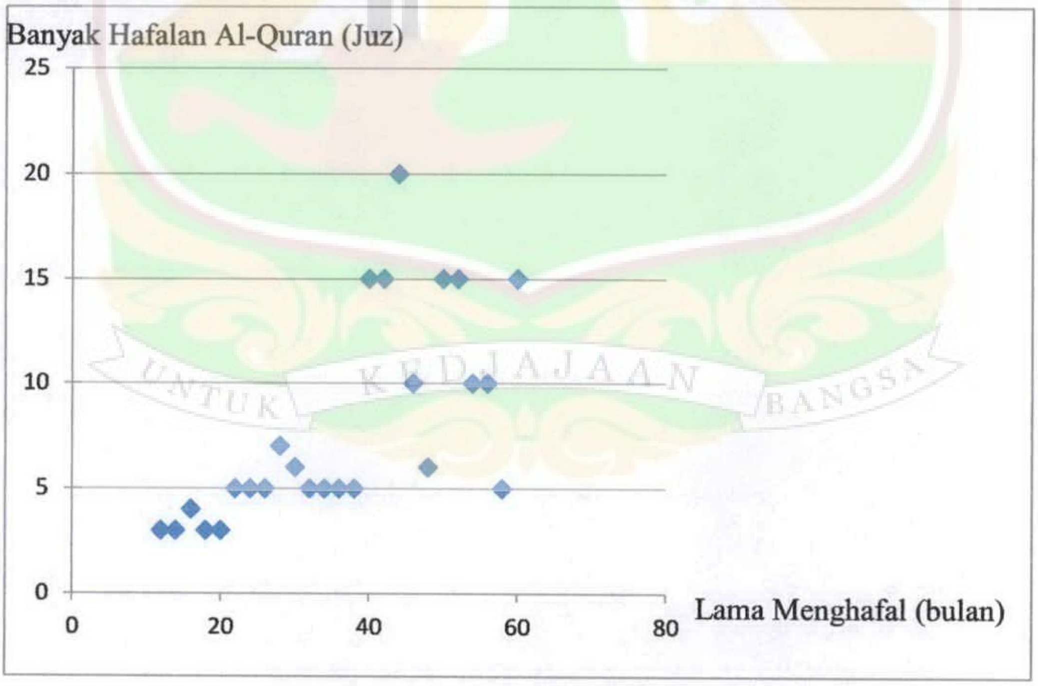
Untuk memberikan gambaran yang lebih jelas tentang metode yang telah dibahas, berikut ini akan disajikan contoh penerapannya. Data untuk ilustrasi ini diambil dari Yayasan Perguruan Islam AR-RISALAH Kecamatan Koto Tangah Kota Padang tentang lama waktu menghafal dan banyak hafalan Al-Quran 25 orang siswa. Data disajikan pada Tabel 3.4.1. Variabel x yang digunakan adalah lama menghafal dan variabel y yaitu banyak hafalan Al-Quran siswa.

Tabel 3.4.1 Data Lama Menghafal dan Banyak Hafalan Al-Quran Siswa

Siswa ke-	Lama Menghafal (bulan)	Banyak Hafalan Al-Quran (Juz)
1	12	3
2	14	3
3	16	4
4	18	3
5	20	3
6	22	5

7	24	5
8	26	5
9	28	7
10	30	6
11	32	5
12	34	5
13	36	5
14	38	5
15	40	15
16	42	15
17	44	20
18	46	10
19	48	6
20	50	15
21	52	15
22	54	10
23	56	10
24	58	5
25	60	15

