PENDETEKSIAN *OUTLIER* PADA REGRESI NONLINIER DENGAN METODE STATISTIK *LIKELIHOOD DISPLACEMENT* (LD)

SKRIPSI

Oleh:
SITI TABI'ATUL HASANAH
NIM. 08610045

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI (UIN) MAULANA MALIK IBRAHIM
MALANG
2012

PENDETEKSIAN *OUTLIER* PADA REGRESI NONLINIER DENGAN METODE STATISTIK *LIKELIHOOD DISPLACEMENT* (LD)

SKRIPSI

Diajukan Kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk Memenuhi Salah Satu Persyaratan Dalam
Memperoleh Gelar Sarjana Sains (S.Si)

Oleh:
SITI TABI'ATUL HASANAH
NIM. 08610045

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI (UIN) MAULANA MALIK IBRAHIM MALANG
2012

PENDETEKSIAN *OUTLIER* PADA REGRESI NONLINIER DENGAN METODE STATISTIK *LIKELIHOOD DISPLACEMENT* (LD)

SKRIPSI

Oleh:
SITI TABI'ATUL HASANAH
NIM. 08610045

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal: 12 Januari 2012

Pembimbing I,	Pembimbing II
Sri Harini, M.Si	Dr. Ahmad Barizi, MA
NIP. 19731014 200112 2 002	NIP. 19731212 199803 1 001

Mengetahui,
Ketua Jurusan Matematika

Abdussakir, M.Pd
NIP. 19751006 200312 1 001

PENDETEKSIAN *OUTLIER* PADA REGRESI NONLINIER DENGAN METODE STATISTIK *LIKELIHOOD DISPLACEMENT* (LD)

SKRIPSI

Oleh:
SITI TABI'ATUL HASANAH
NIM. 08610045

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Sains (S.Si)
Tanggal: 21 Januari 2012

Penguji Utama	: Drs. Turmudi, M.Si NIP. 19571005 198203 1 006	....
Ketua Penguji	: Abdul Aziz, M.Si NIP. 19760318 200604 1 002	....
Sekretaris Penguji	: Sri Harini, M.Si NIP. 19731014 200112 2 002	....
Anggota Penguji	: Dr. Ahmad Barizi, MA NIP. 19731212 199803 1 001	....

Mengesahkan,
Ketua Jurusan Matematika

Abdussakir, M.Pd
NIP. 19751006 200312 1 001

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama	: Siti Tabi'atul Hasanah
NIM	: 08610045
Jurusan	: Matematika
Fakultas	: Sains dan Teknologi

menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 15 Januari 2012
Yang Membuat Pernyataan,

Siti Tabi'atul Hasanah
NIM. 08610045

MOTTO

وَإِذْ تَأَذَّنَ رَبُّكُمْ لَئِن شَكَرْتُمْ لَأَزِيدَنَّكُمْ ۖ وَلَئِن كَفَرْتُمْ إِنَّ عَذَابِي لَشَدِيدٌ ﴿٧﴾

dan (ingatlah juga), tatkala Tuhanmu memaklumkan;
"Sesungguhnya jika kamu bersyukur, pasti Kami akan menambah (nikmat) kepadamu, dan jika kamu mengingkari (nikmat-Ku), Maka Sesungguhnya azab-Ku sangat pedih".

PERSEMBAHAN

Karya ini penulis persembahkan untuk................

Ibu, bapak, keluarga

serta teman-teman semua

KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Puji syukur kehadirat Allah SWT yang telah memberikan rahmat dan berkah-NYA kepada penulis, sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul *"***Pendeteksian *Outlier* pada Regresi Nonlinier dengan Metode Statistik *Likelihood Displacement* (LD)** *"* dengan baik.

Sholawat dan salam semoga tetap tercurahkan kepada Nabi Agung Muhammad SAW. yang telah membimbing dan membawa manusia dari zaman jahiliyah ke zaman yang diridhoi oleh Allah SWT. serta yang dinantikan syafaatnya di hari akhir nanti.

Penyusunan skripsi ini tidak luput dari bantuan, dukungan serta bimbingan dari berbagai pihak. Oleh sebab itu penulis ucapkan terima kasih banyak kepada berbagai pihak yang telah membantu, yaitu:

1. Prof. Dr. H. Imam Suprayogo, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Prof. Drs. Sutiman Bambang Sumitro, S.U, D. Sc, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Abdussakir, M.Pd, selaku Ketua Jurusan Matematika Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Sri Harini, M.Si, dan Dr. H. Ahmad Barizi, MA, selaku Dosen Pembimbing skripsi yang telah banyak memberikan bimbingan dan mengarahkan penulis dalam menyelesaikan penulisan skripsi.

5. Kepada segenap dosen dan karyawan di Jurusan Matematika yang telah memberikan dukungan serta ilmunya.
6. Kedua orang tua, kakak, adik dan seluruh keluarga yang senantiasa memberikan motivasi serta do'anya.
7. Sahabat-sahabat senasib seperjuangan mahasiswa Matematika 2008, terima kasih atas segala pengalaman berharga dan kenangan terindah saat menuntut ilmu bersama.
8. Sahabat-sahabat selama mengerjakan skripsi seperti Shofwan Ali Fauji, Nur Ngaini, Nurus Sakinah dan semua pihak yang tidak mungkin penulis sebut satu persatu, terima kasih atas keikhlasan bantuan moral dan spiritual yang sudah diberikan pada penulis.

Akhir kata, penulis minta maaf atas semua kesalahan yang disengaja ataupun tidak disengaja. Selain itu penulis juga mengharapkan kritik dan saran guna dalam penulisan skripsi. Semoga apa yang penulis paparkan dapat bermanfaat bagi penulis khususnya dan bagi pembaca pada umumnya. Amin

Wassalamu'alaikum Wr.Wb.

Malang, Januari 2012

Penulis.

DAFTAR ISI

DAFTAR TABEL

DAFTAR SIMBOL

Lambang khusus:

I	: matriks identitas
σ^2	: variansi
X^*	: matriks X yang elemennya adalah variabel acak
Y^*	: vektor Y yang elemennya adalah variabel acak
β^*	: vektor β yang elemennya parameter
$\hat{\beta}^*$	: estimator parameter β^*
$\hat{\sigma}^2$	: estimator parameter σ^2
T	: transpose
$Y^*_{A_k}$	: vektor Y^* yang diduga mengandung *outlier*
$Y^*_{[A_k]}$	: vektor Y^* dimana indeks A_k telah dihilangkan
$X^*_{A_k}$	: matriks X^* yang diduga mengandung *outlier*
$X^*_{[A_k]}$	: matriks X^* dimana indeks A_k telah dihilangkan
$\beta^*_{A_k}$	: vektor β^* yang diduga mengandung *outlier*
$\beta^*_{[A_k]}$	: vektor β^* dimana indeks A_k dihilangkan
$\sigma^2_{A_k}$	: variansi σ^2 yang diduga mengandung *outlier*
$\sigma^2_{[A_k]}$	: variansi σ^2 dimana indeks A_k dihilangkan
$\hat{\beta}^*_{[A_k]}$	: estimator parameter β^* yang diduga mengandung *outlier*
$\hat{\sigma}^2_{[A_k]}$	: estimator parameter σ^2 dimana indeks A_k dihilangkan
$\left(\hat{\sigma}^2_{[A_k]}\left(\hat{\beta}^*_{[A_k]}\right)\right)$	: estimator parameter σ^2 ketika β^* diestimasi $\hat{\beta}^*_{[A_k]}$
LD_{A_k}	: *likelihood displacement* yang mengandung *outlier*

$L(x_i, \dots, x_n; \theta)$: fungsi *likelihood*

$f(\theta; x_i, \dots, x_n)$: fungsi distribusi peluang

$f_{x_i}, \dots, x_n(\theta; x_i, \dots, x_n)$: funsgi padat peluang

N : distribusi normal

ABSTRAK

Hasanah, Siti Tabi'atul. 2012. **Pendeteksian *Outlier* pada Regresi Nonlinier dengan Metode Statistik *Likelihood Displacement (LD)***. Skripsi. Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing : (1) Sri Harini, M.Si (II) Dr. H. Ahmad Barizi, MA

Kata kunci: *outlier*, regresi nonlinier multiplikatif, *maximum likelihood estimation*, *likelihood displacement*

Outlier merupakan pengamatan yang jauh berbeda (ekstrim) dari data pengamatan lainnya, atau dapat diartikan data yang tidak mengikuti pola umum model. Adakalanya *outlier* memberikan informasi yang tidak dapat diberikan oleh data yang lainnya. Karena itulah outlier tidak boleh begitu saja dihilangkan. *Outlier* dapat juga merupakan pengamatan berpengaruh.

Banyak sekali metode yang dapat digunakan untuk mendeteki adanya *outlier*. Pada penelitian-penelitian sebelumnya pendeteksian *outlier* dilakukan pada regresi linier. Selanjutnya akan dikembangkan pendeteksian *outlier* pada regresi nonlinier. Regresi nonlinier disini dikhususkan pada regresi nonlinier multiplikatif. Untuk mendeteksi yaitu menggunakan metode statistik *likelihood displacement*. Metode statistik *likelihood displacement* disingkat *(LD)* adalah suatu metode untuk mendeteksi adanya outlier dengan cara menghilangkan data yang diduga *outlier*.

Untuk mengestimasi parameternya maka digunakan metode *maximum likelihood*, sehingga didapatkan hasil etimasi yang maksimal. Dengan metode LD diperoleh LD_{A_k}, yaitu *likelihood displacement* yang diduga mengandung *outlier*. Selanjutnya Keakuratan metode LD dalam mendeteksi adanya *outlier* ditunjukkan dengan cara membandingkan MSE dari LD dengan MSE dari regresi pada umumnya. Statistik uji yang digunakan adalah Λ. Hipotesis awal ditolak ketika $\chi^2_{hitung} > \chi^2_{tabel}$, sehingga terbukti LD_{A_k} adalah suatu *outlier*.

ABSTRACT

Hasanah, Siti Tabi'atul. 2012. **The Detection of Outliers in Nonlinear Regression Using Likelihood Displacement (LD) Statistical Methods**. Thesis. Mathematics programme Faculty of Science and Technology The State of Islamic University Maulana Malik Ibrahim Malang.
Promotor: (1) Sri Harini, M. Si
(II) Dr. H. Barizi Ahmad, MA

Outlier is an observation that much different (extreme) from the other observational data, or data can be interpreted that do not follow the general pattern of the model. Sometimes outliers provide information that can not be provided by other data. That's why outliers should not just be eliminated. Outliers can also be an influential observation.

There are so many methods that can be used to detect the presence of outliers. In previous studies done on outlier detection of linear regression. Next will be developed detection of outliers in nonlinear regression. Nonlinear regression here is devoted to multiplicative nonlinear regression. To detect is use of statistical method likelihood displacement. Statistical methods abbreviated likelihood displacement (LD) is a method to detect outliers by removing the suspected outlier data.

To estimate the parameters are used to the maximum likelihood method, so we get the estimate of the maximum. By using LD method is obtained LD_{A_k} i.e likelihood displacement is thought to contain outliers. Further accuracy of LD method in detecting the outliers are shown by comparing the MSE of LD with the MSE from the regression in general. Statistic test used is Λ. Initial hypothesis was rejected when $\chi^2_{count} > \chi^2_{table}$, proved so LD_{A_k} is an outlier.

Keywords: outlier, multiplicative nonlinear regression,
maximum likelihood estimation, likelihood displacement.

الملخّص

الحسنة، سيتي تبيعة ، ٢١٠٢. **عملية الموجودة أوتلير (*Outlier*) على الترتيب التأخّر غير العمودي بطريقة الإحصاء الممكن الكثير الإنتقالي (*Likelihood Displacement*)**. البحث الجامعي. الشعبة الرياضيات لكلية العلوم والتكنولوجية بجامعة الإسلامية الحكومية مولانا مالك إبراهيم مالانج. المشرفان: (١) سري هاريني الماجستير (٢) د. أحمد بارزي الحاج الماجستير

الكلمة الرئيسية : أوتلير، الترتيب التأخّر غير العمودي أكثر التطبيق، نهاية أكثر الممكنات، الممكن الكثير الإنتقالي الإحصائي

أوتلير هي التأمّل أن يبعد التفريق عن البيانات التأملية الأخرى أو الإضافة هذا هي البيانات التي لاتتّبعها أشكال الطِراز العامة. أما كانت أوتلير تعطي المعلومات التي ليس لها أن تعطي البيانات الأخري فلذالك أوتلير لاتمسح أصلا ولكن أوتلير هي من التأمّل التأثري.

كانت أكثر الطرق أن تستخدم في عملية الموجودة أوتلير، أما كانت في البحث قبله أن عملية الموجودة أوتلير تستخدمه بأوتلير (*Outlier*) على الترتيب التأخّر العمودي ثم هذه تنتمى بعملية الموجودة أوتلير على الترتيب التأخر غير العمودي وهنا تختص بالترتيب التأخر غير العمودي أكثر التطبيقي. وأما كانت عملية الموجودة تستخدم الممكن الكثير الإنتقالي الإحصائي (*Likelihood Displacement statistic*) وهي من الطرق التي تستخدمها أن توجد أوتلير بأن تمسح البيانات التي تظنّها الأوتلير.

لأن تقدّر ممثلة للمجتمع فاستخدمت الباحثة طريقة نهاية أكثر الممكنات حتى تحصل على التقدير الكامل. كما تحصل الممكن الكثير الإنتقالي على الممكن الكثير الإنتقاليك (وهي الممكن الكثير الإنتقالي تظن أن تشتمل أوتلير). ثم تصدق طريقة الممكن الكثير الإنتقالي في عملية الموجودة أوتلير تدل على التقابلي م س أ من الممكن الكثير الإنتقالي ب م س أ من الترتيب التأخر العام. والإختبار الإحصائي تستخدم Λ. فرضية أولى مردود إذا $x^2_{\text{الحساب}} > x^2_{\text{الجدول}}$ فحُقِّقت الممكن الكثير الإنتقاليك هي من الأوتلير.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Al-Qur'an adalah kitabullah yang di dalamnya memuat segala sesuatu yang berhubungan dengan kehidupan. Al-Qur'an tidak hanya membahas tentang masalah agama saja, akan tetapi juga membahas tentang masalah sosial, sains dan masalah-masalah yang lainnya. Dalam kitab Al-Qur'an telah dijelaskan tentang masalah *outlier* (pencilan), yaitu terdapat dalam surat Al-Jinn ayat 14 yang berbunyi:

وَأَنَّا مِنَّا ٱلْمُسْلِمُونَ وَمِنَّا ٱلْقَـٰسِطُونَ ۖ فَمَنْ أَسْلَمَ فَأُو۟لَـٰٓئِكَ تَحَرَّوْا۟ رَشَدًا ﴿١٤﴾

Artinya: *"Dan sesungguhnya di antara kami ada orang-orang yang taat dan ada pula orang-orang yang menyimpang dari kebenaran. Barangsiapa yang taat, maka mereka itu benar-benar telah memilih jalan yang lurus*" (Qs. Al-Jinn/72:14).

