ANALISIS REGRESI *DUMMY VARIABLE* MODEL PROBIT

(Kasus pada Estimasi Hujan di Karangploso Malang)

SKRIPSI

oleh:
FITA ZUNAIDATUS SA'ADAH
NIM. 06510066

JURUSAN MATEMATIKA
FAKULTAS SAIN DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2011

ANALISIS REGRESI *DUMMY VARIABLE* MODEL PROBIT

(Kasus pada Estimasi Hujan di Karangploso Malang)

SKRIPSI

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

oleh:
FITA ZUNAIDATUS SA'ADAH
NIM: 06510066

JURUSAN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2011

ANALISIS REGRESI *DUMMY VARIABLE* MODEL PROBIT

(Kasus pada Estimasi Hujan di Karangploso Malang)

SKRIPSI

oleh :
FITA ZUNAIDATUS SA'ADAH
NIM. 06510066

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal : 11 Januari 2011

Pembimbing I,	Pembimbing II,
Abdul Aziz, M.Si	Dr. H. Ahmad Barizi, M A
NIP. 19760318 200604 1 002	NIP.19731212 199803 1 001

Mengetahui
Ketua Jurusan Matematika

Abdussakir, M. Pd
NIP. 19751006 200312 1 001

ANALISIS REGRESI *DUMMY VARIABLE* MODEL PROBIT

(Kasus pada Estimasi Hujan di Karangploso Malang)

SKRIPSI

oleh:
FITA ZUNAIDATUS SA'ADAH
NIM. 06510064

Telah Dipertahankan di Depan Dewan Penguji Skripsi dan
Dinyatakan Diterima Sebagai Salah Satu Persyaratan Untuk
Memperoleh Gelar Sarjana Sains (S.Si)
Tanggal: 28 Januari 2011

1. Penguji Utama	: Sri Harini, M. Si NIP. 19731014 200112 2 002	(	)
2. Ketua	: Drs. H. Turmudzi, M. Si NIP. 19571005 198203 1 006	(	)
3. Sekretaris	: Abdul Aziz, M.Si NIP. 19760318 200604 1 002	(	)
4. Anggota	: Dr. H. Ahmad Barizi, M.A NIP. 19731212 199803 1 001	(	)

Mengetahui dan Mengesahkan
Ketua Jurusan Matematika

Abdussakir, M.Pd
NIP. 19751006 200312 1 001

MOTTO

إِيَّاكَ نَعۡبُدُ وَإِيَّاكَ نَسۡتَعِينُ ٥

"Hanya kepada Engkaulah hamba menyembah dan hanya kepada Engkaulah hamba meminta pertolongan".

HALAMAN PERSEMBAHAN

Skripsi ini Penulis Persembahkan Teruntuk:
Kedua Orangtua Tercinta
Bapak Drs. Iskandar dan Ibu Aminatuz Zuhriyah
Kakak tersayang
Moh. Taufik Wazuhdi, S. Kom
Adik tersayang
Faidl Mabrurotul Khasanah dan Moh Khusna Ja'far Shodiq
yang tiada henti selalu memberikan dukungan dan do'a

Ananda terkasih
M. Makpudin, Amd.Kep
yang selalu memberikan dorongan, masukan dan do'a

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Fita Zunaidatus Sa'adah

NIM : 06510066

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, 11 Januari 2011
Yang membuat pernyataan,

Fita Zunaidatus Sa'adah
NIM. 06510066

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Segala puji bagi Allah SWT atas rahmat, taufik serta hidayah-Nya, sehingga penulis mampu menyelesaikan dalam penyusunan skripsi ini sebagai salah satu syarat untuk memperoleh gelar sarjana dalam bidang Matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.

Dalam proses penyusunan skripsi ini, penulis banyak mendapat bimbingan dan arahan dari berbagai pihak. Untuk itu ucapan terimakasih yang sebesar-besarnya dan penghargaan yang setinggi-tingginya penulis haturkan kepada semua yang terlibat dan telah membantu terselesaikannya skripsi ini, terutama kepada:

1. Prof. Dr. H. Imam Suprayogo selaku Rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Prof. Drs. Sutiman Bambang Sumitro, SU.D.Sc selaku Dekan Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Abdussakir, M.Pd selaku Ketua Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Drs.Usman Pagalay, M.Si selaku Dosen wali yang telah memberikan nasehat serta semangat kepada penulis selama menjalani perkuliahan.

5. Abdul Aziz, M.Si selaku pembimbing yang telah bersedia meluangkan waktu, tenaga, pikiran serta memberikan arahan dan masukan yang sangat berguna dalam menyelesaikan skripsi ini.
6. Dr. Ahmad Barizi, M.A selaku pembimbing agama yang telah bersedia memberikan pengarahan keagamaan dalam penyelesaian skripsi ini.
7. Segenap dosen Jurusan Matematika terima kasih atas semua ilmu yang telah diberikan.
8. Bapak dan Ibu tercinta yang tiada henti memberikan dukungan baik secara materil maupun moril serta senantiasa memberikan do'a dan restunya kepada penulis dalam menuntut ilmu.
9. Kakak dan adik tersayang yang telah banyak memberikan motivasi kepada penulis untuk menyelesaikan skripsi ini.
10. H. Saiful Munir, M. Pd terima kasih banyak atas ilmu-ilmu agama yang telah diberikan yang sangat berharga buat penulis.
11. Teman-teman matematika angkatan 2006 terima kasih atas segala pengalaman yang berharga dan kenangan indah yang dilewati bersama selama kuliah.
12. Teman-teman (Mufidatul Hasanah S, Fatimatuz Z, dan Siti Munawaroh) terima kasih banyak atas motivasi dan semangat yang diberikan.
13. Teman-teman IPNU-IPPNU PKPT UIN Maulana Malik Ibrahim Malang terima kasih banyak atas dukungan, semangat dan do'anya.
14. Semua pihak yang telah membantu hingga terselesaikannya skripsi ini, baik secara langsung maupun tidak langsung yang tidak dapat penulis sebutkan satu persatu.

Akhir kata penulis berharap semoga skripsi ini memberikan manfaat kepada para pembaca, khususnya kepada penulis secara pribadi. *Amin Yaa Rabbal ‘Alamin*.

Wabillahi taufik wal hidayah,

Wassalamu’alaikum Warahmatullahi Wabarakatuh.

Malang, 10 Januari 2011

Penulis

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR SIMBOL

Lambang Matematika

$\sim$	: Berdistribusi
$\leq$	: Lebih kecil atau sama dengan
$\geq$	: Lebih besar atau sama dengan
∞	: Tak berhingga
$<$	: Lebih kecil daripada
$>$	: Lebih besar daripada
$\prod$	: Pi untuk perkalian
$\sum$	: Sigma untuk penjumlahan

Abjad Yunani

α	: Alpha
β	: Bheta
γ	:Gamma
μ	: Mu
σ	: Sigma
π	: Pi
ϕ	: Phi
Φ	: PHI
∂	: Dho
ε	: Epsilon

Lambang Khusus

$\varnothing$	: Himpunan kososng
μ	: Nilai Tengah (rataan)
$\bar{X}$	: Rata-rata pada pengamatan X
$\bar{Y}$	: Rata-rata pada pengamatan Y
σ^2	: Ragam (varian) untuk populasi
$\mathbf{x}$	: Vektor yang entri-entrinya yang merupakan variabel independen
$\hat{\boldsymbol{\beta}}$	: Estimasi dari vektor β yang entri-entrinya terdiri dari parameter $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, ..., \hat{\beta}_n$
T	: *Transpose*
E	: *Expectation* (nilai harapan)
$X_1, X_2, ..., X_n$	: Peubah acak
$l(x_1, ..., x_n; \beta)$	: Fungsi *likelihood* dengan parameter β
$L(x_1, ..., x_n; \beta)$	: Fungsi *log likelihood* dengan parameter β
$f_x(x)$	: Fungsi padat peluang
$F_x(x)$	: Fungsi distribusi kumulatif
N	: Normal
$\chi^2_{\alpha,1}$	: Khi-Kuadrat dengan derajat bebas = 1

ABSTRAK

Sa'adah, Fita Z. 2011. **Analisis Regresi *Dummy Variable* Model Probit (Kasus pada Estimasi Hujan di Karangploso Malang)**. Skripsi. Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
Pembimbing : (I) Abdul Aziz, M.Si
(II) Dr.H. Ahmad Barizi, M.A

Kata kunci : estimasi parameter, regresi *dummy variable* model probit,*maxsimum likelihood*, estimasi hujan.