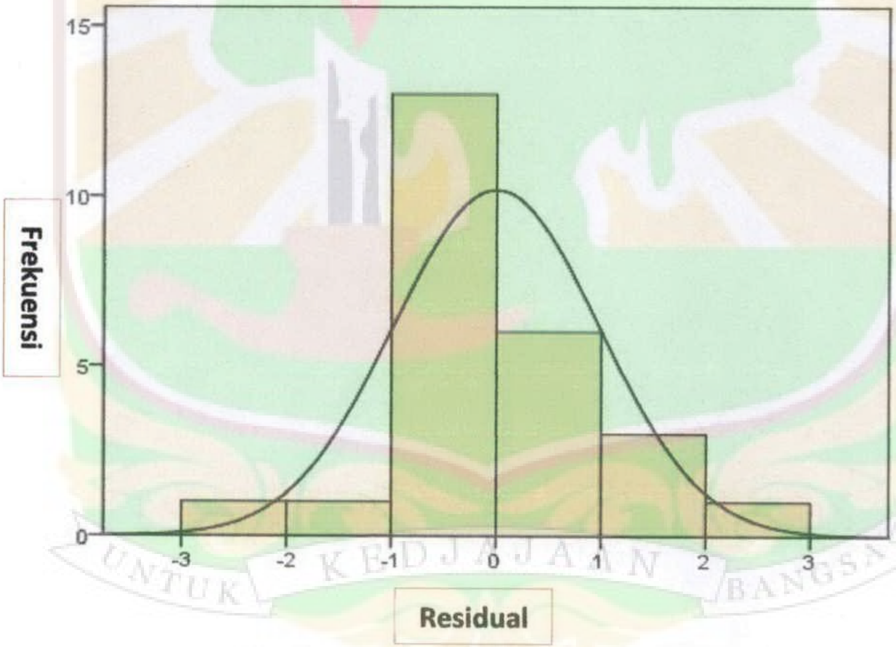
Pertama terlebih dahulu dibuat plot antara variabel x dengan variabel y seperti berikut:



Gambar 3.4.1 Plot Banyak Hafalan Al-Quran terhadap Lama Menghafal

Dari plot pada Gambar 3.4.1, diperkirakan bahwa model yang terbentuk antara variabel bebas x dan variabel terikat y adalah model linier. Selanjutnya karena metode Theil adalah metode statistika nonparametrik yang digunakan apabila kenormalan galat tidak terpenuhi, sehingga sebelum melakukan analisis regresi data terlebih dahulu dilakukan uji kenormalan terhadap residual. Dengan bantuan software SPSS diperoleh data residual dari data banyak hafalan Al-Quran siswa seperti yang terdapat pada Lampiran 1.

Secara umum untuk melihat apakah residual normal atau tidak, yang dapat dilakukan adalah membuat histogram dari data residual sebagaimana yang disajikan pada gambar berikut:



Gambar 3.4.2 Histogram Data Residual

Dari histogram pada Gambar 3.4.2, dapat dikatakan bahwa data residual tidak berdistribusi normal karena pola yang terbentuk tidak mengikuti kurva normal. Secara formal, dapat dilakukan pengujian kenormalan dengan uji *Kolmogorov-Smirnov* dari data residual. Hipotesis yang digunakan pada uji kenormalan adalah:

H_0 : Data residual berdistribusi normal,

H_1 : Data residual tidak berdistribusi normal.

Dengan bantuan *software* SPSS diperoleh output sebagai berikut:

Tabel 3.4.2 Uji Kenormalan residual

	Kolmogorov-Smirnov		
	Statistik	df	Signifikansi
Residual	0.175	25	0.046

Dapat dilihat pada Tabel 3.4.2 uji kenormalan di atas bahwa nilai signifikansi dari pengujian kenormalan menggunakan *Kolmogorov-Smirnov* nilainya lebih kecil dari 0,05. Hal ini berarti bahwa H_0 ditolak, sehingga dapat disimpulkan bahwa data residual tidak berdistribusi normal. Oleh karena itu, cukup alasan untuk menganalisis model regresi bagi data dengan metode Theil.