Ayat di atas merupakan salah satu ayat yang menjelaskan tentang *outlier*. Dalam ayat tersebut dijelaskan bahwa "*di antara kami ada orang-orang yang taat dan ada yang menyimpang"*. Yang dimaksud dengan kata "*kami"* di sini adalah para jin. Seperti yang telah dijelasakan pada ayat di atas bahwa ada jin yang taat dan ada juga yang menyimpang. Kemudian para jin ini dalam statistika dianalogikan seperti suatu data yang mana jin itu ada yang taat (sesuai model) dan ada pula yang menyimpang (*Outlier*). Sama halnya dengan suatu pengamatan, ada kalanya dalam suatu pengamtan akan terjadi penyimpangan (*outlier*).

Outlier adalah pengamatan yang jauh berbeda (ekstrim) dari data pengamatan lainnya. Salah satu penyebab terjadinya *outlier* adalah kesalahan pada pengambilan data sehingga menyebabkan data tersebut menjadi ekstrim. Adakalanya *outlier* ini tidak boleh begitu saja dihilangkan, namun dalam hal ini harus hati-hati karena terkadang *outlier* itu memberikan informasi yang tidak dapat diberikan oleh titik pengamatan lain, misalnya karena adanya kombinasi keadaan yang tidak biasa dan perlu diadakan penyelidikan lebih jauh. Suatu *outlier* dapat dibuang setelah ditelusuri ternyata pengamatan tersebut merupakan akibat dari kesalahan pengukuran atau kesalahan dalam menyiapkan pengukuran. *Outlier* dapat juga merupakan pengamatan berpengaruh. *Outlier* yang bukan pengamatan berpengaruh, tidak memiliki pengaruh yang kuat pada model kecuali *outlier* tersebut sangat besar. Tetapi jika *outlier* merupakan data berpengaruh, maka akan memberikan dampak pada model (Drapper dan Smith, 1992:146).

Misalkan saja pada suatu penelitian tentang sapi penghasil susu. Dari suatu data ternyata diperoleh ada beberapa sapi yang menghasilkan hasil susu yang lebih banyak dari biasanya atau dari sapi normalnya. Sapi penghasil susu yang tidak sesuai dengan normalnya merupakan suatu *outlier*, namun jika mengahapus begitu saja data ini berarti telah menghilangkan bibit sapi unggul yang mampu menghasilkan banyak susu sapi. Oleh sebab itulah penting untuk mengidentifikasi adanya *outlier* agar tidak kehilangan suatu data yang memiliki kualitas yang bagus. Jika dengan adanya *outlier* itu kurang baikmaka perlu diidentifikasi dan kemudian dihilangkan data yang mengandung *outlier*.

Mendeteksi adanya *outlier* dalam suatu pengamatan, harus punya asumsi-asumsi untuk menduga peubah mana yang mengadung *outlier*. Penduganan ini sering disebut dengan istilah estimasi. Dalam ayat Al-Qur'an telah dijelaskan mengenai masalah estimasi yaitu terdapat dalam surat As-Shaffat ayat 147, yaitu:

وَأَرْسَلْنَٰهُ إِلَىٰ مِاْئَةِ أَلْفٍ أَوْ يَزِيدُونَ ﴿١٤٧﴾

Artinya:"*Dan kami utus dia kepada seratus ribu atau lebih* "(Qs. As-Shaffat/ 37:147).

Ayat di atas mengandung unsur ketidakpastian yaitu terdapat pada kata "*seratus ribu atau lebih"*, maksudnya adalah Allah mengutus Nabi Yunus kepada seratus ribu orang atau lebih. Sehingga jika diteliti lebih lanjut berapa detailnya kaum Nabi Yunus tidak diketahui. Sehingga hal ini bisa dinyatakan dengan suatu estimasi.

Banyak sekali metode yang dapat digunakan untuk mendeteksi *outlier*, salah satunya yaitu pendektesian *outlier* pada model linier univariat telah dikemukaan oleh Cook dengan memperkenalkan Jarak Cook (*Cook's Distance*) sebagai ukuran untuk mendeteksi pengamatan berpengaruh dalam model linier univariat. Ukuran Jarak Cook ini dirumuskan sebagai kombinasi dari *studential residual*, variansi *residual*, dan variansi nilai prediksi. Selain metode yang dikemukakan oleh Cook, masih banyak lagi metode yang digunakan untuk pendeteksian *outlier* pada model linier (Makkulau, 2010:95)

Xu, Abraham dan Steiner (2005) mengembangkan Jarak Cook univariat untuk mendeteksi *outlier* pada model linier multivariat (model regresi linier multivariat) dengan menggunakan metode statistik *likelihood* *displacement* yang disingkat LD. Metode LD adalah suatu metode untuk mendeteksi adanya *outlier* dengan cara menghilangkan pengamatan yang diduga *outlier* (Makkulau, 2010:95).

Untuk selanjutnya pada penelitian ini, peneliti akan mengembangkan pendeteksian *outlier* pada model nonlinier dengan menggunkan metode yang sama yaitu metode statistik *likelihood displacement*. Oleh karena itu, dalam penelitian ini akan dikaji tentang **"*Pendeteksian Outlier pada Regresi NonLinier dengan Metode Statistik Likelihood Displacement* (LD)".**

1.2 Rumusan Masalah

Sebagaimana uraian pada latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah bagaimana mendeteksi *outlier* pada regresi nonlinier dengan metode statistik *Likelihood Displacement* (LD)?

1.3 Tujuan Penelitian

Dari uraian rumusan masalah di atas, maka tujuan dari penelitian ini adalah untuk mengetahui cara mendeteksi *outlier* pada regresi nonlinier dengan metode statistik *Likelihood Displacement* (LD).

1.4 Batasan Masalah

Dalam penelitian ini penulis membatasi permasalahan sebagai berikut:

1. *Outlier* yang diteliti yaitu *outlier* dari variabel terikat (Y) pada regresi nonlinier.
2. Regresi nonlinier yang diteliti adalah regresi nonlinier multiplikatif.
3. *Error* diasumsikan berdistribusi normal $\varepsilon \sim N(0, \sigma^2)$.

4. Estimasi parameter β dan σ^2 menggunakan *maximum likelihood estimation.*

1.5 Manfaat Penelitian

1. Penulis

   Manfaat bagi penulis adalah untuk meningkatkan pemahaman penulis mengenai metode-metode yang dapat digunakan untuk mendeteksi *outlier* khususnya *outlier* pada regresi nonlinier dengan metode statistik *likelihood displacement* (LD).

2. Mahasiswa Matematika

   Penelitian ini diharapkan menjadi bahan bacaan dan referensi sekaligus sebagai acuan untuk memahami tentang *outlier* dan sekaligus mengembangkan metode-metode yang dapat digunakan untuk mendeteksi *outlier*, sehingga dapat meningkatkan pengetahuan tentang metode untuk mendeteksi adanya *outlier*.

3. UIN Maulana Malik Ibrahim Malang

   Penelitian ini diharapkan dapat menjadi bahan bacaan referensi bagi pembaca dan peneliti lainnya untuk memahami langkah-langkah menentukan *outlier* dan metode yang digunakannya.

4. Pembaca

   Sebagai tambahan wawasan pengetahuan pembaca tentang *outlier* dan cara pendeteksian *outlier* khususnya pada regresi nonlinier multiplikatif.

1.6 Metode Penelitian

Dalam penelitian ini, penulis menggunakan metode studi literatur. Studi literatur yaitu penelitian yang dalam menunjukkan penelitiannya dilakukan dengan cara mendalami, memahami, menelaah dan mengindentifikasi pengetahuan yang ada dalam kepustakaan. Sumber kajian pustaka dapat berupa jurnal penelitian, disertasi, tesis, skripsi, laporan penelitian atau diskusi-diskusi ilmiah.

Adapun langkah-langkah untuk mendeteksi *outlier* dengan metode statistik *likelihood displacement (LD)* adalah sebagai berikut:

1. Menentukan persamaan regresi nonlinier multiplikatif yang mengandung *outlier*.
2. Mengumpulkan variabel terikat (Y) yang diduga mengandung *outlier*.
3. Mendeteksi *outlier* pada regresi nonlinier multiplikatif dengan asumsi bahwa *error* berdistribusi normal dimulia dengan membuat fungsi *Likelihood.*
4. Mengestimasi parameter regresi nonlinier multiplikatif dengan metode *maximum likelihood estimation.*
5. Memaksimumkan fungsi *likelihood* dengan cara melogaritmanaturalkan dan mendefferensialkan kemudian menetukan estimasi parameter regresi nonlinier multiplikatif ketika variabel terikat (Y) yang diduga mengandung *outlier* dihilangkan.
6. Menentukan fungsi *Likelihood Displacement*.
7. Menentukan statistik uji *outlier* .

8. Membuat kesimpulan sekaligus jawaban dari permasalahan yang dikemukakan pada rumusan masalah.

1.7 Sistematika Penulisan

Untuk memudahkan melihat dan memahami penelitian ini secara keseluruhan, maka penulis menggunakan sistematika penulisan yang terdiri dari empat bab, yaitu :

BAB I : PENDAHULUAN

Pada bab ini berisi tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, metode penelitian dan sistematika penulisan.

BAB II : KAJIAN PUSTAKA

Pada bab ini penulis menjelaskan tentang teori-teori yang berkaitan dengan *Pendeteksian Outlier Pada Regresi Nonlinier Dengan Metode Statistik Likelihood Displacement (LD).*

BAB III : PEMBAHASAN

Pada bab ini penulis menjelaskan tentang hasil penelitian, yaitu tahap-tahap mendeteksi *outlier* pada regresi nonlinier multiplikatif dengan metode LD.

BAB IV : PENUTUP

Pada bab ini penulis memberikan kesimpulan dari pembahasan dan saran-saran yang sesuai untuk penelitian selanjutnya.

BAB II

KAJIAN PUSTAKA

Berikut ini merupakan teori-teori yang berkaitan dengan pendeteksian *outlier* pada regresi nonlinier dengan menggunakan metode statistik *likelihood displacement* (LD).

2.1 *Outlier*

Definisi 1:

Secara umum *outlier* dapat diartikan data yang tidak mengikuti pola umum pada model atau data yang keluar dari model dan tidak berada dalam daerah selang kepercayaan (Sembiring, 1995:62).

Menurut Draper dan Smith (1992:146) sisaan yang merupakan *outlier* adalah yang nilai mutlaknya jauh lebih besar dari pada sisaan lainnya dan terletak tiga atau empat kali simpangan baku atau lebih jauh lagi dari rata-rata sisaannya. *Outlier* merupakan suatu keganjilan dan menandakan suatu titik data yang sama sekali tidak tipikal dibandingkan data lainnya.

2.2 Estimasi Parameter

Dengan statistika populasi berusaha untuk disimpulkan. Untuk ini kelakuan populasi dipelajari berdasarkan data yang diambil baik secara sampling ataupun sensus. Dalam kenyataannya, mengingat berbagai faktor, untuk keperluan tersebut diambil sebuah sampel yang representatif, kemudian berdasarkan pada hasil analisis terhadap data sampel, kesimpulan mengenai populasi dibuat. Kelakuan populasi yang akan ditinjau disini hanyalah mengenai parameter populasi dan sampel yang digunakan adalah sampel acak. Data sampel dianalisis, nilai-nilai yang perlu, yaitu statistik hitung, dan dari nilai-nilai statistik ini disimpulkan bagaimana parameter bertingkah laku. Cara pengambilan kesimpulan tentang parameter yang pertama kali akan dipelajari ialah sehubungan dengan cara-cara mengestimasi harga parameter (Sudjana, 2005:198).

Menurut Yitnosumarto (1990:211) penduga (*estimator*) adalah anggota peubah acak statistik yang mungkin untuk sebuah parameter (anggota peubah yang diturunkkan). Parameter adalah nilai yang mengikuti acuan keterangan atau informasi yang dapat menjelaskan batas-batas atau bagian-bagian tertentu dari suatu sistem persamaan.

Murray dan Larry (1999:166) menyatakan terdapat dua jenis estimasi parameter, yaitu:

1. Estimasi titik

   Estimasi dari sebuah parameter populasi yang dinyatakan oleh bilangan tunggal disebut sebagai estimasi titik dari parameter tersebut. Sebuah nilai yang diperoleh dari sampel dan digunakan sebagai estimasi dari parameter yang nilainya tidak diketahui. Misalkan $X1, X2, \dots, Xn$ merupakan sampel acak berukuran n dari X, maka statistik yang berkaitan dengan θ dinamakan estimasi dari θ. Setelah sampel diambil, nilai-nilai yang dihitung dari sampel itu digunakan sebagai taksiran titik bagi θ.

2. Estimasi Interval

   Estimasi dari parameter populasi yang dinyatakan dengan dua bilangan. Di antara posisi parameternya diperkirakan berbeda, sehinggga disebut estimasi interval. Estimasi interval mengindikasikan adanya tingkat kepresisian atau akurasi dari sebuah estimasi sehingga estimasi interval akan dianggap semakin baik jika mendekati estimasi titik.

Adapun sifat-sifat estimasi titik adalah sebagai berikut:

1. Tak Bias

Yusuf Wibisono (2005:362) dalam bukunya menyatakan bahwa *estimator* tak bias bagi parameter θ, jika

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

dan dikatakan estimator bias bagi parameter θ, jika

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

namun *estimator* bias dapat diubah menjadi *estimator* tak bias jika ruas kanan dikalikan atau ditambahkan dengan konstanta tertentu.

2. Konsisten

Damodar N. Gujarati (2007:98) menerangkan *estimator* parameter $\hat{\theta}$ dikatakan konsisten bila nilai-nilainya mendekati nilai parameter yang sebenarnya meskipun ukuran sampelnya semakin besar.

Suatu statistik $\hat{\theta}$ disebut estimator yang konsisten untuk parameter θ jika dan hanya jika $\hat{\theta}$ konvergen dalam probabilitas ke parameter θ atau $plim\hat{\theta} = \theta$.

Jika $\hat{\theta}_n$ adalah penaksir untuk θ yang didasarkan pada sampel acak berukuran n, maka $\hat{\theta}_n$ dikatakan konsisten bagi parameter θ, jika $lim_{n\to\infty} P\ (|\hat{\theta}_n - \theta| < \varepsilon = 1$. Penentuan penaksir konsisten ini dapat dilakukan dengan menggunakan ketidaksamaan Chebyshev's, $lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < k\sigma \geq 1 - \frac{1}{k^2}\right.$.

3. Efisien

Jika distribusi sampling dari dua statistik memiliki *mean* atau ekspektasi yang sama, maka statistik dengan variansi 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.

2.3 Distribusi

2.3.1 Distribusi Normal

Distribusi normal merupakan model distribusi peluang yang paling banyak digunakan dalam statistika.

Definisi 2:

Suatu variabel acak X berdistribusi normal $N(\mu, \sigma^2)$ bila untuk suatu $\sigma^2 > 0$ dan $-\infty < \mu < \infty$ (Turmudi dan Harini, 2008:204).

Menurut Sudjana (2005:136) jika variabel acak kontinu X mempunyai fungsi densitas pada $X = x$ dengan persamaan:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (2.1)$$

dengan:

π : bilangan irasional (3,1416)

e : bilangan natural (2,7183)

μ : parameter untuk rata-rata distribusi

σ : simpangan baku untuk distribusi

dan nilai x mempunyai batas $-\infty < \mu < \infty$ maka dikatakan bahwa variabel acak X berdistribusi normal.