Salah satu solusi regresi dengan variabel terikat dummy adalah model probit. Model Probit menggunakan *Cumulative Distribution Function* dari fungsi normal. Dengan menggunakan metode *maxsimum likelihood* diharapkan dapat memperoleh nilai parameternya. Dan dapat menerapkan model Probit pada kasus perkiraan curah hujan tahun 2010-2012 di daerah Karangploso Malang.

Pada pencarian estimasi parameter model Probit dengan metode maksimum likelihood, ternyata tidak cukup. Karena model Probit *non liniear*, sehingga tidak dapat dihitung secara analitik. Untuk itu dibantu dengan metode numerik yaitu itersi *Newton-Raphson*. Dari hasil analisis diperoleh,

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}^{n+1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}^{n} -1 \left(n\bar{y} \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \frac{\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \left(\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right) \left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right) \left(\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \right) \right) - \left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right)^2}{\left(1 - \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right) \right)^2} \right)^{-1}$$

$$\left(\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i \left(\left(n\bar{y} \frac{\left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right)}{\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right)} \right) - \left(n\left(1-\bar{y}\right) \frac{\left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right)}{\left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right)} \right) \right) \right)$$

Hasil probabilitas musim hujan 2010-2012 di Karangploso Malang menunjukkan bahwa, pada tahun 2010 musim hujan terjadi pada bulan Januari, April-November, pada tahun 2011 musim hujan terjadi pada bulan April-September dan pada tahun 2012 musim hujan terjadi pada bulan Maret, Mei-Desember.

ABSTRACT

Sa'adah, Fita Z. 2011. **The Regression Analysis of Dummy Variable by Probit Model (Case Study of Rain Estimation at Karangploso Malang)**. Thesis. Mathematic Department, Faculty of Sience and Technology, The State Islamic University of Malang.
Advisor. (I) Abdul Aziz, M.Si
(II) Dr. H. Ahmad Barizi, M.A

Keyword : parameter estimation, regression dummy variable Probit model, Newton-Raphson maximum likelihood, rain estimassion.

One of regression solution of dummy variable bound is Probit model. Probit model uses Cumulative Distribution Function from normal function. By using maximum likelihood method is expexted to get the parameter value. It can applied the Probit model on calculation of heavy shower of rain on 2010-2012 at Karangploso Malang.

On existence of parameter estimation Probit model with maximum likelihood method, the fact that it is not enough. It is because Probit model is not linear. So, it is uncountable analytically. Because of that, it is helped by numeric method that is Newton-Raphson. From the analysis the resultare :

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}^{n+1} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}^{n} -1 \left(n\,\bar{y} \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \frac{\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \left(\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right) \left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right) \left(\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \right) \right) - \left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right)^2}{\left(1 - \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right) \right)^2} \right)^{-1}$$

$$\left(\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i \left(\left(n\,\bar{y} \frac{\left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right)}{\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right)} \right) - \left(n\left(1-\bar{y}\right) \frac{\left(\frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2} \right)}{\left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}} dz \right)} \right) \right) \right)$$

The probability result of rain period on 2010-2012 at Karang Ploso Malang show that on 2010 rain period on January, April-November, on 2011 rain period on April-September and on 2012 rain period on Marth, Mei-December.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Matematika dapat didefinisikan sebagai penelitian bilangan dan angka. Pada dasarnya segala sesuatu yang ada di alam semesta dapat di ilustrasikan ke dalam angka. Menurut Annemarie (2006:11) angka-angka mempengaruhi sifat sesuatu yang ditata dengannya, dengan demikian angka menjadi mediator antara Tuhan dan dunia ciptaan-Nya. Seperti angka 1, artinya tunggal, sendiri, tidak ada yang menyamai. Sifat ketunggalan ada pada Sang Pencipta yaitu Tuhan yang Maha Esa. Oleh karena itu ketika mendengar satu atau ketunggalan akan terpikir Tuhan Yang Maha Esa.

Penelitian angka telah dilakukan oleh berbagai kalangan, dalam dunia Islam pada abad ke-10 sekelompok *Isma'ili Ikhwan ash-Shafa* telah menerapkan angka dalam menentukan penanggalan bulan (*falaqiah*). Angka-angka juga diterapkan dalam berbagai bidang, termasuk bidang peramalan. (Annemarie, 2006:30)

Peramalan dalam ilmu matematika adalah estimasi suatu kejadian. Estimasi tidak memberikan jawaban yang pasti tentang apa yang terjadi, melainkan berusaha mencari pendekatan apa yang akan terjadi dimasa yang akan datang berdasarkan data yang diperoleh di masa sekarang.

Dalam surat Al-Hasyr ayat 18 dijelaskan:

يَٰٓأَيُّهَا ٱلَّذِينَ ءَامَنُواْ ٱتَّقُواْ ٱللَّهَ وَلۡتَنظُرۡ نَفۡسٞ مَّا قَدَّمَتۡ لِغَدٖۖ وَٱتَّقُواْ ٱللَّهَۚ إِنَّ ٱللَّهَ

خَبِيرُۢ بِمَا تَعۡمَلُونَ ١٨

Artinya:

"Hai orang-orang yang beriman, bertakwalah kepada Allah dan hendaklah setiap diri memperhatikan apa yang telah diperbuatnya untuk hari esok (akhirat); dan bertakwalah kepada Allah, sesungguhnya Allah Maha Mengetahui apa yang kamu kerjakan."

Pada lafadz " وَلتَنظُرْنَفْسٌ ماَّ قَدَّمَتْ لِغَدٍ " maksudnya adalah agar setiap orang memperhatikan apa-apa yang telah dikerjakan dari kebaikan dan kejelekan untuk persiapan dirinya kelak di hari kiamat. Syaikh Abu Bakar Jabir Al-Jazairi (2008:374) menjelaskan bahwa setiap jiwa hendaknya memperhatikan apa-apa yang telah diamalkannya (dilakukan) untuk menyongsong masa depan (yaitu hari kiamat). Jadi apa yang telah manusia lakukan saat ini merupan bekal untuk hari yang akan datang.

Estimasi biasa disebut dengan analisis regresi. Regresi mengemukakan tentang keingintahuan apa yang terjadi dimasa depan untuk memberi sumbangan menentukan keputusan yang terbaik. Regresi biasa dinyatakan dalam rumus $Y = \beta_0 + \beta_1 X$, dimana nilai dari variabel bebas (*X*) dan variabel terikat (*Y*) berupa kuantitatif (angka).

Seringkali dalam analisis regresi dijumpai variabel tak bebas tidak hanya dipengaruhi oleh variabel-variabel yang dapat dikuantitatifkan (seperti pendapatan, biaya dan harga), tapi juga oleh variabel yang pada dasarnya bersifat kualitatif (seperti jenis kelamin, pendidikan dan tempat tinggal). Dimana untuk menganalisis data kualitatif tersebut dapat diubah menjadi data kuantitatif dengan bantuan *dummy variable* (variabel boneka). Misalnya, disimbolkan dengan angka 1 untuk lali-laki dan 0 untuk perempuan.

Kenyataannya terdapat juga variabel tak bebas yang berupa data kualitatif, seperti contoh kepemilikan rumah berdasarkan jumlah pendapatan seseorang. Maka variabel tak bebasnya adalah memiliki rumah atau tidak dan variabel bebasnya adalah jumlah pendapatan seseorang. Dengan pemisalan, memiliki rumah disimbolkan dengan angka 1 dan tidak memiliki rumah disimbolkan dengan 0. Permasalahan seperti ini dapat diselesaikan dengan menggunakan model Probit.