Dugaan model regresi linier sederhana untuk data adalah

$$\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_i.$$

Untuk mendapatkan nilai koefisien regresi dari model tersebut menggunakan metode Theil, pertama-tama nilai x diurutkan dari yang kecil ke yang besar seperti yang disajikan dalam Lampiran 2. Dengan $n = 25$, dugaan dari β_1 ditentukan dengan menghitung nilai b_{ij} sebanyak:

$$N = \binom{n}{2} = \frac{n(n-1)}{2} = \frac{25(25-1)}{2} = \frac{600}{2} = 300.$$

Selanjutnya akan dihitung nilai-nilai b_{ij} sebagai berikut:

$$b_{12} = \frac{(Y_2 - Y_1)}{(X_2 - X_1)} = \frac{3-3}{14-12} = 0{,}0000$$

$$b_{13} = \frac{(Y_3 - Y_1)}{(X_3 - X_1)} = \frac{4-3}{16-12} = 0{,}2500$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$b_{2425} = \frac{(Y_{25} - Y_{24})}{(X_{25} - X_{24})} = \frac{15-5}{60-58} = 5{,}0000$$

nilai b_{ij} yang telah diurutkan dapat dilihat pada Lampiran 3. Penduga bagi β_1 dinyatakan sebagai median dari nilai-nilai b_{ij}.

Karena $N = 300$, maka $\tilde{\beta}_1 = median\left(b_{ij}\right)$

$$= \frac{1}{2}\left(b_{\left(\frac{N}{2}\right)} + b_{\left(\frac{N}{2}+1\right)}\right)$$

$$= \frac{1}{2}\left(b_{\left(\frac{300}{2}\right)} + b_{\left(\frac{300}{2}+1\right)}\right)$$

$$= \frac{1}{2}\left(b_{(150)} + b_{(151)}\right)$$

$$= \frac{1}{2}(0{,}1667 + 0{,}1667)$$

$$= 0{,}1667.$$

Selanjutnya nilai $\tilde{\beta}_0$ diperoleh dengan menghitung terlebih dahulu nilai-nilai $a_i = y_i - \tilde{\beta}_1 x_i$, yaitu:

$a_1 = 3 - 0{,}1667(12) = 0{,}9996$

$a_2 = 3 - 0{,}1667(14) = 0{,}6662$

$a_3 = 4 - 0{,}1667(16) = 1{,}3328$

$a_4 = 3 - 0{,}1667(18) = -0{,}0006$

$a_5 = 3 - 0{,}1667(20) = -0{,}3340$

$a_6 = 5 - 0{,}1667(22) = 1{,}3326$

$a_7 = 5 - 0{,}1667(24) = 0{,}9992$

$a_8 = 5 - 0{,}1667(26) = 0{,}6814$

$a_9 = 7 - 0{,}1667(28) = 2{,}3324$

$a_{10} = 6 - 0{,}1667(30) = 0{,}9990$

$a_{11} = 5 - 0{,}1667(32) = -0{,}3344$

$a_{12} = 5 - 0{,}1667(34) = -0{,}6678$

$a_{13} = 5 - 0{,}1667(36) = -1{,}0012$

$a_{14} = 5 - 0{,}1667(38) = -1{,}3346$

$a_{15} = 15 - 0{,}1667(40) = 8{,}3320$

$a_{16} = 15 - 0{,}1667(42) = 7{,}9986$

$a_{17} = 20 - 0{,}1667(44) = 12{,}6652$

$a_{18} = 10 - 0{,}1667(46) = 2{,}3318$

$a_{19} = 6 - 0{,}1667(48) = -2{,}0016$

$a_{20} = 15 - 0{,}1667(50) = 6{,}6650$

$a_{21} = 15 - 0{,}1667(52) = 6{,}3316$

$a_{22} = 10 - 0{,}1667(54) = 0{,}9982$

$a_{23} = 10 - 0{,}1667(56) = 0{,}6648$

$a_{24} = 5 - 0{,}1667(58) = -4{,}6686$

$a_{25} = 15 - 0{,}1667(60) = 4{,}9980.$

Selanjutnya nilai-nilai a_i di atas diurutkan dari yang kecil ke yang besar, dan hasilnya dapat dilihat pada Lampiran 4. Penduga bagi β_0 dinyatakan sebagai median dari nilai-nilai a_i.