Distribusi normal ini mempunyai *mean* μ dan variansi σ^2. Sedangkan distribusi normal dengan rataan 0 dan variansi 1 disebut distribusi normal baku yang dilambangkan dengan $N(0,1)$ (Sembiring, 1995:4-5).

2.3.2 Distribusi *Chi-Square*

Distribusi *chi-square* merupakan distribusi dengan variabel acak kontinue. Simbol untuk *chi-square* adalah χ^2. Distribusi *chi-square* sebenarnya merupakan jumlah kuadrat dari variabel-variabel acak yang bebas dan menyebar normal dengan *mean* 0 dan ragam $1(Z \sim NID(0,1))$. Distribusi ini dapat dinyatakan dengan

$$X^2 = Z_1^2 + Z_2^2 + \cdots + Z_k^2$$

$$\chi^2 = \sum_i Z_i^2 \sum_i \left(\frac{Y_i - \mu_i}{\sigma_i}\right)^2$$

merupakan variabel acak yang tersebar menurut distribusi *chi-square* dengan derajat bebas sebesar k dan dapat dituliskan

$$X^2 \approx \chi_k^2$$

dimana χ_k^2 yaitu distribusi *chi-square* dengan derajat bebas k.

Definisi 3:

Suatu variabel acak X berdistribusi *chi-square* dengan derajat bebas k, dinyatakan dengan $\chi_k^2(0)$ bila untuk suatu bilangan bulat $k > 0$. (Turmudi dan Harini, 2008: 210)

Distribusi ini mempunyai fungsi kepekatan peluang sebagai berikut:

$$f_x(x) = \begin{cases} \dfrac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\left(\frac{k}{2}\right)-1} e^{-\frac{x}{2}}, & x \geq 0 \\ 0, & \text{selainnya} \end{cases} \tag{2.2}$$

Nilai tengah (*mean*) dan ragam untuk distribusi χ^2 adalah $\mu = k$ dan $s^2 = 2k$. Distribusi *chi-square* bergantung pada banyaknya simpangan baku yang bebas antara satu dengan yang lain atau dengan kata lain bergantung pada derajat bebasnya.

2.3.3 Distribusi Peluang Gabungan

Definisi 4:

Jika X dan Y variabel acak, maka peluang terjadinya X dan Y secara serentak dinyatakan sebagai $f(x,y)$ disebut Distribusi Peluang Gabungan untuk setiap pasangan (x,y) (Herrhyanto, 2009:5).

Jika X dan Y adalah dua varibel acak diskrit, maka distribusi peluang terjadinya secara serentak atau bersamaan dinyatakan dengan fungsi $f(x,y)$ dan disebut sebagai distribusi peluang gabungan X dan Y. Fungsi $f(x,y)$ dikatakan distribusi peluang gabungan variabel acak diskrtit X dan Y bila memenuhi:

1. $f(x,y) \geq 0$ untuk semua (x,y)
2. $\sum_x \sum_y f(x,y) = 1$
3. $P[(X,Y) \in A] = \sum_A \sum f(x,y)$ untuk setiap daerah A di bidang xy

Jika X dan Y adalah dua variabel acak kontinu, maka distribusi peluang terjadinya secara serentak atau bersamaan dinyatakan dengan fungsi $f(x,y)$ dan disebut sebagai distribusi peluang gabungan X dan Y. Fungsi $f(x,y)$ dikatakan fungsi peluang atau distribusi peluang gabungan variabel acak kontinu X dan Y bila memenuhi:

1. $f(x,y) \geq 0$ untuk semua (x,y)
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dxdy = 1$
3. $P[(X,Y) \in A] = \int \int f(x,y)\, dxdy$

2.4 Regresi Nonlinier

2.4.1 Pengertian Regresi Nonlinier

Analisis regresi merupakan analisis yang menyangkut studi tentang hubungan antara satu variabel yang disebut variabel terikat atau variabel yang dijelaskan dan satu atau lebih variabel yang lain yang disebut variabel bebas atau variabel penjelas (Gujarati, 2007:115).

Sedangkan tujuan dari analisis regresi adalah untuk memprediksi besarnya variabel terikat (Y) yang dipengaruhi oleh variabel bebas (X) (Hasan, 2002:115).

Definisi 5:

Regresi yang variabel-variabelnya berbentuk tidak biasa. Bentuk grafik regresi nonlinier adalah berupa lekungan (Hasan, 2002:297).

Sedangkan menurut Supranto (1994:262) hubungan fungsi antara dua variabel X dan Y tidak selalu bersifat linier, akan tetapi juga bukan linier (nonlinier). Diagram pencar dari hubungan yang linier akan menunjukkan suatu pola yang dapat didekati dengan garis lurus, sedangkan yang bukan linier harus didekati dengan garis lengkung.

Sedangkan menurut Sugiarto (1992:29) hubungan fungsi di antara dua peubah X dan Y dikatakan tidak linier apabila laju perubahan dalam Y yang berhubungan dengan perubahan satu satuan X tidak konstan untuk satu jangkauan nilai-nilai X tertentu.

2.4.2 Macam-macam Regresi Nonlinier

Model regresi nonlinier dapat digolongkan menjadi dua yaitu model *linier intrinsik* dan model *nonlinier intrinsik*. Jika suatu model dikatakan model *linier intrinsik*, maka model model ini dapat dinyatakan dalam bentuk linier baku dengan mentransformasikan secara tepat terhadap variabelnya. Jika suatu model nonlinier tidak dapat dinyatakan dalam bentuk baku, berarti model ini secara intsinsik adalah nonlinier. Berikut ini adalah beberapa model yang dapat dinyatakan dalam linier baku (Draper dan Smith, 1992:213).

1. **Bentuk Perkalian (Multiplicative)**

Model ini dinyatakan sebagai berikut

$$Y = \propto X_1^{\beta} X_2^{\delta} X_3^{\delta} \varepsilon \qquad (2.2)$$

dimana σ, β, dan δ adalah parameter yang tidak diketahui, dan ε adalah *error* acak yang bersifat multiplikatif. Dengan mengalgoritmakan basis e pada pada persamaan di atas, maka model persamaan di atas menjadi $ln\ Y = ln\,\alpha + \beta\, ln\, X_1 + \gamma\, ln\, X_2 + \delta\, ln\, X_3 + ln\,\varepsilon.$

Model persamaan tersebut menjadi bentuk linier sehingga dapat ditangani dengan prosedur regresi linier. Model tersebut merupakan model linier dalam bentuk $ln\,\varepsilon$. ε tidak berdistribusi normal, sebab yang berdistribusi normal adalah $ln\,\varepsilon$ (Draper dan Smith, 1992:213).

2. **Bentuk Eksponensial**

Berikut ini adalah model atau bentuk umum dari regresi nonlinier eksponensial,

$$Y_i = e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}}\ \varepsilon_i \qquad (2.3)$$

transformasi juga dapat dijalankan dengan mengambil transformasi logaritmanya:

$$ln(Y_i) = ln(e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}}\ \varepsilon_i)$$

$$ln(Y_i) = ln\ e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik}} + ln\,\varepsilon_i$$

$$ln(Y_i) = (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik})\, ln\, e + ln\,\varepsilon_i$$

$$ln(Y_i) = (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik})(1) + ln\,\varepsilon_i$$

$$ln\, Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + ln\, \varepsilon_i \tag{2.3}$$

Model seperti ini adalah model linier dalam bentuk semi log.

3. Bentuk Berkebalikan (Resiprokal)

Bentuk dari model resiprokal secara umumnya adalah sebagai berikut:

$$Y_i = \frac{1}{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i} \tag{2.4}$$

atau transformasi modelnya adalah

$$\frac{1}{Y_i} = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i \tag{2.5}$$

Dua model yang secara intrinsik adalah non linier sebagai berikut (Draper dan Smith, 1992:215).

$$Y = \beta_0 + \beta_1 e^{-(\beta_2 X)} + \varepsilon \tag{2.6}$$

dan

$$Y = \beta_0 + \beta_1 X + \beta_2 (\beta_3)^X + \varepsilon. \tag{2.8}$$

2.5 Regresi dalam Pendekatan Matriks

Model regresi yang paling sederhana adalah model regresi linier. Model regresi linier sederhana terdiri dari satu variabel. Model tersebut dapat digeneralisasikan menjadi lebih dari satu atau dalam k variabel. Persamaan model regresi linier dengan k variabel adalah sebagai berikut:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{2.9}$$

Bila pengamatan mengenai $y, x_1, x_2, \dots, x_k$ dinyatakan masing-masing dengan $y_i, x_{i1}, x_{i2}, \dots, x_{ik}$ dan *error*nya ε_i, maka persamaan (2.9) dapat dituliskan sebagai:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_k + \varepsilon_i, i = 1,2,..,n. \quad (2.10)$$

Dinotasikan dalam bentuk matriks, sehingga menjadi:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}. \quad (2.11)$$

Misalkan

$$\underline{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \quad X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix} \quad \underline{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \quad \underline{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Persamaan (2.11) dapat dinyatakan sebagai:

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon} \quad (2.12)$$

dengan:

$\underline{Y}$: vektor respon n x 1

X : matriks peubah bebas berukuran n x $(k + 1)$

$\underline{\beta}$: vektor parameter berukuran $(k + 1)$ x 1

$\underline{\varepsilon}$: vektor galat ukuran n x 1

(Sembiring,1995:134-135)

2.6 *Maximum Likelihood*

Statistik inferensia dapat dibagi dalam dua bagian besar, estimasi dan pengujian hipotesis. Kedua inferensi tersebut masing-masing bertujuan untuk membuat pendugaan dan pengujian suatu parameter populasi dan informasi sampel yang diambil dari populasi tersebut. Gujarati N. Damodar (2010:131) menjelaskan bahwa metode dari estimasi titik (*point estimation*) dengan sifat-sifat teoritis yang lebih kuat dari pada metode OLS adalah metode *maximum likelihood* (ML).

Definisi 6:

Fungsi *likelihood* dari n variabel acak $x_1, x_2, \dots, x_n$ didefinisikan sebagai fungsi kepadatan bersama dari n variabel acak. Fungsi kepadatan bersama $f_{x_1}, \dots_{x_n}(x_1, \dots, x_n; \theta)$, yang mempertimbangkan fungsi dari θ. Jika $x_1, \dots, x_n$ adalah sampel acak dari fungsi kepadatan $f(x, \theta)$, maka fungsi *likelihood*nya adalah $f(x_1; \theta) f(x_2; \theta) \dots f(x_n; \theta)$ (Mood, Graybill and Boes, 1986:278).

Maximum likelihood dapat diperoleh dengan menentukan turunan dari L terhadap parameternya dan menyatakannya sama dengan nol. Dalam hal ini, akan lebih mudah untuk terlebih dahulu menghitung logaritma kemudian menentukan turunannya. Dengan cara ini diperoleh:

$$\frac{1}{f(X_1, \theta)} \frac{\partial f(x_1, \theta)}{\partial \theta} + \dots + \frac{1}{f(X_n, \theta)} \frac{\partial f(x_n, \theta)}{\partial \theta} = 0$$

Penyelesaian dari persamaan ini, untuk θ dalam bentuk x_k, dikenal sebagai estimator *maximum likelihood* dari θ.

Contoh:

Jika $x_1, x_2, \ldots, x_n$ adalah sampel acak dari distribusi $x \sim N(0{,}1)$. Fungsi *likelihood*nya adalah:

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

karena berdistribusi normal, maka $f(x;\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\theta)^2}$ sehingga fungsi *likelihood*nya adalah

$$\begin{aligned}
L(x_1, x_2, \ldots, x_n; \theta) &= f(x_1;\theta) f(x_2;\theta) \ldots f(x_n;\theta) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_1-\theta)^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_2-\theta)^2} \ldots \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_n-\theta)^2} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}((x_1-\theta)^2+(x_2-\theta)^2+(x_n-\theta)^2)} \\
&= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_n-\theta)^2} \\
&= \left(\frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}\right)^n e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_n-\theta)^2} \\
&= \left(\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}}\right) e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_n-\theta)^2}
\end{aligned}$$

Sehingga fungsi *likelihood*nya dapat ditulis sebagai berikut:

$$L(\theta; x_1, x_2, \ldots, x_n) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_n-\theta)^2}$$

2.7 Metode Statistik *Likelihood Displacement* (LD)

Metode LD adalah suatu metode yang dikembangkan dengan cara menghilangkan pengamatan yang diduga *outlier*. Misalkan k adalah pengamatan dikumpulkan pada pengamatan tertentu, dengan k diduga sebagai *outlier*. Indeks A_k adalah kumpulan dari k yang diduga *outlier*.

Definisi 7:

LD dari pengamatan yang mengandung *outlier* untuk $\hat{\beta}^*$ dengan variansi $\hat{\sigma}^2$ adalah: $LD_{A_k}\ (\hat{\beta}^*|\hat{\sigma}^2) = 2\left(ln\,L(\hat{\beta}^*,\hat{\sigma}^2) - ln\,L\left(\hat{\beta}^*_{[A_k]},\hat{\sigma}^2(\hat{\beta}^*_{[A_k]})\right)\right)$

dimana $\hat{\sigma}^2(\hat{\beta}^*_{[A_k]}$ adalah MLE dari $\hat{\sigma}^2$ ketika $\hat{\beta}^*$ diestimasi oleh $\hat{\beta}^*_{[A_k]}$ (Makkulau, dkk, 2010:97).

$\hat{\beta}^*$dan $\hat{\sigma}^2$ merupakan MLE dari β^* dan σ^2 dengan semua pengamatan dan $\hat{\beta}^*_{[A_k]}$ dan $\hat{\sigma}^2_{[A_k]}$ bahwa tanpa k pengamatan dalam himpunan A_k. Definisi ini adalah analogi secara langsung untuk definisi *Likelihood Displacement* yang digunakan oleh Cook's dan Weisberg (1982) untuk mempertimbangkan pengaruh dari observasi tunggal pada estimasi parameter pada kasus univariat.

Pada kasus khusus yaitu $\theta_1 = (\beta_1^*, \sigma_1^2)$ subset dari $\theta = (\beta^*, \sigma^2)$, maka fungsi *likelihood displacement* dapat dimodifikasi menjadi

$$LD_{A_k}((\beta_1^*, \sigma_1^2)|(\beta_2^*, \sigma_2^2))$$

$$= 2\{ln\ L(\hat{\beta}^*, \hat{\sigma}^2) - \ln L((\hat{\beta}_1^*, \hat{\sigma}_1^2)_{[A_k]}\ ,(\hat{\beta}_2^*, \hat{\sigma}_2^2)(\hat{\beta}_1^*, \hat{\sigma}_1^2)_{[A_k]})\} \quad (2.13)$$

dimana $L((\hat{\beta}_2^*, \hat{\sigma}_2^2)(\hat{\beta}_1^*, \hat{\sigma}_1^2)_{[A_k]}) = max_{(\beta_2^*, \sigma_2^2)} L\left((\hat{\beta}_1^*, \hat{\sigma}_1^2)_{[A_k]}\ ,(\hat{\beta}_2^*, \hat{\sigma}_2^2)\right)$ adalah memaksimumkan fungsi *log likelihood* pada parameter (β_2^*, σ_2^2) dengan $(\beta_1^*, \sigma_1^2) = (\hat{\beta}_1^*, \hat{\sigma}_1^2)_{[A_k]}$ maka $\beta_1^* = \hat{\beta}_{1[A_k]}^*$ dan $\sigma_1^2 = \hat{\sigma}_{1[A_k]}^2$ adalah *maximum* *likelihood estimation* dari (β_1^*, σ_1^2) ketika pengamatan k dihilangkan (Xu, Abraham, dan Steiner, 1998:5-6).