Model Probit adalah model regresi untuk permasalahan variabel terikatnya berupa data kategorik, yaitu bernilai 1 atau 0. Model Probit biasa disebut dengan normit, model ini menggunakan fungsi distribusi kumulatif dari fungsi normal. (Supranto, 2004:331)

Konsep *dummy variable* juga diterapkan dalam Islam. Pada saat umat Islam menjalankan ibadah haji atau umrah, terdapat beberapa ketentuan yang harus dipenuhi. Dalam sebuah hadits yang diriwayatkan oleh Bukhari dan muslim dijelaskan :

وِعن كَعْبٍ بن عُجْرَةَ ،َرضِىَ الله عنه ، قال [حُمِلْتُ الى رسول الله صَلَّى الله

عَليْهِ وَسَلَّمَ والقَمْلُ يَتَنَاثَرُ على وَجْهِي، فَقَال : مَا كُنْتُ أُرَى الوَجَعَ بَلَغَ بِكَ مَا أَرَى، أَتَجِدُ شَاةً؟

فَقُلْتُ : لا، فَقَالَ: فَصُمْ ثَلاَ ثَةَ أَيَّامٍ، أَوْ أَطْعِمُ سِتَّةَ مَسَاكِيْنَ، لِكُلِّ مِسْكِيْنٍ نِصْفَ صَاعٍ]. مُتَّفَقٌ عَلَيْهِ

Artinya :

"Dari Ka'ab Bin 'Ujrah RA, ia berkata : Aku dibawa Rasulullah SAW dan kutu bertebaran di atas wajahku, lalu beliau bersabda : Aku tidak menyangka penyakit yang sampai kepadamu seperti yang kusaksikan, apakah engkau memiliki seekor kambing? Aku menjawab : tidak, Rasulullah bersabda, "Maka berpuasalah kamu tiga hari, atau member makan enam orang miskin, setiap orang setengah sha'. (HR. Muttafaq 'Alaih)

Pada hadits tersebut dijelaskan bahwa bila seseorang sakit atau berhalangan pada saat haji, maka wajib mengganti dengan membayar *dam* (denda) agar haji atau umrahnya tetap sah. *Dam* tersebut dapat berupa bersedekah, berpuasa dan berkurban (menyembelih hewan). Dari perbuatan dapat digantikan dengan nilai, dengan kata lain sesuatu yang kualitatif dapat dirubah menjadi kuantitatif yang merupakan sifat dari *dummy variable.*

Dari pemaparan latar belakang di atas, penulis tertarik untuk meneliti regresi dengan variabel terikat dummy dengan mengambil judul **"Analisis Regresi *Dummy Variable* Model Probit"**.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telang dipaparkan, rumusan masalah yang akan diambil adalah :

1. Bagaimana analisis regresi *dummy variable* model Probit ?
2. Bagaimana estimasi parameter regresi *dummy variable* model Probit ?

3. Bagaimana aplikasi regresi *dummy variable* model Probit pada estimasi hujan di Karangploso Malang untuk tahun 2010-2012 ?

1.3 Batasan Masalah

Penulis memberikan batasan masalah agar pembahasan tidak terlalu lebar, maka dalam hal ini penulis memberikan batasan-batasan sebagai berikut :

1. analisis model Probit pada penelitian ini hanya menjelaskankan bagaimana model probit diperoleh,
2. estimasi parameter menggunakan metode *Newton-Raphson nonlinear maxsimum likelihood*,
3. aplikasi regresi *dummy variable* model Probit pada estimasi hujan memiliki batasan sebagai berikut :
   a. data yang digunakan adalah data yang diambil dari BMKG (Badan Meteorologi Klimatologi dan Geofisika) untuk wilayah Karangploso Malang pada tahun 2007 sampai dengan 2009.
   b. Dalam penelitian ini yang diteliti adalah peluang terjadinya hujan (variabel Y), dengan ketentuan angka 1 dikategorikan musim hujan dan angka 0 dikategorikan tidak musim hujan, dimana dikategorikan sebagai musim hujan apabila curah hujan per bulan minimal 150 milimeter. Dengan variabel bebas X_1 adalah temperatur minimum harian (satuan ^oC) dan variabel bebas X_2 adalah kelembaban relatif harian (satuan %). $\beta_0, \beta_1, \beta_2$ adalah koefisien regresi dan ε adalah kesalahan pengganggu.

c. statistik uji dari parameter β yang digunakan dalam penelitian ini adalah uji Wald (uji signifikansi tiap-tiap parameter),

d. uji kebaikan model yang digunakan dalam penelitian ini adalah dengan melihat nilai dari AIC (*Akaike Information Criterion*) dan SC (*Schawarz Criterion*),

1.4 Tujuan Penelitian

Tujuan yang akan dicapai dalam skripsi ini adalah :

1. untuk mengetahui analisis regresi *dummy variable* model Probit
2. mengetahui estimasi parameter regresi *dummy variable* model Probit
3. mengetahui estimasi peluang hujan di Karangploso Malang untuk tahun 2010-2012.

1.5 Manfaat Penelitian

Penulisan skripsi ini diharapkan memberikan manfaat:

1. Bagi penulis
   a) Mengetahui lebih dalam tentang disiplin ilmu matematika terutama tentang Analisis Regresi *Dummy Variable*.
   b) Sebagai sarana belajar dan latihan untuk mengkaji suatu permasalahan khususnya dalam mengetahui perolehan model Probit dan estimasi dari model Probit dengan metode *Newton-Raphson nonlinear maxsimum likelihood* serta uji hipotesa model Probit.

c) Sebagai panduan untuk melakukan penelitian kualitatif dengan menggunakan model Probit.

2. Bagi pembaca
   a) Sebagai wawasan dan memperdalam pengetahuan terutama dalam bidang estimasi. Khususnya estimasi model Probit dengan metode *Newton-Raphson nonlinear maxsimum likelihood*.
   b) Sebagai bahan studi kasus, terutama bagi yang ingin melakukan penelitian sejenis. Juga menambah khasanah perpustakaan yang akan berguna bagi pembaca.

1.6 Metode Penelitian

Dalam skripsi ini digunakan dua pendekatan, yaitu studi literatur dan studi lapangan/kasus. Studi literatur digunakan dalam menganalisa regresi *dummy variable* dengan model Probit, dan langkah-langkah untuk menentukan estimasi parameter dari model probit. Studi lapangan digunakan untuk mengkaji estimasi hujan. Langkah-langkah penelitian ini adalah sebagai berikut:

a) Menjelaskan perolehan model dan estimasi parameter model Probit dengan langkah sebagai berikut:
   1. Mencari model regresi *dummy variable* model Probit.
   2. Menaksir parameter regresi dummy variabel model probit dengan metode *Newton-Raphson nonlinear maxsimum likelihood* dengan langkah-langkah sebagai berikut:

a. menentukan fungsi distribusi peluang *dummy variable* model Probit,

b. menentukan fungsi *Likelihood* dari fungsi distribusi peluang *dummy variable* model Probit,

c. menentukan fungsi *Log Likelihood* dari fungsi *Likelihood*,

d. memaksimumkan fungsi *Log Likelihood* dengan mendeferensialkan fungsi *Log Likelihood* terhadap parameter dan menyamakannya dengan nol,

e. selanjutnya mencari estimasi parameter dengan menggunakan iterasi *Newton-Raphson*.

b) Mengimplementasikan model probit pada estimasi hujan dengan langkah-langkah sebagai berikut:

1. Statistik deskriptif data (melihat tebaran data) dengan bantuan *software Minitab*.
2. Meregresikan data dengan menggunakan model Probit.
3. Menduga parameter regresi model probit dengan Eviews.
4. Menguji parameter regresi dengan menggunakan uji Wald.
5. Menguji kebaikan model dengan melihat nilai dari AIC dan SC
6. Menginterpretasi model.

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN. Berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian dan sistematika penulisan.