Karena $n = 25$, maka $\tilde{\beta}_0 = median\,(a_i)$

$$= a_{\left(\frac{1}{2}(n+1)\right)}$$

$$= a_{\left(\frac{1}{2}(25+1)\right)}$$

$$= a_{(13)}$$

$$= 0{,}9990.$$

Dengan menggunakan nilai-nilai $\tilde{\beta}_0$ dan $\tilde{\beta}_1$ di atas, jadi diperoleh dugaan model regresinya adalah:

$$\tilde{y}_i = 0{,}9990 + 0{,}1667x_i. \qquad (3.12)$$

Dari model yang diperoleh dapat disimpulkan bahwa setiap kenaikan 1 bulan lama menghafal, maka banyak hafalan Al-Quran siswa akan bertambah 0,1667 juz. Sedangkan koefisien intersep dari model tersebut tidak perlu diinterpretasikan karena data tidak memuat nilai nol. Model tersebut bisa digunakan untuk menyatakan hubungan antara variabel x dan variabel y jika terlebih dahulu dilakukan pengujian apakah koefisiennya berarti atau tidak dengan hipotesis sebagai berikut:

$H_0 : \beta_1 = 0$

$H_1 : \beta_1 \neq 0$

Karena ada nilai y yang sama, maka statistik uji yang digunakan adalah

$$\hat{t} = \frac{S}{\sqrt{\frac{1}{2}n(n-1)}\sqrt{\frac{1}{2}n(n-1) - T_y}}.$$

Untuk memperoleh S dan $\hat{\tau}$, pertama-tama susun data seperti dalam Tabel 3.4.3. Banyaknya pasangan y yang berurutan serasi terhadap setiap nilai y yang ada dibawahnya terlihat dalam kolom tiga dan banyak pasangan y yang berurutan tidak serasi terhadap setiap nilai y yang ada dibawahnya terlihat dalam kolom empat.

Tabel 3.4.3 Penyusunan Data untuk Menghitung S

x	y	Banyaknya Pasangan y yang Berurutan Serasi	Banyaknya Pasangan y yang Berurutan Tidak Serasi
12	3	21	0
14	3	21	0
16	4	20	2
18	3	20	0
20	3	20	0
22	5	12	0
24	5	12	0
26	5	12	0
28	7	9	7
30	6	9	5
32	5	10	0
34	5	10	0
36	5	10	0
38	5	10	0
40	15	1	5
42	15	1	5
44	20	0	8
46	10	3	2
48	6	5	1
50	15	0	3
52	15	0	3
54	10	1	1
56	10	1	1
58	5	1	0
60	15	0	0
Total		$N_c = 209$	$N_d = 43$

Tabel tersebut dapat dijelaskan sebagai berikut:

Kolom Pasangan y yang berurutan serasi:

- Pada baris pertama diketahui $y = 3$. Selanjutnya dicari berapa banyak data yang nilai y nya lebih besar dari 3 pada kolom y yang terletak pada baris di bawahnya, dapat dilihat yang lebih besar dari 3 adalah angka 4, 5, 6, 7, 10, 15 dan 20, sebanyak 21.
- Pada baris keempat diketahui $y = 3$. Selanjutnya dicari berapa banyak data yang nilai y nya lebih besar dari 3 pada kolom y yang terletak pada baris di bawahnya, angka 4 terletak pada baris di atasnya sehingga tidak dihitung, maka angka yang lebih besar dari 3 hanya angka 5, 6, 7, 10, 15 dan 20, yang banyaknya 20. Demikian seterusnya sampai selesai.