2.8 Tafsir Al-Qur'an dan Hadits

2.8.1 Tafsir Al-Qur'an dan Hadits tentang Outlier

Surat Al-Jinn adalah surat ke-72 dari kitab Al-qur'an. Surat Al-Jinn termasuk dalam surat Makkiyah. Pada ayat 14 dari surat Al-Jinn terdapat permasalahan yang sama dengan permasalahan yang ada dalam statistika, yaitu mengkaji tentang adanya *Outlier* dalam sekumpulan data, adapun bunyi surat tesebut yaitu:

وَأَنَّا مِنَّا ٱلْمُسْلِمُونَ وَمِنَّا ٱلْقَٰسِطُونَ ۖ فَمَنْ أَسْلَمَ فَأُو۟لَٰٓئِكَ تَحَرَّوْا۟ رَشَدًا ﴿١٤﴾

Artinya: *"Dan sesungguhnya diantara kami ada orang-orang yang taat dan ada pula orang-orang yang menyimpang dari kebenaran. Barangsiapa yang taat, maka mereka itu benar-benar telah memilih jalan yang lurus"*(Qs. Al-Jinn/72 :14).

Asal mula turunnya surat Al-Jinn ayat 14 yaitu untuk menampik dugaan bahwa semua jin baik yang mendengar langsung ayat-ayat Al-Qur'an maupun yang belum atau tidak mendengarnya kesemuanya telah patuh kepada Allah. Kemudian pada ayat tersebut diterangkan bahwa "*dan sesungguhnya diantara kami*" masyarakat jin "*ada orang-orang taat*" yakni mereka yang benar-benar taat dan penuh kepatuhan kepada Allah "*dan ada pula orang-orang penyimpang*" yakni mereka yang telah sangat jauh dari dari kebenaran lagi sagat mantap kekufurannya. "*Barang siapa yang patuh*" maka mereka itu telah bersungguh-sungguh memilih arah yang mengantar ke jalan kebenaran.

Kata penyimpangan dalam surat di atas pada konsep statistika dapat diartikan sebagai *outlier*. Sebab suatu *outlier* dikatakan sebagai penyimpangan dilihat dari pengertiannya, yaitu:

1. *Outlier* adalah yang nilai mutlaknya jauh lebih besar dari pada sisaan-sisaan lainnya dan bisa jadi terletak tiga atau empat simpangan baku atau lebih jauh dari rata-rata sisaanya.
2. *Outlier* adalah suatu keganjilan dan menandakan suatu titik data yang sama sekali tidak tipikal dibandingkan data yang lainnya. (Drapper dan Smith, 1992:146)
3. *Outlier* adalah data yang tidak mengikuti pola umum model. (Sembiring, 1995:62)

Dari penafsiran ayat ini dijelaskan bahwa para penyimpang yakni mereka yang telah sangat jauh dari kebenaran. Penyimpangan ini mempunyai arti yang sama dengan *outlier* yaitu sama-sama terletak sangat jauh diantara data dalam model tersebut.

Sedangkan menurut *Tafsir Ibnu Katsir Jilid* 8 (2007:310) dijelaskan bahwa diantara hamba-hamba Allah yang hidup di alam semesta ini adalah ada yang muslim ada juga yang melakukan penyimpangan. Maksudnya disini adalah mereka penyimpangan terhadap kebenaran Allah. Berarti mereka jauh dari kebenaran-kebenaran Allah.

2.8.2 Tafsir Al-Qur’an dan Hadits tentang Estimasi

Surat As-Shaffat ayat 147 menjelaskan mengenai estimasi parameter. Ayat tersebut berbunyi sebagai berikut:

وَأَرۡسَلۡنَٰهُ إِلَىٰ مِاْئَةِ أَلۡفٍ أَوۡ يَزِيدُونَ ١٤٧

Artinya:“*Dan kami utus dia kepada seratus ribu atau lebih”* (Qs. As-Shaffat/ 37:147).

Sebab diturunkannya ayat tersebut adalah ketika Nabi Yunus mulai sembuh dari penderitaannya, badannya sudah mulai segar, Allah SWT mengutus kembali kepada kaumnya yang pada waktu itu jumlahnya seratus ribu orang atau lebih. Kedatangan Nabi Yunus as. disambut dengan baik dan mereka beriman kepadanya. Sesungguhnya mereka telah menyadari bahwa mereka dahulunya telah melakukan kesalahan sehingga Nabi Yunus as. pergi meninggalkan mereka. Bilamana mereka tidak beriman dan mematuhinya, tentulah mereka akan ditimpa azab seperti halnya kaum-kaum yang dahulu yang mengingkari Nabi-nabinya. Ketika Nabi Yunus kembali ke tengah-tengah mereka dan mengajak mereka ke jalan yang benar, beriman kepada Allah dan RasulNya, mereka segera menerimanya dengan penuh ketaatan. Karena itu, Allah SWT menganugerahkan kenikmatan kepada mereka dengan hidup bahagia, aman sentosa sampai ajal mereka.

Para ulama’ mempunyai penafsiran yang berbeda. Perbedaan para ulama’ dalam menafsirkan dapat dilihat dalam tabel berikut:

Tabel 2.1. Tafsir Surat As-Shaffat Ayat 147 Menurut Para Ulama’ Tafsir

No.	Ulama	Tafsir	Penafsiran ulama’
1.	M. Quraish Shihab	Tafsir Al-Misbah (2003:84)	Kata (أَو) auw pada firman-Nya: (أَوۡيَزِيدُونَ)yang berarti *bahkan* atau *dan*, maka itu berarti beliau (Yunus) diutus kepada dua kelompok yaitu: 100.000 diutus

			pada orang yahudi penduduk Nainawa dan 20.000 selain orang Yahudi penduduk Nainawa, sedangkan Kata (أَو) auw pada firman-Nya: (أَويَزِيدُونَ) yang berarti *atau*, maka nabi Yunus disuga diutus kepada 100.000 orang atau lebih
2.	Imam Jalaluddin Al-Mahally dan ImamJalaluddin As-Suyuthi	Tafsir Jalalain (1990:1946)	أَويَزِيدُونَ (*atau lebih dari itu*) yakni diduga lebih dari dua puluh atau tiga puluh atau tujuh puluh ribu orang
3.	Abdulmalik Abdulkarim Amrullah (HAMKA)	Tafsir Al-Azhar (1981:194)	Menceritakan bahwa setelah Nabi Yunus sehat dan kuat kembali, dia diperintahkan Tuhan melaksanakan perintah yang dipikulkan kepadanya, yaitu mendatangi dan melakukan dakwah kepada kaumnya di negeri Ninive ini, yang berjumlah 100.000 orang atau lebih, artinya lebih dari 100.000, kurang tidak.

Tabel (2.1) telah menjelaskan secara rinci perbedaan pendapat para ulama' dalam menafsirkan surat As-Shaffat ayat 147. Dari tabel di atas dapat dijelaskan bahwa para ulama memperkirakan jumlah umat Nabi Yunus dengan jumlah yang berbeda-beda, meskipun demikian tidak ada yang mengatakan kurang dari 100.000 orang. Dari ketiga penafsiran di atas dapat disimpulkan bahwa terdapat suatu penggunaan istilah estimasi pada surat Ash-Shaffat ayat 147.

Selain ayat surat Ash-Shaffat ayat 147, ayat lain yang mengkaji tentang estimasi yaitu Al-Qur'an surat Al-Jatsiyah ayat 24 yang bunyinya adalah sebagai berikut:

وَقَالُواْ مَا هِيَ إِلَّا حَيَاتُنَا ٱلدُّنۡيَا نَمُوتُ وَنَحۡيَا وَمَا يُهۡلِكُنَآ إِلَّا ٱلدَّهۡرُۚ
وَمَا لَهُم بِذَٰلِكَ مِنۡ عِلۡمٍۖ إِنۡ هُمۡ إِلَّا يَظُنُّونَ ٢٤

Artinya:*"Dan mereka berkata," kehidupan ini tidak lain hanyalah kehidupan didunia saja, kita mati dan kita hidup dan tidak ada yag membinasakan kita selain masa", dan mereka sekali-kali tidak mengetahui pengetahuan tentang itu, mereka tidak lain adalah menduga-duga"* (Qs. Al-Jatsiyah/ 45:24).

Allah menyebutkan bahwa orang-orang musyrik menjadikan hawa nafsu mereka sebagai Tuhan mereka, dan Allah telah menyesatkan mereka karena Dia mengetahui hal ihwal mereka dan memasang tutup pada penglihatan mereka, maka Dia sebutkan di sini salah satu kebodohan mereka. Yaitu bahwa mereka mengingkari kebangkitan, bahkan mengatakan yang ada itu adalah kehidupan dunia saja, kita mati dan kita hidup, dan tidak ada yang membinasakan kita selain masa. Perbuatan mereka yang seperti itu tidak ada sandarannya. Mereka tidak mendapatkan hujjah yang dapat mereka ucapkan kecuali kata-kata mereka.

Menyatakan bahwa kehidupan ini hanya kehidupan dunia saja, dan yang membinasakan mereka adalah masa, mereka tidaklah mempunyai ilmu yang didasarkan pada akal maupun naql (kitab). Mereka hanyalah menyangka, membuat perkiraan saja tanpa adanya hujjah yang jitu yang dapat mereka jadikan pegangan.

Kemudian Allah mengutus Rasul-Nya supaya memberikan jawaban kepada mereka bahwa Allah yang menghidupkan bapak-bapak mereka, kemudian mematikan mereka, kemudian membangkitkan mereka di hari yang tidak diragukan.

Dari penafsiran ayat di atas dijelaskan bahwa " mereka hanyalah menyangka, membuat perkiraan" yang dimaksud di sini adalah sama dengan mengenai estimasi parameter yaitu hanya dapat menduga saja dan tidak diketahui seberapa tepat pendugaannya (Al-Maraghi, 1989:276-279).

Abdussyakir (2007:155-156) mengatakan bahwa estimasi adalah ketrampilan untuk menentukan sesuatu tanpa melakukan proses perhitungan secara eksak. Dalam matematika terdapat tiga jenis estimasi, yaitu estimasi banyak, estimasi pengukuran, dan estimasi komputasional.

1. Estimasi banyak

   Estimasi banyak adalah menentukan banyaknya objek tanpa menghitung secara eksak. Objek di sini maknanya sangat luas. Objek dapat berupa orang, uang, kelereng, titik, dan mobil. Estimasi yang terdapat dalam surat Ass-Shaffat, 37:147 adalah estimasi banyak orang.

2. Estimasi pengukuran

   Estimasi pengukuran adalah menentukan ukuran sesuatu tanpa menghitung secara eksak. Ukuran di sini maknanya sangat luas. Ukuran dapat bermakna ukuran waktu, panjang, luas, dan volume. Ketika melihat orang berjalan tanpa menanyakan tanggal lahirnya, pemabaca dapat menebak, menaksir usianya.

3. Estimasi komputasional

Estimasi komputasional adalah menentukan hasil operasi hitung tanpa menghitungnya secara eksak. Seseorang mungkin akan menghitung dengan cara membulatkan kepuluhan terdekat.

Dalam surat Al-Baqoroh ayat 80 yang bunyinya:

وَقَالُوا لَن تَمَسَّنَا ٱلنَّارُ إِلَّآ أَيَّامًا مَّعۡدُودَةً ۚ قُلۡ أَتَّخَذۡتُمۡ عِندَ ٱللَّهِ عَهۡدًا
فَلَن يُخۡلِفَ ٱللَّهُ عَهۡدَهُۥٓ ۖ أَمۡ تَقُولُونَ عَلَى ٱللَّهِ مَا لَا تَعۡلَمُونَ ﴿٨٠﴾

Artinya*: “Dan mereka berkata: "Kami sekali-kali tidak akan disentuh oleh api neraka, kecuali selama beberapa hari saja." Katakanlah: "Sudahkah kamu menerima janji dari Allah sehingga Allah tidak akan memungkiri janji-Nya, ataukah kamu Hanya mengatakan terhadap Allah apa yang tidak kamu ketahui?:* (Qs. Al-Baqoroh/ 1:80).

Ayat ini juga membahas masalah estimasi. Makna estimasi yaitu terletak pada potongan ayat yang artinya “*selama beberapa hari saja”* . dari ayat ini tidak diketahui dengan tepat berapa hari akan tetapi hanya bisa mengestimasinya.

Hadits Shahih Bukari no 21888 berikut merupakan hadits yang berhubungan dengan estimasi, hadits tersebut berbunyi:

عَنْ مَا لِكٍ نَا فِعٍ عَنْ ابْنِ عُمَرَ عَنْ زَيْدِ بْنِ ثَا بِتٍ رَضِيَ اللهُ عَنْهُمْ أَنَّ رَسُولُ اللهِ صَلَّى
اللهُ عَلَيْهِ وَسَلَّمَ أَ رْخَصَ لِصَا حِبِ الْعَرِيَّةِ أَن يَبِيعَهَا بِخَرْ صِهَا

Artinya:*“Dari Malik, dari Nafi’, dari Ibnu Umar, dari Zaid bin Tsabit ra, “sesungguhnya Rasulullah SAW memberi keringanan kepada mereka yang mempunyai ariyah untuk menjualnya dengan kira-kira (ditaksir)”*(HR. Shohih Bukhari: 2188).

Para ulama salaf berbeda pendapat. Apakah anggur atau selainnya masuk kategori kurma dalam hal *ariyah*?. Sebagian mengatakan tidak diikutkan dari madzhab Azh-Zhahiri. Sebagian pendapat lagi mengatakan diikutkan, pendapat ini adalah pendapat yang masyhur dalam madzhab Syafi'i. Ada yang berpendapat bahwa semua buah-buahan dan semua yang dapat disimpan lama dapat diikutkan di dalamnya, ini adalah pendapat madzhab Maliki (Al-Asqalani, 2007:312).

Dari potongan hadits di atas yaitu أَن يَبِيعَهَا بِخَرْ صِهَا "*untuk dijual sesuai taksirannya*". Ath-Thabrani menambahkan dari Ali bin Abdul Aziz. Dari Al-Qa'nabi (guru Imam Bukhari dalam riwayat ini), كَيْلاً "*berdasarkan takaran*". Imam Muslim juga meriwayatkan dari Yahya bin Yahya, dari Malik بِخَرْصِهَا مِنَ التَّمْرِ "*berdasarkan taksiran setelah menjadi kurma kering*". Imam Muslim juga meriwayatkan hal serupa dari Sulaiman bin Bilal, dari Yahya bin Sa'ad dengan lafazh

رَخَّصَ فِي العَرِيَةِ يَأْخُذُ هَا أَهْلُ البَيْتِ بِخَرْصِهَا تَمْرًا يَأْكُلُو نَهَا رَطْبًا

"*memberi keringanan dalam jual beli Ariyah, diambil oleh penghuni rumah berdasarkan taksirannya setelah menjadi kuma kering yang mana mereka memakannya dalam keadaan masih basah*"

Yahya berkata, "*Ariyah* adalah seorang membeli kurma kering dan menukarnya dengan kurma basah miliknya dengan memperkirakan atau menaksir berapa banyak jumlahnya setelah kering."