BAB II KAJIAN PUSTAKA. Berisi hal-hal yang mendasar dalam teori yang dikaji, meliputi: dasar pemikiran hujan dalam Islam, teori dasar probabilitas, fungsi distribusi kumulatif, regresi dengan *dummy variable*, estimasi parameter (metode *maximum likelihood* dan *Newton-Raphson nonlinear maxsimum likelihood*), uji hipotesis (uji Wald dan uji AIC, SC) dan curah hujan, temperatur, kelembaban.

BAB III PEMBAHASAN. Berisi pembahasan tentang perolehan model Probit yang berupa bentuk umum, estimasi parameter, pengujian model dan contoh penerapan.

BAB IV PENUTUP. Berisi kesimpulan akhir penelitian dan saran untuk pengembangan penelitian selanjutnya yang lebih baik.

BAB II

KAJIAN PUSTAKA

2.1 Dasar Pemikiran Hujan dalam Islam

Allah berfirman :

أَفَرَءَيْتُمُ ٱلْمَآءَ ٱلَّذِى تَشْرَبُونَ ﴿٦٨﴾ ءَأَنتُمْ أَنزَلْتُمُوهُ مِنَ ٱلْمُزْنِ أَمْ نَحْنُ
ٱلْمُنزِلُونَ ﴿٦٩﴾ لَوْ نَشَآءُ جَعَلْنَٰهُ أُجَاجًا فَلَوْلَا تَشْكُرُونَ ﴿٧٠﴾

Artinya :

"Maka apakah kamu melihat air yang kamu minum. Kamukah yang menurunkannya dari awan ataukah Kami Para Penurun(nya)?. Kalau Kami menghendaki, Kami menjadikannya asin, maka mengapakah kamu tidak bersyukur?"(QS.Al-Waqi'ah : 68-70)

Kata (المُزْن) *al-muzn* adalah bentuk jamak dari kata *al-muznah* yaitu awan yang mengandung air. Ada juga yang mengartikannya awan putih yang mengandung air (yaitu, air yang paling jernih dan sedap). Ayat ini mengisyaratkan bahwa tidak semua awan dapat mengakibatkan turunnya hujan, tetapi hanya awan tertentu yang mengandung lahirnya benih-benih. Pengunaan bentuk (المُنزِلُون) *al-munzilun* selain untuk menunjukkan kuasa dan kebesaran Allah swt, juga untuk mengisyaratkan bahwa ada malaikat yang ditugaskan Allah mengatur turunnya hujan, dan ada juga sistem dan hukum-hukum alam yang dapat dimanfaatkan manusia untuk maksud tersebut. (Shihab, 2003: 569)

Ayat di atas dikomentari oleh tim penyusun *Tafsir al-Muntakhab* bahwa: untuk terjadinya hujan diperlukan keadaan cuaca tertentu yang berada diluar kemampuan manusia, seperti adanya angin dingin yang berhembus di atas angin panas, atau keadaan cuaca yang tidak stabil. Adapun hujan buatan yang kita kenal itu sampai saat ini masih merupakan percobaan yang prosentase keberhasilannya masih sangat kecil dan masih memerlukan beberapa kondisi alam tertentu.

Selanjutnya Allah juga berfirman surat Ar-Ruum ayat 48, sebagai berikut :

ٱللَّهُ ٱلَّذِى يُرۡسِلُ ٱلرِّيَـٰحَ فَتُثِيرُ سَحَابًا فَيَبۡسُطُهُۥ فِى ٱلسَّمَآءِ كَيۡفَ يَشَآءُ وَيَجۡعَلُهُۥ كِسَفًا فَتَرَى ٱلۡوَدۡقَ يَخۡرُجُ مِنۡ خِلَـٰلِهِۦ ۖ فَإِذَآ أَصَابَ بِهِۦ مَن يَشَآءُ مِنۡ عِبَادِهِۦٓ إِذَا هُمۡ يَسۡتَبۡشِرُونَ ﴿٤٨﴾

Artinya :

"Allah, Dialah yang mengirim angin, lalu angin itu menggerakkan awan dan Allah membentangkannya di langit menurut yang dikehendaki-Nya, dan menjadikannya bergumpal-gumpal; lalu kamu Lihat hujan keluar dari celah-celahnya, Maka apabila hujan itu turun mengenai hamba-hamba-Nya yang dikehendakiNya, tiba-tiba mereka menjadi gembira."

Dari ayat tersebut dapat dijelaskan bahwa proses fenomena hujan terjadi dengan beberapa tahapan, yaitu :

TAHAP KE-1: *"Dialah Allah Yang mengirimkan angin..."*

Gelembung-gelembung udara yang jumlahnya tak terhitung yang dibentuk dengan pembuihan di lautan, pecah terus-menerus dan menyebabkan partikel-partikel air tersembur menuju langit. Partikel-partikel ini, yang kaya akan garam, lalu diangkut oleh angin dan bergerak ke atas di atmosfir. Partikel-partikel ini, yang disebut aerosol, membentuk awan dengan mengumpulkan uap air di sekelilingnya, yang naik lagi dari laut, sebagai titik-titik kecil dengan mekanisme yang disebut "perangkap air". Hal ini yang akan meyebabkan terjadinya kelembaban pada udara.

TAHAP KE-2: *"...lalu angin itu menggerakkan awan dan Allah membentangkannya di langit menurut yang dikehendaki-Nya, dan menjadikannya bergumpal-gumpal..."*

Awan-awan terbentuk dari uap air yang mengembun di sekeliling butir-butir garam atau partikel-partikel debu di udara. Karena air hujan dalam hal ini sangat kecil (dengan diamter antara 0,01 dan 0,02 mm), awan-awan itu bergantungan di udara dan terbentang di langit. Jadi, langit ditutupi dengan awan-awan.

TAHAP KE-3: *"...lalu kamu lihat air hujan keluar dari celah-celahnya..."*

Partikel-partikel air yang mengelilingi butir-butir garam dan partikel-partikel debu itu mengental dan membentuk air hujan. Jadi, air hujan ini, yang menjadi lebih berat daripada udara, bertolak dari awan dan mulai jatuh ke tanah sebagai hujan.

Semua tahap pembentukan hujan telah diceritakan dalam ayat-ayat Alqur'an. Selain itu, tahap-tahap ini dijelaskan dengan urutan yang benar. Sebagaimana fenomena-fenomena alam lain di bumi, lagi-lagi Alqur'an lah yang menyediakan penjelasan yang paling benar mengenai fenomena ini dan juga telah mengumumkan fakta-fakta ini kepada orang-orang pada ribuan tahun sebelum ditemukan oleh ilmu pengetahuan.

2.2 Teori Dasar Probabilitas

Teori peluang sering disebut dengan teori kemungkinan yang merupakan konsep dasar dari ilmu statistika. Dalam kehidupan sehari-hari sering dijumpai suatu kejadian kadang terjadi atau tidak terjadi. Terjadinya suatu peristiwa tersebut mempunyai tingkat yang berbeda-beda, ada yang kemungkinan terjadinya besar dan ada yang kemungkinan terjadinya kecil. Pernyataan kemungkinan ini menunjukkan ukuran ketidakpastian. Kemungkinan yang menyangkut ketidakpastian ini dinamakan peluang (Harini, 2007:55-56).