Kolom Pasangan y yang berurutan tidak serasi:

- Pada baris pertama diketahui $y = 3$, akan dihitung berapa banyak data yang lebih kecil dari 3 pada kolom y yang terletak pada baris di bawahnya, dapat dilihat tidak ada angka yang lebih kecil dari 3, sehingga banyaknya nol.
- Pada baris kesembilan diketahui $y = 7$, akan dihitung banyak data yang lebih kecil dari 7 Pada kolom y yang terletak pada baris di bawahnya, dapat dilihat yang lebih kecil dari 7 adalah angka 5 dan 6, yang banyaknya 7.
- Pada baris ketiga diketahui $y = 4$, selanjutnya dicari berapa banyak data yang nilai y nya lebih kecil dari 4 pada kolom y yang terletak pada baris di bawahnya, ada angka 3 yang terletak pada baris di atasnya sehingga tidak dihitung, maka angka yang lebih kecil dari 4 hanya angka 3 yang ada pada baris di bawahnya, yang banyaknya 2. Demikian seterusnya sampai selesai.

$$\tilde{y}_{12} = 0,9990 + 0,1667(34) = 6,6668$$

$$\tilde{y}_{13} = 0,9990 + 0,1667(36) = 7,0002$$

$$\tilde{y}_{14} = 0,9990 + 0,1667(38) = 7,3336$$

$$\tilde{y}_{15} = 0,9990 + 0,1667(40) = 7,6670$$

$$\tilde{y}_{16} = 0,9990 + 0,1667(42) = 8,0004$$

$$\tilde{y}_{17} = 0,9990 + 0,1667(44) = 8,3338$$

$$\tilde{y}_{18} = 0,9990 + 0,1667(46) = 8,6672$$

$$\tilde{y}_{19} = 0,9990 + 0,1667(48) = 9,0006$$

$$\tilde{y}_{20} = 0,9990 + 0,1667(50) = 9,3340$$

$$\tilde{y}_{21} = 0,9990 + 0,1667(52) = 9,6674$$

$$\tilde{y}_{22} = 0,9990 + 0,1667(54) = 10,0008$$

$$\tilde{y}_{23} = 0,9990 + 0,1667(56) = 10,3342$$

$$\tilde{y}_{24} = 0,9990 + 0,1667(58) = 10,6676$$

$$\tilde{y}_{25} = 0,9990 + 0,1667(60) = 11,0010.$$

Kemudian dicari:

$$\bar{y} = \frac{\sum y_i}{n} = \frac{3 + 3 + 4 + \cdots + 15}{25} = \frac{200}{25} = 8,$$

maka

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

$$R^2 = 1 - \frac{(3 - 2,9994)^2 + (3 - 3,3328)^2 + \cdots + (15 - 11,0010)^2}{(3 - 8)^2 + (3 - 8)^2 + \cdots + (15 - 8)^2}$$

$$R^2 = 1 - \frac{377,414}{598}$$

$$R^2 = 0,36887.$$

Jadi, diperoleh $R^2 = 0,36887$, artinya presentase total keragaman variabel terikat y (banyak hafalan Al-Quran siswa) yang dapat diterangkan oleh model regresi yang diperoleh pada Persamaan (3.12) adalah sebesar 36,89 % dan sisanya sebesar 63,11 % merupakan kontribusi dari faktor-faktor lain yang mempengaruhi banyak hafalan Al-Quran.

Dalam hal ini bukan berarti model regresi yang diperoleh tidak baik, namun ada variabel-variabel bebas lain selain lama menghafal yang tidak dimasukkan dalam penelitian ini yang mungkin memberikan kontribusi lebih besar dalam menentukan banyak hafalan Al-Quran siswa.

Jika prosedur regresi dilakukan menggunakan Metode Kuadrat Terkecil (MKT) terhadap data tersebut, dengan bantuan *software* SPSS yang outputnya dapat dilihat pada Lampiran 7, diperoleh nilai $R^2 = 41,6\%$. Hasil Metode Kuadrat terkecil tidak terlalu jauh berbeda dengan metode Theil. Jadi dapat disimpulkan metode Theil hampir seefesien Metode Kuadrat Terkecil.