Dari konsep hadits jual beli *Ariyah*, yang mana membeli kurma yang kering kemudian ditaksir dengan kurma basah yang dimilikinya, hal ini mengandung konsep estimasi yaitu suatu perkiraan tentang harga kurma kering dibeli dengan kurma basah dengan jalan memperkirakan banyaknya kurma basah tersebut ketika sudah kering.

Konsep taksiran ini sama halnya dengan konsep taksiran yang ada dalam statistika, yang mana dalam menaksir suatu parameter berarti mengestimasi nilai parameter tersebut. Jika hasil dari estimasi tersebut diaplikasikan dalam kehidupan nyata nilai yang sesungguhnya, maka nilai taksiran tersebut adalah mendekati nilai sebenarnya atau berkisar di sekitar nilai tersebut. Konsep-konsep tentang *outlier* dan estimasi ini sudah termaktup dalam Al-Qur'an dan Hadits.

BAB III

PEMBAHASAN

Dalam bab ini akan dibahas tentang pendeteksian *outlier* pada regresi nonlinier dengan metode statistik *Likelihood Displacement (LD).*

3.1 Regresi Nonlinier Multiplikatif

Bentuk umum dari regresi nonlinier multiplikatif adalah dinyatakan sebagai berikut:

$$y_i = \beta_0 x_{i1}^{\beta_1} x_{i2}^{\beta_2} x_{i3}^{\beta_3} \dots x_{nk}^{\beta_k} \dots \varepsilon_i \qquad (3.1)$$

Persamaan (3.1) dapat dilinierkan dengan melogaritmanaturalkan persamaannya, sehingga modelnya menjadi:

$$ln(y_i) = ln\left(\beta_0 x_{1i}^{\beta_1} x_{2i}^{\beta_2} x_{3i}^{\beta_3} \dots x_{nk}^{\beta_k} \dots \varepsilon_i\right)$$

$$ln(y_i) = ln(\beta_0) + ln\left(x_{1i}^{\beta_1}\right) + ln\left(x_{2i}^{\beta_2}\right) + \cdots + ln\left(x_{nk}^{\beta_k}\right) + \cdots + ln(\varepsilon_i)$$

$$ln\, y_i = ln\, \beta_0 + \beta_1\, ln\, x_{1i} + \beta_2\, ln\, x_{2i} + \cdots + \beta_k\, ln\, x_{nk} + \cdots + ln\, \varepsilon_i \qquad (3.2)$$

dengan $i = 1,2,\dots,n$ dan $k = 1,2,\dots,n$

Dalam penelitian ini diasumsikan bahwa variabel terikat $(ln\, y)$ berdistribusi normal dengan *mean* μ dan variansi σ^2. Sehingga dalam persamaan (3.1) ε berdistribusi log normal, karena yang berdistribusi normal adalah $\ln \varepsilon$.

Dengan menggunakan pendekatan matriks, maka persamaan (3.2) dapat dituliskan sebagai berikut:

$$\begin{pmatrix} \ln y_1 \\ \ln y_2 \\ \ln y_3 \\ \vdots \\ \ln y_n \end{pmatrix} = \begin{pmatrix} 1 & \ln x_{11} & \ln x_{12} & \ln x_{13} & \cdots & \ln x_{1k} \\ 1 & \ln x_{21} & \ln x_{22} & \ln x_{23} & \cdots & \ln x_{2k} \\ 1 & \ln x_{31} & \ln x_{32} & \ln x_{33} & \cdots & \ln x_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln x_{n1} & \ln x_{n2} & \ln x_{n3} & \cdots & \ln x_{nk} \end{pmatrix} \begin{pmatrix} \ln \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} \ln \varepsilon_1 \\ \ln \varepsilon_2 \\ \ln \varepsilon_3 \\ \vdots \\ \ln \varepsilon_n \end{pmatrix}.$$

Misalkan :

$$Y^* = \begin{pmatrix} \ln y_1 \\ \ln y_2 \\ \ln y_3 \\ \vdots \\ \ln y_n \end{pmatrix} \qquad X^* = \begin{pmatrix} 1 & \ln x_{11} & \ln x_{12} & \ln x_{13} & \cdots & \ln x_{1k} \\ 1 & \ln x_{21} & \ln x_{22} & \ln x_{23} & \cdots & \ln x_{2k} \\ 1 & \ln x_{31} & \ln x_{32} & \ln x_{33} & \cdots & \ln x_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \ln x_{n1} & \ln x_{n2} & \ln x_{n3} & \cdots & \ln x_{nk} \end{pmatrix}$$

$$\beta^* = \begin{pmatrix} \ln \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} \qquad \varepsilon^* = \begin{pmatrix} \ln \varepsilon_1 \\ \ln \varepsilon_2 \\ \ln \varepsilon_3 \\ \vdots \\ \ln \varepsilon_n \end{pmatrix}$$

maka bentuk linier dari regresi nonlinier multiplikatif dengan pendekatan matriks adalah:

$$Y^*_{n\,x\,1} = X^*_{n\,x\,(k+1)} \beta^*_{(k+1)\,x\,1} + \varepsilon^*_{n\,x\,1} \tag{3.4}$$

3.2 Estimasi Parameter Regresi Nonlinier Multiplikatif

Untuk menentukan estimasi parameter dari regresi multiplikatif dengan pendekatan matriks, yaitu menggunakan persamaan (3.4)

$Y^* = X^* \beta^* + \varepsilon^*$.

Dari persamaan (3.4) diketahui bahwa $Y^* = (\ln y_1, \ln y_2, \dots, \ln y_n)^T$ adalah variabel acak, karena diasumsikan berdistribusi normal, maka $Y^* \sim N(X^*\beta^*, I\sigma^2)$ dengan $X^* = (\ln x_{0i}, \ln x_{1i}, \dots, \ln x_{1k})$ dan

$\beta^* = (ln\,\beta_0\,,\beta_1,\dots,\beta_n)^T$ dimana $i = 1,2,\dots,n$ dan I adalah matriks identitas. Sehingga fungsi distribusi peluang gabungannya adalah:

$$f(Y^*|\beta^*,\sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(Y^*-X^*\beta^*)^T(Y^*-X^*\beta^*)} \tag{3.5}$$

Untuk menentukan estimasi parameter menggunakan *Maximum Likelihood Estimation*, maka terlebih dahulu menentukan fungsi *likelihood*. Fungsi *likelihood* adalah fungsi gabungan dari fungsi distribusi peluangnya. Maka dari persamaan (3.5) diperoleh:

$$L(\beta^*,\sigma^2|Y^*) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(Y^*-X^*\beta^*)^T(Y^*-X^*\beta^*)}$$

$$= \left(\frac{1}{(2\pi)^{\frac{1}{2}}(\sigma^2)^{\frac{1}{2}}}\right)^n e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^*-2\beta^{*T}X^{*T}Y^*+\beta^{*T}X^{*T}X^*\beta^*)}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^*-2\beta^{*T}X^{*T}Y^*+\beta^{*T}X^{*T}X^*\beta^*)} \tag{3.6}$$

Jadi persamaan (3.6) fungsi *likelihood* dari persamaan (3.5)

Kemudian untuk memudahkan mengestimasi parameternya maka persamaan (3.6) diubah ke dalam bentuk *log likelihood*, yaitu

$ln\,L(\beta,\sigma^2|Y)$

$$= ln\left(\frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^*-2\beta^{*T}X^{*T}Y^*+\beta^{*T}X^{*T}X^*\beta^*)}\right)$$

$$= ln\left(\frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}}\right) + ln\, e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^*-2\beta^{*T}X^{*T}Y^*+\beta^{*T}X^{*T}X^*\beta^*)}$$

$$= ln\,1 - ln\left((2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}\right) + ln\, e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^*-2\beta^{*T}X^{*T}Y^*+\beta^{*T}X^{*T}X^*\beta^*)}$$

$$= ln\ 1 - ln\left((2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}\right) + ln\ e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^* - 2\beta^{*T}X^{*T}Y^* + \beta^{*T}X^{*T}X^*\beta^*)}$$

$$= -\left(\ln(2\pi)^{\frac{n}{2}} + ln\ (\sigma^2)^{\frac{n}{2}}\right) + ln\ e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^* - 2\beta^{*T}X^{*T}Y^* + \beta^{*T}X^{*T}X^*\beta^*)}$$

$$= -\left(\ln(2\pi)^{\frac{n}{2}} + ln\ (\sigma^2)^{\frac{n}{2}}\right) + ln\ e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^* - 2\beta^{*T}X^{*T}Y^* + \beta^{*T}X^{*T}X^*\beta^*)}$$

$$= -ln(2\pi)^{\frac{n}{2}} - ln(\sigma^2)^{\frac{n}{2}} \ - \frac{1}{2\sigma^2}(Y^{*T}Y^* - 2\beta^{*T}X^{*T}Y^* + \beta^{*T}X^{*T}X^*\beta^*)$$

$$= -\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \frac{1}{2\sigma^2}Y^{*T}Y^* + \frac{1}{2\sigma^2}2\beta^{*T}X^{*T}Y^* -$$

$$\frac{1}{2\sigma^2}\beta^{*T}X^{*T}X^*\beta^* \qquad (3.7)$$

3.2.1 Estimasi Parameter β^*

Untuk mengestimasi parameter β^* dengan *maximum likelihood estimation* yaitu dengan cara mendeferensialkan *log likelihood* pada persamaan (3.7) terhadap β^* kemudian disamadengankan nol.

Mendeferensialkan persamaan (3.7) terhadap β^*:

$$\frac{\partial\ ln\ L(\beta^*, \sigma^2|Y^*)}{\partial\beta^*}$$

$$= \frac{\partial(-\frac{n}{2}ln(2\pi) - \frac{n}{2}ln(\sigma^2) - \frac{1}{2\sigma^2}Y^{*T}Y^* + \frac{1}{2\sigma^2}2\beta^{*T}X^{*T}Y^* - \frac{1}{2\sigma^2}\beta^{*T}X^{*T}X^*\beta^*)}{\partial\beta^*}$$

$$= -0 - 0 - 0 + \frac{1}{2\sigma^2}2X^{*T}Y^* - \frac{1}{2\sigma^2}X^{*T}X^*\beta^* - \frac{1}{2\sigma^2}(\beta^{*T}X^{*T}X^*)^T$$

$$= \frac{1}{2\sigma^2}2X^{*T}Y^* - \frac{1}{2\sigma^2}X^{*T}X^*\beta^* - \frac{1}{2\sigma^2}(\beta^{*T}X^{*T}X^*)^T$$

$$= \frac{1}{\sigma^2}X^{*T}Y^* - \frac{1}{\sigma^2}X^{*T}X^*\beta^* \qquad (3.8)$$

Kemudian persamaan (3.8) disamadengankan nol, maka diperoleh

$$\frac{\partial \ln L(\beta^*, \sigma^2 | Y^*)}{\partial \beta^*} = 0$$

$$\frac{1}{\sigma^2} X^{*T} Y^* - \frac{1}{\sigma^2} (X^{*T} X^* \beta^*) = 0$$

$$\frac{1}{\sigma^2} (X^{*T} X^* \beta^*) = \frac{1}{\sigma^2} X^{*T} Y^*$$

$$(X^{*T} X^* \beta^*) = X^{*T} Y^*$$

$$(X^{*T} X^*)^{-1} (X^{*T} X^*) \beta^* = (X^{*T} X^*)^{-1} X^{*T} Y^*$$

$$I \beta^* = (X^{*T} X^*)^{-1} X^{*T} Y^*$$

$$\hat{\beta}^* = (X^{*T} X^*)^{-1} X^{*T} Y^* \qquad (3.9)$$

yang merupakan hasil estimasi dari β^* *secara maximum likelihood.*

3.2.2 Estimasi Parameter σ^2

Untuk mengestimasi parameter σ^2 dengan *maximum likelihood estimation* yaitu dengan cara mendiferensialkan *log likelihood* pada persamaan (3.7) terhadap σ^2 kemudian disama dengankan nol.

Mendeferensialkan persamaan (3.7) terhadap σ^2:

$$\frac{\partial \ln L(\beta, \sigma^2 | Y)}{\partial \sigma^2}$$

$$= \frac{\partial \left(-\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} Y^{*T} Y^* + \frac{1}{2\sigma^2} 2\beta^{*T} X^{*T} Y^* - \frac{1}{2\sigma^2} \beta^{*T} X^{*T} X^* \beta^* \right)}{\partial \sigma^2}$$

$$= -0 - \frac{n}{2\sigma^2} - \frac{1}{2(\sigma^2)^2} Y^{*T} Y^* + \frac{1}{2(\sigma^2)^2} 2\beta^{*T} X^{*T} Y^* -$$

$$\frac{1}{2(\sigma^2)^2} \beta^{*T} X^{*T} X^* \beta^*$$

$$= -\frac{n}{2\sigma^2} - \frac{1}{2(\sigma^2)^2} Y^{*T} Y^* + \frac{1}{2(\sigma^2)^2} 2\beta^{*T} X^{*T} Y^* -$$

$$\frac{1}{2(\sigma^2)^2} \beta^{*T} X^{*T} X^* \beta^* \qquad (3.10)$$

Kemudian persamaan (3.10) disamadengankan nol, sehingga diperoleh

$$\frac{\partial \ln L(\beta, \sigma^2 | y)}{\partial \sigma^2} = 0$$

$$0 = -\frac{n}{2\sigma^2} - \frac{1}{2(\sigma^2)^2} Y^{*T} Y^* + \frac{1}{2(\sigma^2)^2} 2\beta^{*T} X^{*T} Y^* - \frac{1}{2(\sigma^2)^2} \beta^{*T} X^{*T} X^* \beta^*$$

$$\frac{n}{2\sigma^2} = -\frac{1}{2(\sigma^2)^2} Y^{*T} Y^* + \frac{1}{2(\sigma^2)^2} 2\beta^{*T} X^{*T} Y^* - \frac{1}{2(\sigma^2)^2} \beta^{*T} X^{*T} X^* \beta^*$$

$$\frac{n\sigma^2}{2(\sigma^2)^2} = -\frac{1}{2(\sigma^2)^2} Y^{*T} Y^* + \frac{1}{2(\sigma^2)^2} 2\beta^{*T} X^{*T} Y^* - \frac{1}{2(\sigma^2)^2} \beta^{*T} X^{*T} X^* \beta^*$$

$$n\sigma^2 = -Y^{*T} Y^* + 2\beta^{*T} X^{*T} Y^* - \beta^{*T} X^{*T} X^* \beta^*$$

$$n\sigma^2 = Y^{*T} Y^* - 2\beta^{*T} X^{*T} Y^* + \beta^{*T} X^{*T} X^* \beta^*$$

$$\sigma^2 = \frac{1}{n} (Y^{*T} Y^* - 2\beta^{*T} X^{*T} Y^* + \beta^{*T} X^{*T} X^* \beta^*)$$

$$\hat{\sigma}^2 = \frac{1}{n} (Y^* - X^* \beta^*)^T (Y^* - X^* \beta^*) \qquad (3.11)$$