Bain dan Engelhard (1991:9) mendefinisikan peluang sebagai berikut:
Jika sebuah percobaan E mempunyai ruang sampel S dan sebuah kejadian A didefinisikan pada S, maka $P(A)$ adalah suatu angka riil yang disebut probabilitas dari peristiwa A atau probabilitas A. Dan fungsi $P(.)$ mempunyai syarat-syarat sebagai berikut:

1. $0 \leq P(A) \leq 1$ untuk setiap kejadian A dari S
2. $P(\varnothing) = 0$
3. $P(S) = 1$

Menurut Dudewicz dan Mishra (1995:19) bila suatu percobaan yang dapat menghasilkan n macam hasil yang berkemungkinan sama dan bila terdapat sebanyak $n(A)$ dari hasil yang berkaitan dengan kejadian A, maka probabilitas A adalah:

$$P(A) = \frac{n(A)}{n} \tag{2.1}$$

dimana:

$n(A)$ = jumlah hasil yang termasuk kejadian A

n = jumlah keseluruhan hasil

Menurut Gujarati (2007:27) peluang terjadiya suatu kejadian B bila diketahui kejadian A telah terjadi disebut peluang bersyarat (*conditional* *probability*) dan dinyatakan dengan $P(B|A)$. Lambang $P(B|A)$ biasanya dibaca "peluang B terjadi bila diketahui A terjadi", atau lebih sederhana lagi "peluang B, bila A terjadi", yang dapat dituliskan sebagai berikut:

$$P(B|A) = \frac{P(AB)}{P(B)}; P(A) > 0. \tag{2.2}$$

Walpole dan Myers (1995:38) menyatakan dalam teorema bahwa :

Dua kejadian A dan B bebas jika dan hanya jika,

$$P(AB) = P(A)P(B) \tag{2.3}$$

Bukti :

$$\begin{aligned} P(AB) &= \frac{n(AB)}{n} \\ &= \frac{n(A).n(B)}{n} \\ &= \frac{n(A)}{n}.\frac{n(B)}{n} \\ &= P(A)P(B) \end{aligned} \qquad \begin{aligned} P(A)P(B) &= \frac{n(A)}{n}.\frac{n(B)}{n} \\ &= \frac{n(A).n(B)}{n} \\ &= \frac{n(AB)}{n} \\ &= P(AB) \end{aligned}$$

Dengan demikian apabila A dan B kejadian bebas dan bila $P(A) > 0$, maka

$$P(B|A) = P(B).$$

Bukti :

Karena A dan B kejadian bebas, $P(AB) = P(A)P(B)$ dan menurut definisi Persamsaan (2.2) $P(B|A) = \dfrac{P(AB)}{P(A)}$. Maka,

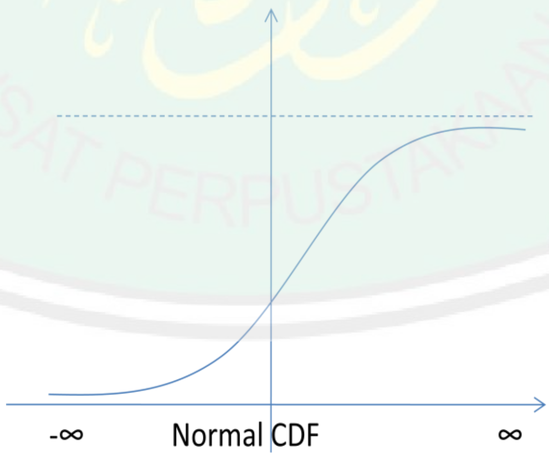
$$\begin{aligned} P(B|A) &= \frac{P(AB)}{P(A)} \\ &= \frac{P(A)P(B)}{P(A)} \\ &= P(B). \end{aligned}$$

2.3 Fungsi Distribusi Kumulatif

Dudewicz dan Mishra (1995:149) mendefinisikan fungsi distribusi kumulatif atau *Cumulative Distribution Function (CDF)* sebagai berikut :

Fungsi distribusi kumulatif atau probabilitas kumulatif sering disebut fungsi distribusi saja. Fungsi distribusi variabel acak kontinu X yang dinotasikan $F_x(x) = P(X \le x)$ untuk semua bilangan riil x, didefinisikan dengan:

$$F_x(x) = \int_{-\infty}^{x} f(t)\,dt. \tag{2.4}$$



Gambar 2.1 Kurva CDF Normal

Salah satu distribusi yang penting dalam statistik adalah adalah distribusi normal. Menurut Gujarati (2007:68) variabel acak X yang berdistribusi normal biasanya di notasikan dengan $X \sim N\left(\mu_x, \sigma_x^2\right)$. Dimana $\sim$ berarti didistribusikan sebagai, N menyatakan distribusi normal, dan notasi di dalam tanda kurung menyatakan *parameter* distribusi, yakni nilai rata-rata atau nilai harapan (populasi)nya, μ_x , dan variansinya, σ_x^2. Dengan X merupakan variabel acak yang bersifat kontinu dan dapat memiliki nilai antara $-\infty$ sampai ∞. Sedangkan Dudewicz dan Mishra (1995:169) mendefinisikan distribusi normal sebagai berikut:

Suatu peubah acak X dikatakan berdistribusi normal $N\left(\mu_x, \sigma_x^2\right)$ bila (untuk suatu $\sigma^2 > 0$ dan $-\infty < \mu < \infty$) berlaku

$$f_x\left(x\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty. \tag{2.5}$$

Fungsi $f_x\left(x\right)$ pada Persamaan (2.5) menunjukkan nilai *Probability Density Function (PDF)* dari distribusi normal. Sehingga fungsi distribusi kumulatif atau *Cumulative Distribution Function (CDF)* dari distribusi normal dapat dinyatakan sebagai berikut:

$$F_x\left(x\right) = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \tag{2.6}$$

Dari Persamaan (2.6) dapat ditransformasikan ke dalam normal baku yang dinyatakan dengan Z. Dalam teorema, Bain (1991:121) menyatakan:

Jika $X \sim N\left(\mu_x, \sigma_x^2\right)$, maka:

1. $Z = \dfrac{X-\mu}{\sigma} \sim N\left(0,1\right)$

2. $F_x(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)$

Bukti:

$$\begin{aligned}F_z(z)&=P[Z\le z]\\&=P\left[\frac{X-\mu}{\sigma}\le z\right]\\&=P[X\le \mu+z\sigma]\\&=\int_{-\infty}^{\mu+z\sigma}\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]dx\end{aligned}$$

kemudian substitusikan $w=\frac{(x-\mu)}{\sigma}$, maka diperoleh:

$$\begin{aligned}F_z(z)&=\int_{-\infty}^{\mu+z\sigma}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2}\,dw\\&=\Phi(z).\end{aligned} \tag{2.7}$$

Simbol $\Phi(z)$ dinotasikan untuk standar normal kumulatif atau *Cumulative Distribution Function* (CDF) yang berdistribusi normal. Dan turunan dari standar normal kumulatif disebut standar normal *Probability Density Function (PDF)* yang dinyatakan sebagai berikut :

$$\begin{aligned}f_z(z)&=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2}\\&=\phi(z).\end{aligned} \tag{2.8}$$

2.4 Regresi dengan *Dummy Variable*

Persaman regresi, biasanya menggunakan simbol Y untuk variabel tak bebas (*dependent variable*) dan X variabel bebas (*independent variable*). Variabel X bisa lebih dari satu (*multivariate*). Baik X maupun Y bisa berupa variabel kualitatif (Nachrowi, 2004: 167).

Variabel dalam persamaan regresi yang sifatnya kualitatif ini biasanya menunjukkan ada tidaknya (*presence or absence*) suatu "*quality*" atau suatu "*atribute*", misalnya laki–laki atau perempuan, Jawa atau luar Jawa, sarjana atau bukan, sudah menikah atau masih membujang dan sebagainya. Salah satu metode untuk membuat *kuantifikasi* (berbentuk angka) dari data kualitatif (tidak berbentuk angka) adalah dengan membentuk *variable-variable artificial* yang memperhitungkan nilai-nilai 0 atau 1, 0 menunjukkan ketiadaan sebuah atribut dan 1 menujukkan keberadaan (kepemilikan) atribut itu. Misalnya, 1 mungkin menunjukkan bahwa seseorang adalah wanita dan 0 mungkin menunjukkan laki-laki, atau 1 mungkin menunjukkan bahwa seseorang adalah sarjana dan 0 menunjukkan bahwa seseorang bukan seorang sarjana. Variabel-variabel yang mengasumsikan nilai-nilai seperti 0 dan 1 ini disebut dengan *dummy variable* (variabel buatan) (Gujarati, 2007:1).