BAB IV

KESIMPULAN

Berdasarkan pembahasan sebelumnya dapat diambil kesimpulan, metode Theil adalah metode regresi yang digunakan sebagai alternatif dari Metode Kuadrat Terkecil (MKT) pada kondisi asumsi kenormalan galat tidak terpenuhi. Metode ini digunakan hanya untuk menduga model regresi linier sederhana dengan nilai x_i yang berbeda untuk semua data. Dugaan model regresi pada metode ini adalah $\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_i$, dengan $\tilde{\beta}_1 = median\,(b_{ij})$ dan $\tilde{\beta}_0 = median\,(a_i)$, dimana $b_{ij} = \frac{y_j - y_i}{x_j - x_i}$ untuk satu pasangan $(x_i, y_i), (x_j, y_j)$, dengan $i < j$ dan $a_i = y_i - \tilde{\beta}_1 x_i$.

Dugaan model regresi yang diperoleh dari contoh penerapan metode Theil pada data banyak hafalan Al-Quran siswa adalah $\hat{y}_i = 0{,}9990 + 0{,}1667 x_i$, dengan $\hat{y}_i$ adalah nilai dugaan dari variabel terikat y_i (banyak hafalan Al-Quran siswa) dan x_i adalah variabel bebas (lama menghafal). Hasil pengujian menunjukkan bahwa pada taraf nyata 5 % lama menghafal disimpulkan berpengaruh terhadap banyak hafalan Al-Quran siswa. Kemudian selang kepercayaan 95% koefisien *slope* berada pada selang $0{,}0833 < \beta_1 < 0{,}3$, artinya dengan keyakinan 95% disimpulkan bahwa banyak hafalan Al-Quran siswa akan bertambah antara 0,0833 sampai 0,3 juz untuk setiap pertambahan 1 bulan lama menghafal.

DAFTAR PUSTAKA

[1] Conover, W.J. 1980. *Practical Nonparametric Statistics*. John Wiley and Sons, New York

[2] Daniel, W.W. 1989. *Statistika Nonparametrik Terapan*. Gramedia, Jakarta

[3] Draper, N.R dan H. Smith. 1992. *Analisis Regresi Terapan*. PT. Gramedia, Jakarta

[4] Gujarati, D. 1997. *Ekonometrika Dasar*. Erlangga, Jakarta

[5] Montgomery, P. 1991. *Introduction to Linear Regression Analysis*. New York

[6] Nachrowi, N.D. 2008. *Penggunaan Teknik Ekonometri*. PT. Raja Grafindo Persada, Jakarta

[7] Sembiring, R. K. 1995. *Analisis Regresi*. ITB, Bandung

[8] Siegel, S. 1985. *Statistika Nonparametrik untuk Ilmu-ilmu Sosial*. PT. Gramedia Pustaka Utama, Jakarta

[9] Soemartini. 2008. Regresi Linier Nonparametrik Melalui Metode Theil. http://resources.unpad.ac.id/unpadcontent/uploads/publikasi_dosen/THEIL%27S%20METHOD.pdf, diakses tanggal 28 Febrauri 2012

[10] Sprent, P. 1991. *Metode Statistik Nonparametrik Terapan*. Universitas Indonesia, Jakarta

[11] Walpole, R.E. 1995. *Pengantar Statistika*. PT. Gramedia Pustaka Utama, Jakarta

Lampiran 1. Data Residual yang Diperoleh dari SPSS

NO	Residu
1	0,40
2	-0,05
3	0,50
4	-0.95
5	-1,40
6	0,15
7	-0,30
8	-0,75
9	0,80
10	-0,65
11	-2,10
12	-2,55
13	-3,00
14	-3,45
15	6,10
16	5,65
17	10,20
18	-0,25
19	-4,70
20	3,85
21	3,40
22	-2,05
23	-2,50
24	-7,95
25	1,60

Lampiran 2. Nilai x Diurutkan dari yang Kecil ke yang Besar

No	x	y
1	12	3
2	14	3
3	16	4
4	18	3
5	20	3
6	22	5
7	24	5
8	26	5
9	28	7
10	30	6
11	32	5
12	34	5
13	36	5
14	38	5
15	40	15
16	42	15
17	44	20
18	46	10
19	48	6
20	50	15
21	52	15
22	54	10
23	56	10
24	58	5
25	60	15