yang merupakan hasil estimasi dari σ^2 secara *maximum llikelihood.*

3.2.3 Menentukan Fungsi *Likelihood* dan *Log Likelihood* dari Hasil Estimasi Parameter

Fungsi *likelihood* dari estimator $\hat{\beta}^*$ dan $\hat{\sigma}^2$ adalah sebagai berikut:

$$L(\hat{\beta}^*, \hat{\sigma}^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2}(Y^* - X^*\beta^*)^T (Y^* - X^*\beta^*)}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^* - 2\beta^{*T}X^{*T}Y^* + \beta^{*T}X^{*T}X^*\beta^*)}$$

sehingga fungsi *likelihood* dari parameter $\hat{\beta}^*$ dan $\hat{\sigma}^2$ adalah

$$L(\hat{\beta}^*, \hat{\sigma}^2) = \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^* - X^*\beta^*)^T(Y^* - X^*\beta^*)} \tag{3.12}$$

dan fungsi *Log likelihood* dari persamaan (3.12) adalah

$$ln\, L(\hat{\beta}^*, \hat{\sigma}^2) = ln \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma^2}(Y^* - X^*\beta^*)^T(Y^* - X^*\beta^*)}$$

$$= \ln \frac{1}{(2\pi)^{\frac{n}{2}}(\sigma^2)^{\frac{n}{2}}} + \ln e^{-\frac{1}{2\sigma^2}(Y^{*T}Y^* - 2\beta^{*T}X^{*T}Y^* + \beta^{*T}X^{*T}X^*\beta^*)}$$

$$= ln\, 1 - \left(ln(2\pi)^{\frac{n}{2}} + ln(\sigma^2)^{\frac{n}{2}}\right) + ln\, e^{-\frac{1}{2\sigma^2}(n\sigma^2)}$$

$$= -\left(\ln(2\pi)^{\frac{n}{2}} + \ln(\sigma^2)^{\frac{n}{2}}\right) + \ln e^{-\frac{n}{2}}$$

$$= -\ln(2\pi)^{\frac{n}{2}} - \ln(\sigma^2)^{\frac{n}{2}} - \frac{n}{2}\ln e$$

$$= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{n}{2} \tag{3.13}$$

3.3 Menentukan Sifat-Sifat Estimasi Parameter Regresi Nonlinier

3.3.1 Sifat Estimator Parameter $\boldsymbol{\beta}^*$

Estimator $\hat{\beta}^*$ adalah estimator tak bias dari parameter β^* karena $E(\hat{\beta}^*) = \beta^*$. Bukti:

$$\hat{\beta}^* = (X^{*T}X^*)^{-1}X^{*T}Y^*$$

$$\begin{aligned} E(\hat{\beta}^*) &= E((X^{*T}X^*)^{-1}X^{*T}Y^*) \\ &= (X^{*T}X^*)^{-1}X^{*T}E(Y^*) \\ &= (X^{*T}X^*)^{-1}X^{*T}X^*\beta^* \\ &= I\beta^* \\ &= \beta^* \end{aligned} \tag{3.14}$$

sehingga terbukti bahwa $\hat{\beta}^*$ adalah estimator tak bias (unbias).

Selanjutnya akan dibuktikan bahwa estimator $\hat{\beta}^*$ adalah estimator efisien. Dikatakan estimator efisien apabila mempunyai nilai variansi yang terkecil. Bukti:

$$\begin{aligned} var\ (\hat{\beta}^*) &= E\left[\left(\hat{\beta}^* - E(\hat{\beta}^*)\right)\left(\hat{\beta}^* - E(\hat{\beta}^*)\right)^T\right] \\ &= E\left[(\hat{\beta}^* - \beta^*)(\hat{\beta}^* - \beta^*)^T\right] \end{aligned}$$

karena:

$$\begin{aligned} \hat{\beta}^* &= (X^{*T}X^*)^{-1}X^{*T}Y^* \\ &= (X^{*T}X^*)^{-1}X^{*T}(X^*\beta^* + \varepsilon^*) \\ &= (X^{*T}X^*)^{-1}X^{*T}X^*\beta^* + (X^{*T}X^*)^{-1}X^{*T}\varepsilon^* \\ &= \beta^* + (X^{*T}X^*)^{-1}X^{*T}\varepsilon^* \\ \hat{\beta}^* - \beta^* &= (X^{*T}X^*)^{-1}X^{*T}\varepsilon^* \end{aligned}$$

maka:

$$var\ (\hat{\beta}^*) = E\left[(\hat{\beta}^* - \beta^*)(\hat{\beta}^* - \beta^*)^T\right]$$

$$= E[((X^{*T}X^*)^{-1}X^{*T}\varepsilon^*)((X^{*T}X^*)^{-1}X^{*T}\varepsilon^*)^T]$$

$$= E[((X^{*T}X^*)^{-1}X^{*T}\varepsilon^*)(\varepsilon^{*T}X^*(X^{*T}X^*)^{-1})]$$

$$= E[((X^{*T}X^*)^{-1}X^{*T}\varepsilon^*\varepsilon^{*T}X^*(X^{*T}X^*)^{-1}]$$

$$= (X^{*T}X^*)^{-1}X^{*T}E(\varepsilon^*\varepsilon^{*T})X^*(X^{*T}X^*)^{-1}$$

$$= (X^{*T}X^*)^{-1}X^{*T}\sigma^2 I X^*(X^{*T}X^*)^{-1}$$

$$= (X^{*T}X^*)^{-1}X^{*T}X^*(X^{*T}X^*)^{-1}\sigma^2$$

$$= (X^{*T}X^*)^{-1}\sigma^2 \quad (3.15)$$

Sehingga $var\left(\hat{\beta}^*\right) = (X^{*T}X^*)^{-1}\sigma^2$ harus sekecil mungkin agar estimator $\hat{\beta}^*$ efisien. Kemudian sifat estimator yang ketiga yaitu konsisten. Dikatakan estimator yang konsisten jika $lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| < \varepsilon = 1\right.$ sehingga $lim_{n\to\infty} E\left(\hat{\theta} - E\left(\hat{\theta}\right)\right)^2 = \lim_{n\to\infty}(X^{*T}X^*)^{-1}\sigma^2 = 0$. Sehingga dapat dikatakan bahwa $\hat{\beta}^*$merupakan estimator yang konsisten.

3.3.2 Sifat Estimator Parameter σ^2

Estimator $\hat{\sigma}^2$ adalah estimator tak bias dari parameter σ^2 karena $E(\hat{\sigma}^2) = \sigma^2$. Bukti:

$$\hat{\sigma}^2 = \frac{1}{n}(Y^* - X^*\beta^*)^T(Y^* - X^*\beta^*)$$

$$E(\hat{\sigma}^2) = E\left(\frac{1}{n}(Y^* - X^*\beta^*)^T(Y^* - X^*\beta^*)\right)$$

$$= \frac{1}{n}E\big((X^*\beta^* + \varepsilon^* - X^*\beta^*)^T(X^*\beta^* + \varepsilon^* - X^*\beta^*)\big)$$

$$= \frac{1}{n}E\big((\beta^{*T}X^{*T} + \varepsilon^{*T} - \beta^{*T}X^{*T})(X^*\beta^* + \varepsilon^* - X^*\beta^*)\big)$$

$$= \frac{1}{n} E((\beta^{*T} X^{*T} - \beta^{*T} X^{*T} + \varepsilon^{*T})(\varepsilon^* + X^* \beta^* - X^* \beta^*))$$

$$= \frac{1}{n} E(\beta^{*T} X^{*T} - \beta^{*T} X^{*T}) + E(\varepsilon^{*T} \varepsilon^*) + E(X^* \beta^* - X^* \beta^*)$$

$$= \frac{1}{n} (\varepsilon^{*T} \varepsilon^*)$$

$$= \frac{1}{n} \sigma^2 \qquad (3.16)$$

Jadi $\hat{\sigma}^2$ merupakan estimator bias untuk σ^2.

3.4 Pendeteksian *Outlier*

Pendeteksian *outlier* pada regresi nonlinier dengan metode statistik *likelihood displacement* (LD) dilakukan dengan cara menghilangkan pengamatan yang diduga mengandung *outlier* pada model. Misalkan ada k pengamatan yang dikumpulkan dalam suatu himpunan tertentu, dengan k adalah pengamatan yang diduga mengandung *outlier*. Dimana $k < n$. Dan misalkan indeks A_k adalah kumpulan dari k pengamatan yang diduga *outlier* dengan $A_k = i_1, i_2, \dots, i_k$, dan misalkan indeks $i_k \in \{1,2, \dots, n\}$.

Dengan mempertimbangkan pengamatan k dalam estimasi parameter, maka likelihood displacement untuk $\hat{\theta} = \hat{\beta}^*, \hat{\sigma}^2$dan $\hat{\theta}_{[A_k]} = \hat{\beta}^*_{[A_k]}, \hat{\sigma}^2_{[A_k]}$adalah:

$$LD_{A_k}(\beta^* | \sigma^2) = 2\left(\ln L\left(\hat{\beta}^*, \hat{\sigma}^2\right) - \ln L\left(\hat{\beta}^*_{[A_k]}, \hat{\sigma}^2_{[A_k]}\right)\right)$$

dimana $\hat{\beta}^*$adalah *maximum likelihood estimation* dari β^* dan $\hat{\sigma}^2$ adalah *maximum likelihood estimation* pada keseluruhan pengamatan dan $\hat{\beta}^*_{[A_k]}$ dan $\hat{\sigma}^2_{[A_k]}$ adalah MLE dari β^* dan σ^2 ketika pengamatan dengan indeks A_k dihilangkan.

Pada kasus khusus yaitu $\theta_1 = (\beta^*_1, \sigma^2_1)$ subset dari $\theta = (\beta^*, \sigma^2)$, maka fungsi *likelihood displacement* dapat dimodifikasi menjadi

$$LD_{A_k}((\beta^*_1, \sigma^2_1)|(\beta^*_2, \sigma^2_2))$$
$$= 2\{ln\ L(\hat{\beta}^*, \hat{\sigma}^2) - \ln L((\hat{\beta}^*_1, \hat{\sigma}^2_1)_{[A_k]}\ ,(\hat{\beta}^*_2, \hat{\sigma}^2_2)(\hat{\beta}^*_1, \hat{\sigma}^2_1)_{[A_k]})\} \quad (3.17)$$

dimana $L((\hat{\beta}^*_2, \hat{\sigma}^2_2)(\hat{\beta}^*_1, \hat{\sigma}^2_1)_{[A_k]}) = max_{(\beta^*_2, \sigma^2_2)} L\left((\hat{\beta}^*_1, \hat{\sigma}^2_1)_{[A_k]}\ ,(\hat{\beta}^*_2, \hat{\sigma}^2_2)\right)$ adalah memaksimumkan fungsi *log likelihood* pada parameter (β^*_2, σ^2_2) dengan $(\beta^*_1, \sigma^2_1) = (\hat{\beta}^*_1, \hat{\sigma}^2_1)_{[A_k]}$ maka $\beta^*_1 = \hat{\beta}^*_{1[A_k]}$ dan $\sigma^2_1 = \hat{\sigma}^2_{1[A_k]}$ adalah *maximum likelihood estimation* dari (β^*_1, σ^2_1) ketika pengamatan k dihilangkan.

Selanjutnya untuk keseluruhan data ketika k pengamatan pada himpunan A_k dihilangkan maka modelnya menjadi:

$$Y^*_{[A_k]} = X^*_{[A_k]}\beta^*_{[A_k]} + \varepsilon_{[A_k]} \quad (3.18)$$

dengan $Y^*_{[A_k]} \sim N(0, I\sigma^2)$

3.4.1 Estimasi Parameter $\boldsymbol{\beta^*_{[A_k]}}$

Dari persamaan (3.18) maka *maximum likelihood estimation* dari $\hat{\beta}^*_{[A_k]}$ sesuai persamaan (3.9) adalah:

$$\hat{\beta}^*_{[A_k]} = \left(X^{*T}_{[A_k]}X^*_{[A_k]}\right)^{-1}X^{*T}_{[A_k]}Y^*_{[A_k]} \tag{3.19}$$

dimana:

$$X^{*T}_{[A_k]}X^*_{[A_k]} = X^{*T}X^* - X^{*T}_{A_k}X^*_{A_k} \text{ dan } X^{*T}_{[A_k]}Y^*_{[A_k]} = X^{*T}Y^* - X^{*T}_{A_k}Y^*_{A_k} \tag{3.20}$$

Maka dari (3.19) dan (3.20)

$$\begin{aligned}\hat{\beta}^*_{[A_k]} &= \left(X^{*T}X^* - X^{*T}_{A_k}X^*_{A_k}\right)^{-1}X^{*T}Y^* - X^{*T}_{A_k}Y^*_{A_k}\\ &= \left((X^{*T}X^*)^{-1} + (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - X^*_{A_k}(X^{*T}X^*)^{-1}X^{*T}_{A_k}\right)^{-1}X^*_{A_k}(X^{*T}X^*)^{-1}\right)\\ &\quad (X^{*T}Y^* - X^{*T}_{A_k}Y^*_{A_k})\end{aligned}$$

Misal: $X^*_{A_k}(X^{*T}X^*)^{-1}X^{*T}_{A_k} = Q_{A_k}$ maka,

$$\begin{aligned}\hat{\beta}^*_{[A_k]} &= \left((X^{*T}X^*)^{-1} + (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}X^*_{A_k}(X^{*T}X^*)^{-1}\right) \text{ x}\\ &\quad (X^{*T}Y^* - X^{*T}_{A_k}Y^*_{A_k})\end{aligned}$$

$$\begin{aligned}&= (X^{*T}X^*)^{-1}X^{*T}Y^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}Y^*_{A_k} +\\ &\quad (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}X^*_{A_k}(X^{*T}X^*)^{-1}X^{*T}Y^* -\\ &\quad (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}X^*_{A_k}(X^{*T}X^*)^{-1}X^{*T}_{A_k}Y^*_{A_k}\\ &= \hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}Y^*_{A_k} + (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}X^*_{A_k}\hat{\beta}^* -\\ &\quad (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}Q_{A_k}Y^*_{A_k}\\ &= \hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(Y^*_{A_k} - \left(I - Q_{A_k}\right)^{-1}X^*_{A_k}\hat{\beta}^* + \left(I - Q_{A_k}\right)^{-1}Q_{A_k}Y^*_{A_k}\right)\\ &= \hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I + \left(\left(I - Q_{A_k}\right)^{-1}Q_{A_k}\right)Y^*_{A_k} - \left(I - Q_{A_k}\right)^{-1}X^*_{A_k}\hat{\beta}^*\right)\end{aligned}$$

Misal:

$$I + \left(\left(I - Q_{A_k}\right)^{-1} Q_{A_k}\right) = (I - Q_{A_k})(I - Q_{A_k})^{-1} + \left(\left(I - Q_{A_k}\right)^{-1} Q_{A_k}\right)$$

$$= \left(I - Q_{A_k}\right)^{-1}\{(I - Q_{A_k}) + Q_{A_k}\}$$

$$= \left(I - Q_{A_k}\right)^{-1}\{(I - Q_{A_k} + Q_{A_k}\}$$

$$= \left(I - Q_{A_k}\right)^{-1} I$$

$$= \left(I - Q_{A_k}\right)^{-1}$$

Sehingga:

$$\hat{\beta}^*_{[A_k]} = \hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(\left(I - Q_{A_k}\right)^{-1} Y^*_{A_k} - \left(I - Q_{A_k}\right)^{-1} X^*_{A_k}\hat{\beta}^*\right)$$

$$\hat{\beta}^*_{[A_k]} = \hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}(Y^*_{A_k} - X^*_{A_k}\hat{\beta}^*)$$

misal $\hat{\varepsilon}^*_{A_k} = Y^*_{A_k} - X^*_{A_k}\hat{\beta}^*$, maka

$$\hat{\beta}^*_{[A_k]} = \hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}^*_{A_k} \quad (3.21)$$

Selanjutnya menentukan sifat-sifat *estimator* $\hat{\beta}^*_{[A_k]}$.