Menurut Supranto (2004:175) *dummy variable* disebut juga variabel indikator, biner, kategorik, kualitatif, boneka atau variabel dikotomi. Suatu persamaan regresi tidak hanya menggunakan variabel kategorik sebagai variabel bebas, tetapi dapat pula disertai oleh variabel bebas lain yang numerik. Persamaan regresi dengan variable bebas berupa dummy dapat dituliskan sebagai berikut :

$$Y = \beta_1 + \beta_2 D + \varepsilon \qquad (2.9)$$

dimana:

Y = variabel terikat

D = variabel dummy sebagai variabel bebas yang bernilai 1 atau 0

ε = kesalahan random

Variabel dummy bisa saja digunakan pada variabel tak bebas (Y), sehingga Y bernilai 0 atau 1, yang memiliki arti ya atau tidak (bersifat dikotomi). Misalkan pada penelitian partisipasi angkatan kerja pria dewasa sebagai fungsi tingkat pengangguran, pendapatan keluarga, tingkat pendidikan dan lain-lain. Seseorang bisa berada di dalam atau di luar angkatan kerja. Jadi keberadaan orang ini di dalam atau di luar angkatan kerja cuma memiliki dua nilai saja : 1 jika orang ini ada dalam angkatan kerja dan 0 jika tidak. Persamaan model ini dapat ditulis:

$$Y = \beta_1 + \beta_2 X + \varepsilon. \qquad (2.10)$$

Model Persamaan (2.10) terlihat seperti regresi linear pada umumnya, tapi teryata bukan, karena koefisien kemiringan β_2 yang menunjukkan tingkat perubahan Y untuk setiap perubahan unit X tidak dapat ditafsirkan, karena Y hanya menggunakan dua nilai, 1 dan 0. Maka persamaan (2.7) disebut dengan model probabilitas linier (LPM, *Linear Probability Model*) karena ekspetasi bersyarat Y bila X diketahui, $E(Y|X)$, bisa ditafsirkan sebagai probabilitas bersyarat, mengingat kejadian tersebut akan terjadi bila X diketahui, yakni $P(Y=1|X)$ (Gujarati, 2007: 21).

2.5 Estimasi Parameter

2.5.1 Pengertian Estimasi Parameter

Dalam statistik, estimasi (penaksiran) adalah suatu metode untuk mengetahui sekitar nilai-nilai suatu populasi dengan menggunakan nilai-nilai sampel. Dalam kasus sebuah variabel acak X di asumsikan berdistribusi normal dengan dua parameternya nilai rata-rata (μ_x) dan varians (σ_x^2), yang mana nilai dari kedua parameter ini tidak diketahui. Untuk menaksir nilai parameter yang tidak diketahui ini, dapat diasumsikan terdapat sampel acak sebesar *n* dari distribusi probabilitas yang diketahui dan menggunakan sampel tersebut untuk menaksir parameter yang tidak diketahui. Jadi, rata-rata sampel dapat dijadikan sebagai taksiran atas rata-rata populasi dan varians sampel sebagai taksiran atas varians populasi (Gujarati, 2007: 91).

Prinsip penggunaan metode estimasi pada sebuah observasi ke-*i*, dengan persamaan regresi $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, dapat diperoleh nilai error sebagai berikut :

$$\varepsilon_i = Y_i - \beta_0 - \beta_1 X_i \tag{2.11}$$

Sehingga dapat diperoleh rata-rata error ke-*n* dari sampel sebagai berikut:

$$\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \beta_0 - \beta_1 X_i\right) \tag{2.12}$$

diasumsikan nilai rata-rata error ke-*n* adalah nol, sehingga berakibat nilai parameter $\beta_1 = 0$. Karena pada model ini hanya memiliki satu parameter, yaitu β_0, β_0 pada Persamaan (2.12) sama dengan nol, maka diperoleh:

$$\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{\beta}_0\right) = 0. \tag{2.13}$$

Dan karena nilai β_0 tidak terikat indek *i*. Persamaan (2.13) dapat ditulis:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{\beta}_0 = 0$$

$$\hat{\beta}_0 = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{2.14}$$

$$= \bar{Y}$$

dimana $\hat{\beta}_0$ adalah estimasi dari β_0. Estimasi ini hanya rata-rata dari nilai observasi variabel bebas X_i (Davidson, 1999:32-33).

2.5.2 Metode *Maximum Likelihood*

Metode *maxsimum likelihood* adalah suatu penaksir titik yang mempunyai sifat teoritis yang lebih kuat dibandingkan dengan metode penaksir kuadrat terkecil. Metode *maxsimum likelihood* merupakan salah satu cara untuk mengestimasi parameter yang tidak diketahui. Prosedur estimasi *maxsimum likelihood* menguji apakah estimasi *maxsimum* yang tidak diketahui dari fungsi *likelihood* suatu sampel nilainya sudah memaksimumkan fungsi *likelihood* (Gujarati, 2004: 112).

Menurut Greene (2002: 468-469) fungsi PDF (*Probability Density Function*) dari variabel *y* acak dengan parameter β, dinotasikan $f(y\,|\,\beta)$. Probabilitas sampel random dari *joint* PDF untuk $y_1, y_2, \ldots, y_n$ (dimana *n* saling bebas dan berdistribusi sama) dapat dihitung:

$$f(y_1,\cdots,y_n\,|\,\beta)=\prod_{i=1}^{n} f(y_i\,|\,\beta)=L(\beta\,|\,y). \tag{2.15}$$

Metode *maxsimum likelihood* akan memilih nilai β yang diketahui sedemikian hingga memaksimumkan nilai probabilitas dari gambaran sampel secara acak yang telah diperoleh secara aktual. Fungsi *log likelihood*-nya adalah :

$$L(\beta\,|\,y)=\ln l(\beta\,|\,y)=\sum_{i=1}^{n}\ln f(y_i\,|\,\beta). \tag{2.16}$$

Menurut Davidson dan Mackinnon (1999:32-33) bila fungsi *likelihood* terdefernsialkan terhadap β, maka estimasi *maxsimum likelihood* dapat diperoleh melalui persamaan berikut:

$$\frac{\partial l\left(x_1,x_2,\ldots,x_n\right)}{\partial\beta_i} \tag{2.17}$$

untuk $i=1,2,\ldots,n.$

Dalam banyak kasus, penggunaan deferensiasi akan lebih mudah bekerja pada logaritma natural dari $l\left(x_1,x_2,\ldots,x_n\,|\,\beta\right)$, yaitu:

$$L\left(x_1,x_2,\ldots,x_n\,|\,\beta\right)=\ln l\left(x_1,x_2,\ldots,x_n\,|\,\beta\right). \tag{2.18}$$

2.5.3 Newton-Raphson Nonlinear Maximum Likelihood

Menurut Judge (1985:497) secara umum, iterasi untuk mendapatkan taksiran β dengan *nonlinear maximum likelihood* dapat ditulis dengan :

$$\beta^{n+1}=\beta^{n}-t_n\,P_n\,\gamma_n \tag{2.19}$$

dimana

$$t_n(koefisienpengali)=1,\ P_n=\left(\frac{\partial^2 L}{\partial\beta\partial\beta'}|_{\beta^n}\right)^{-1},\ \gamma_n=\frac{\partial L}{\partial\beta}|_{\beta^n} \quad (2.20)$$

Nilai n menunjukkan iterasi ke-n. Fungsi distribusi peluang (pdf) dari y_i diberikan oleh X_i, β dan σ^2 adalah

$$f\,(y_i\,|X_i,\beta,\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(y_i-f\,(X_i,\beta))^2\right) \quad (2.21)$$

dimana $f\,(X_i,\beta)$ adalah fungsi nonlinier yang bersifat umum.