Lampiran 3. Nilai b_{ij} yang Telah Diurutkan dari Terkecil sampai Terbesar

Urut	b_{ij}	Urut	b_{ij}	Urut	b_{ij}
1	-5,0000	36	-0,2500	71	0,0000
2	-3,5000	37	-0,2000	72	0,0000
3	-2,5000	38	-0,1667	73	0,0000
4	-2,5000	39	-0,1250	74	0,0000
5	-2,0000	40	-0,1000	75	0,0000
6	-1,6667	41	-0,0500	76	0,0000
7	-1,5000	42	-0,0357	77	0,0000
8	-1,2500	43	0,0000	78	0,0000
9	-1,2500	44	0,0000	79	0,0000
10	-1,2500	45	0,0000	80	0,0000
11	-1,2500	46	0,0000	81	0,0000
12	-1,2500	47	0,0000	82	0,0000
13	-1,1250	48	0,0000	83	0,0000
14	-1,0714	49	0,0000	84	0,0000
15	-1,0000	50	0,0000	85	0,0000
16	-0,8333	51	0,0000	86	0,0000
17	-0,8333	52	0,0000	87	0,0000
18	-0,8333	53	0,0000	88	0,0000
19	-0,8333	54	0,0000	89	0,0000
20	-0,6250	55	0,0000	90	0,0238
21	-0,6250	56	0,0000	91	0,0385
22	-0,5555	57	0,0000	92	0,0417
23	-0,5000	58	0,0000	93	0,0435
24	-0,5000	59	0,0000	94	0,0454
25	-0,5000	60	0,0000	95	0,0454
26	-0,5000	61	0,0000	96	0,0454
27	-0,4167	62	0,0000	97	0,0500
28	-0,4167	63	0,0000	98	0,0500
29	-0,3571	64	0,0000	99	0,0526
30	-0,3571	65	0,0000	100	0,0555
31	-0,3333	66	0,0000	101	0,0625
32	-0,3125	67	0,0000	102	0,0625
33	-0,3125	68	0,0000	103	0,0625
34	-0,2500	69	0,0000	104	0,0667
35	-0,2500	70	0,0000	105	0,0714

Lampiran 4. Nilai a_i yang Telah Diurutkan dari Terkecil sampai Terbesar

Urut	a_i
1	-4,6686
2	-2,0016
3	-1,3346
4	-1,0012
5	-0,6678
6	-0,3344
7	-0,3340
8	-0,0006
9	0,6648
10	0,6662
11	0,6814
12	0,9982
13	0,9990
14	0,9992
15	0,9996
16	1,3326
17	1,3328
18	2,3318
19	2,3324
20	4,9980
21	6,3316
22	6,6650
23	7,9986
24	8,3320
25	12,6652