$E\left(\hat{\beta}^*_{[A_k]}\right)$

$$= E(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}^*_{A_k})$$

$$= E(\hat{\beta}^*) - E\left((X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}^*_{A_k}\right)$$

$$= E((X^{*T}X^*)^{-1}X^{*T}Y^*) - E\left((X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}(Y^*_{A_k} - X^*_{A_k}\hat{\beta}^*)\right)$$

$$= (X^{*T}X^*)^{-1}X^{*T}E(Y^*) - \left((X^{*T}X^*)^{-1}X^{*T}_{A_k}\left(I - Q_{A_k}\right)^{-1}(E(Y^*_{A_k}) - X^*_{A_k}\hat{\beta}^*)\right)$$

$$= (X^{*T}X^{*})^{-1}X^{*T}X^{*}\beta^{*} - \left((X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\left(X_{A_k}^{*}\hat{\beta}^{*} - X_{A_k}^{*}\hat{\beta}^{*}\right)\right)$$

$$= I\beta^{*} - \left((X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}(0)\right)$$

$$= \beta^{*} - 0$$

$$= \beta^{*}$$

Jadi $\hat{\beta}_{[A_k]}^{*}$ adalah estimator tak bias *(unbias).*

3.4.2 Estimasi Parameter $\hat{\sigma}_{[A_k]}^{2}$

Dari persamaan (3.11) diketahui bahwa $\hat{\sigma}^2 = \frac{1}{n}\left(Y^{*} - X^{*}\hat{\beta}^{*}\right)^{T}\left(Y^{*} - X^{*}\hat{\beta}^{*}\right)$, maka $\hat{\sigma}_{[A_k]}^{2}$ dari persamaan (3.18) adalah:

$$\hat{\sigma}_{[A_k]}^{2} = \frac{1}{n-k}\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\hat{\beta}_{[A_k]}^{*}\right)^{T}\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\hat{\beta}_{[A_k]}^{*}\right)$$

$$= \frac{1}{n-k}\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)\right)^{T}$$
$$\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)\right)$$

$$= \frac{1}{n-k}\left(Y_{[A_k]}^{*T} - \left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)^{T}X_{[A\,k]}^{*T}\right)$$
$$\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)\right)$$

$$= \frac{1}{n-k}\left(Y_{[A_k]}^{*T}Y_{[A_k]}^{*} - Y_{[A_k]}^{*T}X_{[A\,k]}^{*}\left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right) - \right.$$
$$\left(\left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)^{T}X_{[A\,k]}^{*T}Y_{[A_k]}^{*}\right)^{T} +$$
$$\left(\left(\hat{\beta}^{*} - (X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)^{T}X_{[A\,k]}^{*T}\right)\mathrm{x}$$

$$\Big(X^*_{[A\,k]}\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big)\Big)\Big)$$

$$= \frac{1}{n-k}\Big(Y^{*T}_{[A_k]}Y^*_{[A_k]} - 2Y^{*T}_{[A_k]}X^*_{[A\,k]}\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big) +$$

$$\Big(\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big)^T X^{*T}_{[A\,k]}\Big) \mathrm{x}$$

$$\Big(X^*_{[A\,k]}\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big)\Big)\Big)$$

$$= \frac{1}{n-k}\Big(Y^{*T}_{[A_k]}Y^*_{[A_k]} - 2Y^{*T}_{[A_k]}X^*_{[A\,k]}\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big) +$$

$$\Big(\hat{\beta}^* - \Big((X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big)^T\Big) X^{*T}_{[A\,k]} \mathrm{x}$$

$$X^*_{[A\,k]}\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big)\Big)$$

$$= \frac{1}{n-k}(Y^{*T}_{[A_k]}Y^*_{[A_k]} - 2Y^{*T}_{[A_k]}X^*_{[A\,k]}\hat{\beta}^* +$$

$$2Y^{*T}_{[A_k]}X^*_{[A\,k]}(X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k} -$$

$$\Big(\hat{\varepsilon}^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}X^*_{A_k}(X^{*T}X^*)^{-1} - \hat{\beta}^{*T}\Big) X^{*T}_{[A\,k]} \mathrm{x}$$

$$X^*_{[A\,k]}\Big(\hat{\beta}^* - (X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k}\Big)\Big)$$

$$= \frac{1}{n-k}(Y^{*T}_{[A_k]}Y^*_{[A_k]} - 2Y^{*T}_{[A_k]}X^*_{[A\,k]}\hat{\beta}^* +$$

$$2Y^{*T}_{[A_k]}X^*_{[A\,k]}(X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k} -$$

$$\Big(\hat{\varepsilon}^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}X^*_{A_k}(X^{*T}X^*)^{-1}X^{*T}_{[A\,k]} - \hat{\beta}^{*T}X^{*T}_{[A\,k]}\Big) \mathrm{x}$$

$$(X^*_{[A\,k]}\hat{\beta}^* - X^*_{[A\,k]}(X^{*T}X^*)^{-1}X^{*T}_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}^*_{A_k})\Big)$$

$$= \frac{1}{n-k}(Y^{*T}_{[A_k]}Y^*_{[A_k]} - 2Y^{*T}_{[A_k]}X^*_{[A\,k]}\hat{\beta}^* +$$

$$2Y_{[A_k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*} -$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*} +$$

$$\hat{\beta}^{*T}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*} - \hat{\beta}^{*T}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big)$$

$$= \frac{1}{n-k}\Big(Y_{[A_k]}^{*T}Y_{[A_k]}^{*} - 2Y_{[A_k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*} + \hat{\beta}^{*T}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*} +$$

$$2Y_{[A_k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*} -$$

$$\left(\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*}\right)^{T} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*} -$$

$$\hat{\beta}^{*T}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big)$$

$$= \frac{1}{n-k}\Big(Y_{[A_k]}^{*T}Y_{[A_k]}^{*} - 2Y_{[A_k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*} + \hat{\beta}^{*T}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}\hat{\beta}^{*} +$$

$$2Y_{[A_k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*} -$$

$$2\hat{\beta}^{*T}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big) +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{[A\,k]}^{*T}X_{[A\,k]}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}$$

$$= \frac{1}{n-k}\Big(\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\hat{\beta}^{*}\right)^{T}\left(Y_{[A_k]}^{*} - X_{[A\,k]}^{*}\hat{\beta}^{*}\right) +$$

$$2\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{[A\,k]}^{*T}Y_{[A_k]}^{*} - 2\hat{\beta}^{*T}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big) +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}X_{A_k}^{*}I(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\left(I-Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}$$

$$= \frac{1}{n-k}\left(n\hat{\sigma}^2 + 2\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*}\hat{\beta}^{*} - 2\hat{\beta}^{*T} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right) +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} I (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}$$

$$= \frac{1}{n-k}\left(n\hat{\sigma}^2 + \hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} I (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right.$$

$$\hat{\sigma}_{[A_k]}^2 = \frac{n}{n-k}\sigma^2 + \frac{1}{n-k}\hat{\varepsilon}_{A_m}^{T} Q_{A_m}\left(I - Q_{A_m}\right)^{-1}\hat{\varepsilon}_{A_k} \tag{3.22}$$

3.4.3 Estimasi Parameter $\hat{\sigma}^2\left(\hat{\beta}^{*}{}_{[A_k]}\right)$

Untuk kasus khusus pada persamaan (3.17) jika $\theta_1 = \beta_1^*$ dan $\theta_2 = \sigma_2^2$ sehingga $\hat{\theta} = (\hat{\beta}^*, \hat{\sigma}^2)$ maka persamaan (3.17) menjadi

$$LD_{A_k}(\beta^*|\sigma^2) = 2\left\{\ln L(\hat{\beta}^*, \hat{\sigma}^2) - \ln L\left(\hat{\beta}^{*}{}_{[A_k]}, \hat{\sigma}^2\left(\hat{\beta}^{*}{}_{[A_k]}\right)\right)\right\} \tag{3.23}$$

dimana $\hat{\sigma}^2\left(\hat{\beta}^{*}{}_{[A_k]}\right)$ adalah *maximum likelihood estimation* dari $\hat{\sigma}^2$ ketika β^* diestimasi dengan $\hat{\beta}^{*}{}_{[A_k]}$. Subtitusikan $\hat{\beta}^{*}{}_{[A_k]}$ untuk β^* pada $\hat{\sigma}^2$.

Maka:

$$\hat{\sigma}^2\left(\hat{\beta}^{*}{}_{[A_k]}\right) = \frac{1}{n}\left(Y^* - X^*\hat{\beta}^{*}{}_{[A_k]}\right)^T (Y^* - X^*\hat{\beta}^{*}{}_{[A_k]})$$

Karena:

$$Y^* - X^*\hat{\beta}_{[A_k]}{}^{*} = Y^* - X^*\left(\left(\hat{\beta}^* - (X^{*T}X^*)^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)\right)$$

$$= Y^* - X^*\hat{\beta}^* + X^*(X^{*T}X^*)^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}$$

dan

$$\left(Y^* - X^*\hat{\beta}_{[A_k]}{}^{*}\right)^T = \left(Y^* - X^*\hat{\beta}^* + X^*(X^{*T}X^*)^{-1}X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}\right)^T$$

$$= Y^{*T} - \hat{\beta}^{*T} X^{*T} + \hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T}$$

Sehingga diperoleh:

$$\hat{\sigma}^2\left(\hat{\beta}_{[A_k]}\right)$$

$$= \frac{1}{n}\Big[\left(Y^{*T} - \hat{\beta}^{*T} X^{*T} + \hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T}\right)$$

$$\left(Y^{*} - X^{*}\hat{\beta}^{*} + X^{*}(X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*}\right)\Big]$$

$$= \frac{1}{n}\Big[Y^{*T}Y^{*} - Y^{*T}X^{*}\hat{\beta}^{*} + Y^{*T}X^{*}(X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*} - \hat{\beta}^{*T} X^{*T} Y^{*} +$$

$$\hat{\beta}^{*T} X^{*T} X^{*} \hat{\beta}^{*} - \hat{\beta}^{*T} X^{*T} X^{*} (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} Y^{*} -$$
$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} X^{*} \hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} X^{*} (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[\left(Y^{*} - X^{*}\hat{\beta}^{*}\right)^{T}\left(Y^{*} - X^{*}\hat{\beta}^{*}\right) + Y^{*T}X^{*}(X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*} -$$

$$\hat{\beta}^{*T} X^{*T} X^{*} (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} Y^{*}$$
$$- \hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} X^{*} \hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} X^{*} (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[(\varepsilon^{T}\varepsilon) + 2\left(Y^{*T}X^{*}(X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*}\right)$$
$$- \hat{\beta}^{*T} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*} -$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} X^{*} \hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} X_{A_k}^{*} (X^{*T}X^{*})^{-1} X^{*T} X^{*} (X^{*T}X^{*})^{-1} X_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} \hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[(\varepsilon^T\varepsilon) + 2\Big(\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}\hat{\beta}^{*}\Big) - \hat{\beta}^{*T}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*} -$$

$$\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X^{*T}X^{*}\hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X^{*T}X^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[(\varepsilon^T\varepsilon) + \hat{\beta}^{*T}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*} - \hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X^{*T}X^{*}\hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X^{*T}X^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[(\varepsilon^T\varepsilon) + \hat{\beta}^{*T}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*} - \hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}\hat{\beta}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}X_{A_k}^{*}(X^{*T}X^{*})^{-1}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[(\varepsilon^T\varepsilon) + \hat{\beta}^{*T}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*} - \hat{\beta}^{*T}X_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*} +$$

$$\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}Q_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}\Big[(\varepsilon^T\varepsilon) + \hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}Q_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*}\Big]$$

$$= \frac{1}{n}(\varepsilon^T\varepsilon) + \frac{1}{n}\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}Q_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*}$$

yang merupakan estimator $\hat{\sigma}^2$ ketika β^* diestimasi dengan $\hat{\beta}^*{}_{[A_k]}$

$$\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right) = \hat{\sigma}^2 + \frac{1}{n}\hat{\varepsilon}_{A_k}^{*T}\big(I - Q_{A_k}\big)^{-1}Q_{A_k}\big(I - Q_{A_k}\big)^{-1}\hat{\varepsilon}_{A_k}^{*} \tag{3.24}$$

3.4.4 Menentukan Fungsi Likelihood Ketika Pengamatan yang Diduga Mengandung *Outlier* Dihilangkan

Fungsi *likelihood* dari $\hat{\beta}^*{}_{[A_k]}, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)$ adalah:

$$L\left(\hat{\beta}^*{}_{[A_k]}, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)$$

$$=\left(\frac{1}{\sqrt{2\pi\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)}}\right)^n e^{-\frac{1}{2\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)}\left(Y^*-X^*_{[A_k]}\beta^*_{[A_k]}\right)^T\left(Y^*-X^*_{[A_k]}\beta^*_{[A_k]}\right)}$$

$$=\frac{1}{\left(2\pi^{\frac{n}{2}}\right)\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)^{\frac{n}{2}}\right)} e^{-\frac{1}{2\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)}\left(Y^*-X^*_{[A_k]}\beta^*_{[A_k]}\right)^T\left(Y^*-X^*_{[A_k]}\beta^*_{[A_k]}\right)}$$

kemudian dilogaritmanaturalkan:

$$ln\, L\left(\hat{\beta}^*{}_{[A_k]}, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)$$

$$= ln\left(\frac{1}{\left(2\pi^{\frac{n}{2}}\right)\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)^{\frac{n}{2}}\right)} e^{-\frac{1}{2\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)}\left(Y^*-X^*_{[A_k]}\beta^*_{[A_k]}\right)^T\left(Y^*-X^*_{[A_k]}\beta^*_{[A_k]}\right)}\right)$$

$$= ln\frac{1}{(2\pi)^{\frac{n}{2}}\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)^{\frac{n}{2}}} + ln\, e^{-\frac{1}{2\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)}\left(n\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)\right)}$$

$$= ln\frac{1}{(2\pi)^{\frac{n}{2}}\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)^{\frac{n}{2}}} + ln\, e^{-\frac{n}{2}}$$

$$= ln\, 1 - \left(ln(2\pi)^{\frac{n}{2}} + ln\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)^{\frac{n}{2}}\right) + ln\, e^{-\frac{n}{2}}$$

$$= -\left(ln(2\pi)^{\frac{n}{2}} + ln\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)^{\frac{n}{2}}\right) + ln\, e^{-\frac{n}{2}}$$

$$= -\, ln(2\pi)^{\frac{n}{2}} - ln\left(\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)^{\frac{n}{2}} - \frac{n}{2}\; ln\, e$$

$$= -\frac{n}{2} ln\, 2\pi - \frac{n}{2} ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right) - \frac{n}{2} \qquad (3.25)$$

Persamaan (3.25) merupakan fungsi *log likelihood* dari fungsi *likelihood* $L\left(\hat{\beta}^*{}_{[A_k]}, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)$.