Likelihood function dari β dan σ^2 diberikan oleh y_i dan X_i adalah

$$l_i\,(\beta,\sigma^2)=(2\pi\sigma^2)^{-1/2}\exp\left(-\frac{1}{2\sigma^2}(y_i-f\,(X_i,\beta))^2\right) \quad (2.22)$$

Log likelihood function dari β dan σ^2 diberikan oleh y_i dan X_i adalah

$$\log\ l_i=L_i=-\frac{1}{2}(\log(2\pi)+\log(\sigma^2))-\frac{1}{2\sigma^2}(y_i-f\,(X_i,\beta))^2. \quad (2.23)$$

P.d.f gabungan dari $(y_1,...,y_i)$ diberikan oleh X , β dan σ^2 adalah

$$\begin{aligned} f\left(y_1,...,y_n\,|X,\beta,\sigma^2\right)&=(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-f\,(X_i,\beta))^2\right)\\ &=(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}(y-f\,(X_i,\beta))^T\,(y-f\,(X_i,\beta))\right) \end{aligned} \quad (2.24)$$

Likelihood function dari β dan σ^2 diberikan X dan y adalah

$$l\left(\beta,\sigma^2\right)=(2\pi\sigma^2)^{-n/2}\exp\left(-\frac{1}{2\sigma^2}(y-f\,(X_i,\beta))^T\,(y-f\,(X_i,\beta))\right) \quad (2.25)$$

dan *log likelihood function* dari β dan σ^2 diberikan X dan y adalah

$$
\begin{aligned}
L &= \sum_{i=1}^{n} L_t \\
&= \log\, l\left(\beta,\sigma^2\right) \\
&= -\frac{n}{2}\left(\log(2\pi)+\log\left(\sigma^2\right)\right)-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta) \\
&= -\frac{n}{2}\left(\log(2\pi)+\log\left(\sigma^2\right)\right)-\frac{1}{2\sigma^2}S(\beta).
\end{aligned} \tag{2.26}
$$

Maka

$$\frac{\partial L}{\partial \sigma^2}=\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}S(\beta). \tag{2.27}$$

dimana $S(\beta)$ adalah fungsi jumlah kuadrat dari error. Selanjutnya menyamakan Persamaan (2.27) dengan nol akan diperoleh

$$\hat{\sigma}^2=\frac{S(\beta)}{n}. \tag{2.28}$$

Sekarang Persamaan (2.26) dapat ditulis menjadi :

$$
\begin{aligned}
L(\beta) = L &= -\frac{n}{2}\left(\log(2\pi)+\log\left(\frac{S(\beta)}{n}\right)\right)-\frac{1}{2\left(\frac{S(\beta)}{n}\right)}S(\beta) \\
&= -\frac{n}{2}\left(\log(2\pi)+\log\left(\frac{S(\beta)}{n}\right)\right)-\frac{n}{2}
\end{aligned} \tag{2.29}
$$

Aproksimasi $L(\beta)$ di sekitar β^1 dengan deret Taylor orde 2, yaitu :

$$L(\beta)=L\left(\beta^{(1)}\right)+\frac{\partial L}{\partial \beta^T}\bigg|_{\beta^{(1)}}\left(\beta-\beta^{(1)}\right)+\frac{1}{2}\left(\beta-\beta^{(1)}\right)^T\frac{\partial^2 L}{\partial\beta\partial\beta^T}\bigg|_{\beta^{(1)}}\left(\beta-\beta^{(1)}\right), \tag{2.30}$$

diperoleh

$$\frac{\partial L(\beta)}{\partial \beta}=\left(\frac{\partial L}{\partial \beta^T}\bigg|_{\beta^{(1)}}\right)^T+\frac{\partial^2 L}{\partial\beta\partial\beta^T}\bigg|_{\beta^{(1)}}\left(\beta-\beta^{(1)}\right). \tag{2.31}$$

Karena

$$\left(\frac{\partial L}{\partial \beta^T} \Big|_{\beta^{(1)}} \right)^T = \frac{\partial L}{\partial \beta} \Big|_{\beta^{(1)}}$$

maka

$$\frac{\partial L(\beta)}{\partial \beta} = \frac{\partial L}{\partial \beta} \Big|_{\beta^{(1)}} + \frac{\partial^2 L}{\partial \beta \partial \beta^T} \Big|_{\beta^{(1)}} \left(\beta - \beta^{(1)} \right) \tag{2.32}$$

Menyamakan persamaan (2.32) dengan nol, akan diperoleh

$$\frac{\partial L}{\partial \beta} \Big|_{\beta^{(1)}} + \frac{\partial^2 L}{\partial \beta \partial \beta^T} \Big|_{\beta^{(1)}} \left(\beta^{(2)} - \beta^{(1)} \right) = 0$$

atau

$$\beta^{(2)} = \beta^{(1)} - \left(\frac{\partial^2 L}{\partial \beta \partial \beta^T} \Big|_{\beta^{(1)}} \right)^{-1} \frac{\partial L}{\partial \beta} \Big|_{\beta^{(1)}}$$

Secara umumnya diperoleh iterasi

$$\begin{aligned} \beta^{(n+1)} &= \beta^{(n)} - \left(\frac{\partial^2 L}{\partial \beta \partial \beta^T} \Big|_{\beta^{(n)}} \right)^{-1} \frac{\partial L}{\partial \beta} \Big|_{\beta^{(n)}} \\ &= \beta^{(n)} - \left(-\frac{1}{2\sigma^2} \frac{\partial^2 S}{\partial \beta \partial \beta^T} \Big|_{\beta^{(n)}} \right)^{-1} \left(-\frac{1}{2\sigma^2} \frac{\partial S}{\partial \beta} \Big|_{\beta^{(n)}} \right) \\ &= \beta^{(n)} - \left(\frac{\partial^2 L}{\partial \beta \partial \beta^T} \Big|_{\beta^{(n)}} \right)^{-1} \left(\frac{\partial S}{\partial \beta} \Big|_{\beta^{(n)}} \right) \end{aligned} \tag{2.33}$$

2.6 Uji Hipotesis

2.6.1 Pengertian Uji Hipotesis

Yang dimaksud dengan pengujian hipotesis adalah salah satu cara dalam statistika untuk menguji ‘parameter’ populasi berdasarkan statistik sampelnya, untuk dapat diterima atau ditolak pada tingkat signifikansi tertentu. Pada prinsipnya pengujian hipotesis ini adalah membuat kesimpulan sementara untuk melakukan penyanggahan atau pembenaran dari permasalahan yang akan ditelaah. Sebagai wahana untuk menetapkan kesimpulan sementara tersebut kemudian ditetapkan hipotesis nol dan hipotesis alternatifnya (Supangat, 2008:293).

Hipotesis nol (H_0) untuk memprediksi bahwa variabel bebas tidak mempunyai efek pada variabel terikat dalam populasi. H_0 juga untuk memprediksi tidak adanya perbedaan antara suatu kondisi dengan kondisi yang lain. Sedangkan hipotesis alternatif, biasa dilambangkan dengan H_1, yang memprediksi bahwa variabel bebas mempunyai efek pada variabel terikat dalam populasi. H_1 juga untuk memprediksi adanya perbedaan antara suatu kondisi dengan kondisi yang lainnya (Irianto, 2006:97-98).

Menurut Supangat (2008:294) pernyataan hipotesis nol ini merupakan dugaan terhadap parameter suatu permasalahan yang akan dilakukan kajian untuk membenarkan atau menyanggah informasi dari suatu populasinya, berdasarkan statistik sampel pada tingkat signifikansi tertentu. Ada beberapa pengertian dalam pelaksanaan pengujian hipotesis, diantaranya:

- Tingkat signifikansi / taraf nyata (α)

  Tingkat signifikansi (taraf nyata) adalah luas daerah di bawah kurva yang merupakan daerah penolakan hipotesis nolnya.

- Tingkat keyakinan / tingkat kepercayaan $(1-\alpha)$

Tingkat keyakinan (tingkat kepercayaan) adalah luas daerah di bawah kurva yang merupakan daerah penerimaan hipotesis nolnya.

2.6.2 Uji Wald

Pengujian hipotesa dapat dilakukan dengan metode Uji Wald, yaitu uji signifikansi tiap-tiap parameter.

$H_0 : \beta_j = 0$ untuk suatu j tertentu ; j = 0,1,…,p

$H_1 : \beta_j \neq 0$

Statistik uji yang digunakan adalah

$$W_j = \left[\frac{\tilde{\beta}_1}{SE(\beta_1)} \right]; \; j = 0,1,2,\ldots.p. \qquad (2.34)$$

Statistik ini berdistribusi Khi-Kuadrat dengan derajat bebas 1 atau secara simbolis ditulis $W_j \sim \chi^2$. H_0 ditolak jika $W_j > \chi^2_{\alpha,1}$; dengan α adalah tingkat signifikansi yang dipilih. Bila H_0 ditolak, artinya parameter tersebut signifikan secara statistik pada tingkat signifikansi α (Nachrowi, 2004: 256).