Lampiran 5. Tabel Statistik τ Kendall

n	α									
	0.005		0.010		0.025		0.050		0.100	
	S	τ^*	S	τ^*	S	τ^*	S	τ^*	S	τ^*
4	8	1.000	8	1.000	8	1.000	6	1.000	6	1.000
5	12	1.000	10	1.000	10	1.000	8	.800	8	.800
6	15	1.000	13	.867	13	.867	11	.733	9	.600
7	19	.905	17	.810	15	.714	13	.619	11	.524
8	22	.786	20	.714	18	.643	16	.571	12	.429
9	26	.722	24	.667	20	.556	18	.500	14	.389
10	29	.644	27	.600	23	.511	21	.467	17	.378
11	33	.600	31	.564	27	.491	23	.418	19	.345
12	38	.576	36	.545	30	.455	26	.394	20	.303
13	44	.564	40	.513	34	.436	28	.359	24	.308
14	47	.516	43	.473	37	.407	33	.363	25	.275
15	53	.505	49	.467	41	.390	35	.333	29	.276
16	58	.483	52	.433	46	.383	38	.317	30	.250
17	64	.471	58	.426	50	.368	42	.309	34	.250
18	69	.451	63	.412	53	.346	45	.294	37	.242
19	75	.439	67	.392	57	.333	49	.287	39	.228
20	80	.421	72	.379	62	.326	52	.274	42	.221
21	86	.410	78	.371	66	.314	56	.267	44	.210
22	91	.394	83	.359	71	.307	61	.264	47	.203
23	99	.391	89	.352	75	.296	65	.257	51	.202
24	104	.377	94	.341	80	.290	68	.246	54	.196
25	110	.367	100	.333	86	.287	72	.240	58	.193
26	117	.360	107	.329	91	.280	77	.237	61	.188
27	125	.356	113	.322	95	.271	81	.231	63	.179
28	130	.344	118	.312	100	.265	86	.228	68	.180
29	138	.340	126	.310	106	.261	90	.222	70	.172
30	145	.333	131	.301	111	.255	95	.218	75	.172
31	151	.325	137	.295	117	.252.	99	.213	77	.166
32	160	.323	144	.290	122	.246	104	.210	82	.165
33	166	.314	152	.288	128	.242	108	.205	86	.163
34	175	.312	157	.280	133	.237	113	.201	89	.159
35	181	.304	165	.277	139	.234	117	.197	93	.156
36	190	.302	172	.273	146	.232	122	.194	96	.152
37	198	.297	178	.267	152	.228	128	.192	100	.150
38	205	.292	185	.263	157	.223	133	.189	105	.149
39	213	.287	193	.260	163	.220	139	.188	109	.147
40	222	.285	200	.256	170	.218	144	.185	112	.144

Sumber: Daniel, W.W. 1989. *Statistika Nonparametrik Terapan*. Gramedia, Jakarta

Lampiran 6. Tabel Nilai-nilai Kritis Statistik Uji *Kolmogorov-Smirnov*

n	α				
	0.20	0.15	0.10	0.05	0.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.214	.228	.244	.272	.326
30	.195	.208	.222	.248	.297
35	.181	.193	.210	.230	.275
$n > 35$	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Sumber: Siegel, Sidney. 1985. *Statistika Nonparametrik Untuk Ilmu-ilmu Sosial.* PT. Gramedia Pustaka Utama, Jakarta

Lampiran 7. Output SPSS untuk Metode Kuadrat Terkecil (MKT)

```
The regression equation is
banyak hafalan Al-Quran = - 0,10 + 0,225 lama menghafal


Predictor          Coef  SE Coef      T      P
Constant         -0,100    2,052  -0,05  0,962
lama menghafal  0,22500  0,05290   4,25  0,000


S = 3,81502    R-Sq = 44,0%    R-Sq(adj) = 41,6%
```

RIWAYAT HIDUP

Penulis bernama Aldila Sarti, dilahirkan di Padang pada tanggal 05 Juli 1992 dari pasangan Suardi Onyot dan Elmi. Penulis adalah anak kelima dari enam bersaudara. Penulis menamatkan pendidikan Sekolah Dasar di SDN 28 Padang pada tahun 2002, SMPN 26 Padang pada tahun 2005 dan SMA Negeri 8 Padang pada tahun 2008. Pada tahun yang sama, penulis diterima sebagai mahasiswi jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Andalas melalui jalur SPMB (Seleksi Penerimaan Mahasiswa Baru).

Selama menjadi mahasiswa di Universitas Andalas, penulis aktif dalam organisasi Forum Studi Islam (FSI) FMIPA Unand, organisasi Pengembangan Ilmu dan Kandungan Al-Quran (PIKA) Universitas Andalas, dan pengajar privat mata pelajaran Matematika dan seni baca Al-Quran. Penulis melaksanakan Kuliah Kerja Nyata (KKN) pada tahun 2011 di Jorong Balai Rupih, Kenagarian Simalanggang, Kecamatan Payakumbuh, Kabupaten Limapuluh Kota dalam rangka menyelesaikan salah satu mata kuliah wajib fakultas.