Likelihood Displacement dari β^* dan σ^2 yang diberikan pada persamaan (3.23) adalah:

$$LD_{A_k}(\beta^*|\sigma^2) = 2\left\{\ln L(\hat{\beta}^*, \hat{\sigma}^2) - \ln L\left(\hat{\beta}^*{}_{[A_k]}, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)\right\}$$

Subtitusikan persamaan (3.13) dan (3.25) ke persamaan (3.23) maka:

$$LD_{A_k}(\beta^*|\sigma^2)$$

$$= 2\left\{\left(-\frac{n}{2} ln(2\pi) - \frac{n}{2} ln(\sigma^2)\right) - \left(-\frac{n}{2} ln\, 2\pi - \frac{n}{2} ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right)\right\}$$

$$= 2\left\{-\frac{n}{2} ln(2\pi) - \frac{n}{2} ln(\sigma^2) + \frac{n}{2} ln\, 2\pi + \frac{n}{2} ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right\}$$

$$= 2\left\{-\frac{n}{2} ln(\sigma^2) + \frac{n}{2} ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right\}$$

$$= -n\, ln\, \sigma^2 + n\;\; ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)$$

$$= n\left\{-\, ln\, \sigma^2 + ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)\right\}$$

$$= n\left\{ln\, \hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right) - ln\, \sigma^2\right\}$$

$$= n\, ln\left(\frac{\hat{\sigma}^2\left(\hat{\beta}^*{}_{[A_k]}\right)}{\sigma^2}\right)$$

$$= n\, ln\left\{\frac{\hat{\sigma}^2 + \frac{1}{n}\hat{\varepsilon}_{A_k}^{*T}\left(I - Q_{A_k}\right)^{-1} Q_{A_k}\left(I - Q_{A_k}\right)^{-1}\hat{\varepsilon}_{A_k}^{*}}{\sigma^2}\right\}$$

misal $C_{A_k} = \left(I - Q_{A_k}\right)^{-1} Q_{A_k}\left(I - Q_{A_k}\right)^{-1}$, maka:

$$LD_{A_k} = n\, ln\left\{\frac{\hat{\sigma}^2 + \frac{1}{n}\hat{\varepsilon}_{A_k}^{*T} C_{A_k}\hat{\varepsilon}_{A_k}^{*}}{\sigma^2}\right\}$$

$$= n\, ln\left\{\frac{\hat{\sigma}^2}{\hat{\sigma}^2} + \frac{\frac{1}{n}\hat{\varepsilon}_{A_k}^{*T} C_{A_k}\hat{\varepsilon}_{A_k}^{*}}{\hat{\sigma}^2}\right\}$$

$$= n\, ln\left\{1 + \frac{1}{n\hat{\sigma}^2}\hat{\varepsilon}_{A_k}^{*T} C_{A_k}\hat{\varepsilon}_{A_k}^{*}\right\}$$

Sehingga *Likelihood Displacement* yang diduga mengandung *outlier* adalah sebagai berikut:

$$LD_{A_k} = n\, ln\left\{1 + \frac{1}{n\hat{\sigma}^2}\hat{\varepsilon}_{A_k}^{*T} C_{A_k}\hat{\varepsilon}_{A_k}^{*}\right\} \quad (3.26)$$

3.5 Menentukan Uji *Outlier*

Untuk menunjukkan keakuratan dari hasil metode LD dalam mendeteksi adanya *outlier* pada persamaan (3.26)

$$LD_{A_k} = n\, ln\left\{1 + \frac{1}{n\hat{\sigma}^2}\hat{\varepsilon}_{A_k}^{*T} C_{A_k}\hat{\varepsilon}_{A_k}^{*}\right\}$$

maka digunakan uji statistik. Uji statistik disini dilakukan dengan cara membandingkan MSE dari metode LD dengan MSE dari regresi pada umumnya (regresi tanpa *outlier*). Statistik uji yang digunakan adalah

$$\Lambda = \frac{MSE_{LD_{A_k}}}{MSE_{reg}} \sim \chi^2_{\lambda_i}$$

dimana λ_i, $i = 1, 2, \dots, k$, adalah nilai eigen dari C_{A_k}. Ketika nilai $MSE_{LD_{A_k}}$ lebih besar dari pada MSE_{reg} maka nilai Λ akan semakin besar.

Dari hasil uji statistik yang telah dijelaskan, maka diberikan uji hipotesis sebagai berikut:

$H_0 : A_k =$ adalah bukan *outlier*

$H_1 : A_k =$ adalah *outlier*

H_0 ditolak jika $\chi^2_{hitung} > \chi^2_{tabel}$ dan H_0 diterima jika $\chi^2_{hitung} < \chi^2_{tabel}$.

3.6 Keterkaitan Hasil Penelitian dengan Kajian Agama

Outlier adalah suatu data yang ekstrim atau data yang tidak sesuai pada model pada umumnya. Seperti yang disajikan dalam bab II, bahwa terdapat ayat Al-Qur'an yang berkenaan dengan masalah *outlier*. Dalam statistika adanya *outlier* perlu dideteksi, karena terkadang *outlier* merupakan suatu data yang tidak mampu dijelaskan oleh data yang lain. Ayat Al-Qur'an tersebut yaitu surat Al-Jinn ayat 14 yang berbunyi:

وَأَنَّا مِنَّا ٱلْمُسْلِمُونَ وَمِنَّا ٱلْقَـٰسِطُونَ ۖ فَمَنْ أَسْلَمَ فَأُو۟لَـٰٓئِكَ تَحَرَّوْا۟ رَشَدًا ﴿١٤﴾

Artinya: *"Dan sesungguhnya di antara kami ada orang-orang yang taat dan ada pula orang-orang yang menyimpang dari kebenaran. Barangsiapa yang taat, maka mereka itu benar-benar telah memilih jalan yang lurus"*(Qs. Al-Jinn/72 :14).

Pengertian *outlier* dari ayat di atas terdapat pada kata "*menyimpang dari kebenaran*". Konsep yang dapat diambil dari ayat di atas jika dikaitkan dengan pengertian *outlier* yaitu, jika dalam statistika *outlier* merupakan suatu data yang tidak mengikuti pola umum pada model atau data yang keluar dari model dan tidak berada dalam daerah selang kepercayaan, maka dalam potongan ayat yang berarti "*dan sesungguhnya di antara kami*" dianalogikan menjadi suatu kumpulan data. Kemudian potongan ayat selanjutnya yang berarti "*ada orang-orang yang taat dan ada pula orang-orang yang menyimpang dari kebenaran*" orang-orang (jin) yang menyimpang ini dianalogikan sebagai *outlier* sedangkan orang-orang (jin) yang taat merupakan sekumpulan data yang mengikuti pola umumnya.

Kemudian untuk mendeteksi adanya *outlier* tersebut maka digunakan estimasi terhadap parameternya. Mengenai masalah estimasi ini telah termaktup dalam ayat Al-Qur'an surat As-Shaffat ayat 147 yang berbunyi:

وَأَرْسَلْنَٰهُ إِلَىٰ مِاْئَةِ أَلْفٍ أَوْ يَزِيدُونَ ﴿١٤٧﴾

Artinya: "*Dan kami utus dia kepada seratus ribu atau lebih"* (Qs. As-Shaffat/ 37:147).

Makna estimasi yang termaktup dalam ayat di atas yaitu terdapat pada potongan ayat yang artinya "*seratus ribu atau lebih*". Potongan ini mengandung unsur estimasi, karena dalam ayat ini Kami (Allah) tidak memberikan kepastian kepada Nabi Yunus mengenai jumlah kaum yang diutus kepada beliau. Di sini berarti hanya memperkirakan atau menduga-duga jumlah kaum Nabi Yunus. Hal ini sama dengan estimasi yang ada dalam statistika. Jika diketahui ada suatu model tertentu yang belum diketahui parameternya, maka untuk mencari nilai parameter tersebut dengan cara mengestimasi parameternya.

Kemudian disebutkan pula dalam hadits yang diriwayatkan oleh Bukhari tentang masalah estimasi yaitu *"sesungguhnya Rasulullah SAW* *memberi keringanan kepada mereka yang mempunyai ariyah untuk menjualnya dengan kira-kira (ditaksir)".* Arti kata mengira-ngira dalam hadits ini dalam statistika dapat diartikan sebagai estimasi. Hadits ini menyatakan bahwa Rasulullah memberi keringanan terhadap jual beli *ariyah*, yaitu jual beli kurma kering dan menukarnya dengan kurma basah dengan memperkirakan atau menaksir berapa banyak jumlahnya setelah kering. Jadi konsep *outlier* dan estimasi sudah termaktup dalam Al-Qur'an dan Hadits sebagaimana yang telah dijelaskan.

BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan pembahasan pada bab III, dapat disimpulkan bahwa metode statistik *Likelihod Displacement* (LD) mampu mendeteksi adanya *outlier* pada regresi nonlinier multiplikatif.

Sebelum menerapkan metode LD terlebih dahulu harus melinierkan model dengan asumsi bahwa *error* berdistribusi normal kemudian mengestimasi parameter regresi nonlinier multiplikatif dengan metode *maximum likelihood estimation*. Kemudian menerapkan metode statistik *likelihood displacement*, sehingga diperoleh hasil perumusan LD_{A_k} *likelihood displacement* untuk pengamatan yang diduga mengandung *outlier*.

Keakuratan metode LD dalam mendeteksi adanya *outlier* ditunjukkan dengan uji statistik. Yaitu dengan membandingkan MSE dari LD dengan MSE dari regresi pada umumnya. Statistik uji yang digunakan adalah Λ. Hipotesis awal ditolak ketika $\chi^2_{hitung} > \chi^2_{tabel}$, Sehingga terbukti LD_{A_k} adalah *outlier*.

4.2 saran

Dalam penelitian ini peneliti mendeteksi *outlier* pada regresi nonlinier multiplikatif, yang mana regresi nonlinier multiplikatif ini masih dapat ditransformasikan dalam bentuk linier. Oleh sebab itu disarankan untuk penelitian selanjutnya menggunakan regresi nonlinier yang tidak dapat ditranformasikan dalam bentuk linier.

DAFTAR PUSTAKA

Abdusysyakir. 2007. *Ketika Kyai Mengajar Matematika*. Malang : UIN-Malang Press.

Al-Asqolani, Ibnu hajar dan Al-Imam Al-Hafizh. 2007. *Fathul Baari penjelas Kitab Shahih Al-Bukhari (12).* Penj. Amiruddin. Jakarta: Pustaka azzam anggota IKAPI DKI.

Al-Maraghi, Ahmad Musthofa. 1989. *Tafsir Al-Maraghi*. Semarang: CV. Thoha Putra.

Al-Mahally, Imam Jalalud-din dan Imam Jalalud-din As-Suyuthi. 1990. *Terjemah Tafsir Jalalain Berikut Asbaabun Nuzul*. Bandung: Sinar Baru.

Amrullah, Abdulmalik Abdulkarim. 1981. *Tafsir Al-Azhar*. Surabaya: Yayasan Latimojong

Draper, Norman dan Harry Smith. 1992. *Analsis Regresi Terapan (edisi kedua)*. Jakarta: PT. Gramedia Pustaka Utama.

Ghoffur, Abdul, dkk. 2007. *Tafsir Ibnu Katsir (8)*. Bogor: Pustaka Imam Syafi'i.

Gujarati, Damodar N. 2007. *Dasar-dasar Ekonometri jilid 1 edisi ke-3*. Jakarta: Penerbit Erlangga.

Hasan, M.Iqbal. 2002. *Pokok-pokok materi metodologi penelitian dan aplikasinya*. Jakarta:Ghalia Indonesia.

Herrhyanto, Nar. 2007. http://www.Herryanto.blog/Statistika.Matematika.I.html (diunduh pada tanggal 26 januari 2012).

Makkulau, Susanti Linuwih, Purhadi, Muhammad Mashuri. 2010. *Pendeteksian Outlier dan Penentuan Faktor-Faktor yang Mempengaruhi Produksi Gula dan Tetes Tebu dengan Metode Likelihood Displacement Statistic-Lagrange*. Jurnal Teknik Industri, vol12.No.2 Desember.2010,95-100.

Murray dan Larry. 2007. Statistik edisi ke-3. Jakarta: Erlangga.

Mood, M Alexander dkk.1986. *Introduction to the Theory of Statistics*. McgrawHill Book Company.

Sembiring, RK. 1995. *Analisis Regresi*. Bandung: ITB.

Shihab, M Quraish. 2003. *Tafsir Al-Mishbah Volume* 14. Jakarta: Lentera Hati.

Sudjana. 2005. *Metoda Statistika*. Bandung: Penerbit Transito.

Turmudi dan Harini, Sri. 2008. *Metode Statistika Pendekatan Teoritis dan Aplikatif*. Malang: UIN-Press.

Wibisono, Yusuf. 2005. *Metode Statistik*. Yogyakarta: Gadjah Mada University Press.

Xu, Abraham, dan Steiner, 1998. *Outlier Detection Methods in Multivariate Regression Models*. Journal of Multivariate Analysis, 65, 1998, pp. 195-208.

Yitnosumarto, Sunyoto. 1990. *Dasar-Dasar Statistika*. Jakarta: Rajawali.

KEMENTERIAN AGAMA RI
UNIVERSITAS ISLAM NEGERI (UIN)
MAULANA MALIK IBRAHIM MALANG
FAKULTAS SAINS DAN TEKNOLOGI
Jl.Gajayana No. 50 Dinoyo Malang (0341) 551345 Fax. (0341) 572533

BUKTI KONSULTASI SKRIPSI

Nama : **Siti Tabi'atul Hasanah**
NIM : **08610045**
Fakultas/ Jurusan : **Sains dan Teknologi/ Matematika**
Judul : **Pendeteksian *Outlier* Pada Regresi Nonlinier Dengan Metode Statistik *Likelihood Displacement* (LD)**
Dosen pembimbing I : **Sri Harini, M.Si**
Dosen pembimbing II : **Dr. Ahmad Barizi, MA**

No.	Tanggal	Hal	Tanda tangan	
1	18 September 2011	Proposal Skripsi	1	
2	24 September 2011	ACC proposal Skripsi		2
3	19 November 2011	Konsultasi BAB I	3	
4	26 November 2011	Konsultasi BAB I dan BAB II		4
5	08 Desember 2011	Konsultasi BAB I dan BAB II keagamaan	5	
6	09 Desember 2011	Konsultasi BAB I, BAB II dan BAB III		6
7	12 Desember 2011	Konsultasi BAB I, BAB II dan BAB III	7	
8	21 Desember 2011	ACC BAB I dan BAB II		8
9	02 Januari 2012	Konsultasi BAB III	9	
10	07 Januari 2012	Konsultasi BAB III		10
11	12 Januari 2012	ACC BAB III	11	
12	12 Januari 2012	Konsultasi BAB III		12
13	12 Januari 2012	ACC keagamaan BAB III	13	
14	12 Januari 2012	ACC keseluruhan		14

Malang,13 Januari 2012

Mengetahui,
Ketua Jurusan Matematika,

Abdussakir, M.Pd
NIP: 19751006 200312 1 001