2.6.3 Uji Kebaikan Model AIC Dan SC

Seorang ahli statistik dari Jepang, professor Hirotugu Akaike, pada tahun 1974 mengusulkan suatu metode untuk menguji ketepatan suatu model. Suatu metode tersebut yang kemudian dinamakan sebagai *Akaike Information Criterion* (AIC) dengan rumus :

$$AIC = \frac{RSS}{n} \exp\left(\frac{2k}{n}\right) \tag{2.35}$$

Persamaan (2.35) mengandung bilangan *e* dan *residual sum square* (RSS) sehingga dapat diubah ke dalam bentuk logaritma sebagai berikut:

$$\ln AIC = \ln\left(\frac{1}{n}\sum_{i=1}^{n}\hat{e}_i^{\,2}\right) + \frac{2k}{n} \tag{2.36}$$

dimana k adalah banyaknya variabel independen dan n adalah banyaknya observasi (Wahyu, 2007: 4.6).

Sedangkan *Schawarz Criterion* (SC) juga digunakan untuk menilai kualitas suatu model dengan rumus,

$$SC = \frac{RSS}{n} n^{k/n} \tag{2.37}$$

Persamaan (2.37) mengandung fungsi eksponensial dan *residual sum square* (RSS) sehingga dapat diubah ke dalam bentuk logaritmanya sebagai berikut :

$$\log SC = \log\left(\frac{1}{n}\sum_{i=1}^{n}\hat{e}_i^{\,2}\right) + \frac{k}{n}\log n \tag{2.38}$$

dimana k adalah banyaknya variabel independen dan n adalah banyaknya observasi. Kedua metode pengujian model di atas dapat digunakan salah satu atau gabungan untuk satu model regresi. Semakin kecil nilai AIC dan SC, semakin baik pula model tersebut (Wahyu, 2007: 4.21).

2.7 Curah Hujan, Temperatur Udara, dan Kelembapan Udara

2.7.1. Curah Hujan

Curah hujan adalah jumlah air hujan yang turun pada suatu daerah dalam waktu tertentu. Curah hujan normalnya berkisar 150 milimeter (Siwi Tri Puji B, 2010). Alat untuk mengukur banyaknya curah hujan disebut *Rain Gauge*. Curah hujan diukur dalam harian, bulanan, dan tahunan.

Curah hujan yang jatuh di wilayah Indonesia dipengaruhi oleh beberapa faktor antara lain:

1) Bentuk medan atau topografi;

2) Arah lereng medan;

3) Arah angin yang sejajar dengan garis pantai;

4) Jarak perjalanan angin di atas medan datar.

Hujan adalah butiran-butiran air yang dicurahkan dari atmosfer turun ke permukaan bumi (Soko, 2009).

2.7.2. Temperatur Udara

Temperatur udara adalah derajat panas dari aktivitas molekul dalam atmosfer. Alat untuk mengukur temperatur udara atau disebut *Thermometer*. Biasanya pengukuran temperatur udara dinyatakan dalam skala *Celcius* (C), *Reamur* (R), dan *Fahrenheit* (F). Udara timbul karena adanya radiasi panas matahari yang diterima bumi. Persebaran temperatur udara dapat dibedakan menjadi dua, yaitu persebaran horizontal dan vertikal.

1. *Persebaran temperatur udara horizontal.* Temperatur udara di permukaan bumi untuk berbagai tempat tidak sama. Untuk mempermudah membandingkannya, maka dibuat peta isotherm. Isotherm yaitu garis khayal dalam peta yang menghubungkan tempat-tempat yang mempunyai temperatur udara rata-rata sama. Persebaran horizontal secara tidak teratur dipengaruhi oleh kondisi lingkungannya, misalnya perbedaan temperatur udara daratan dan lautan.
2. *Persebaran temperatur udara vertikal*. Semakin tinggi suatu tempat dari atas permukaan laut maka temperatur udara akan semakin turun. Secara umum, setiap naik 100 meter, temperatur udara turun 0,5°C. Ketentuan ini tergantung pada letak dan ketinggian suatu tempat. Adanya perairan, seperti selat dan laut sangat besar peranannya pada pengendalian temperatur, sehingga tidak terjadi perbedaan suhu terendah dan temperatur tertinggi yang sangat besar. (Soko, 2009).

2.7.3. Kelembapan Udara

Kelembaban udara adalah banyaknya uap air yang terkandung dalam massa udara pada saat dan tempat tertentu. Alat untuk mengukur kelembaban udara disebut *psychrometer* atau *hygrometer*. Kelembaban udara dapat dinyatakan dalam beberapa cara yaitu:

1. Kelembaban *absolute* ialah bilangan yang menunjukkan berat uap air tiap kesatuan volume udara. Biasanya dinyatakan dengan gram uap air/m^3 udara.

2. Kelembaban spesifik ialah berat uap air tiap kesatuan berat udara. Biasanya dinyatakan dengan gram uap air/kg udara.
3. Tekanan uap ialah besarnya tekanan yang diberikan oleh uap air sebagai bagian dari udara. Separti tekanan udara, tekanan uap air dinyatakan dalam mb atau mm air raksa.
4. Kelembaban nisbi atau kelembaban relatif ialah perbandingan antara banyaknya air yang terdapat di udara dengan banyaknya uap air maksimum yang dapt dikandung oleh udara pada suhu dan tekanan yang sama. Kelembaban relatif dinyatakan dengan %. Besarnya kelembaban udara relatif ditentukan oleh banyaknya uap air di udara dan suhu udara (Waryono, 1987: 58).

BAB III

PEMBAHASAN

3.1 Analisis Model Regresi Probit

Misalkan terdapat model regresi linier sebagai berikut:

$$y_i = \beta_0 + \beta_1 x_{1i} + \ldots + \beta_k x_{ki} + \varepsilon_i = \beta_j x_{ji} + \varepsilon_i. \quad (3.1)$$

untuk i = 1, 2, …, n dan j = 0, 1, 2, …, k. Atau dapat ditulis :

$$y_i = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_k \end{bmatrix}_i^T \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \varepsilon_i \quad (3.2)$$

Persamaan 3.2 dapat ditulis sebagai matrik $y_i = \mathbf{x}_i^T \boldsymbol{\beta}$.

Pada model regresi, Gauss telah membuat asumsi mengenai variabel ε bahwa nilai rata-rata atau harapan variabel ε adalah sama dengan nol atau

$$\mathrm{E}(\varepsilon) = 0$$

Yang berarti nilai bersyarat ε yang diharapkan adalah sama dengan nol. Sehingga nilai harapan Persamaan (3.1) dapat dituliskan menjadi :

$$\begin{aligned} E(y_i \mid \mathbf{x}_i^T) &= 1.P(y_i = 1 \mid \mathbf{x}_i^T) + 0.P(y_i = 0 \mid \mathbf{x}_i^T) \\ &= P(y_i = 1 \mid \mathbf{x}_i^T) + 0 \\ &= P(y_i = 1 \mid \mathbf{x}_i^T) \\ &= p_i. \end{aligned} \quad (3.3)$$

dimana $\mathbf{x}_i^T = [1 \quad x_{1i} \quad \ldots \quad x_{ki}]$.

Bila p_i adalah probabilitas bahwa $y_i = 1$ dan $q_i = 1 - p_i$ adalah probailitas bahwa $y_i = 0$, maka variabel y_i memiliki probabilitas $p_i + 1 - p_i = 1$. Jika probabilitas p_i harus berada antara angka 0 dan 1 dan y_i harus bernilai 0 atau 1, maka y_i mengikuti distribusi probabilitas Bernoulli dengan syarat

$$0 \le E(y_i \mid \mathbf{x}_i^T) \le 1. \tag{3.4}$$

Untuk memudahkan dalam estimasi model Probit, diasumsikan tiap data memiliki *utility indeks* $\mathrm{I}_i; i = 1, 2, ..., n$ yang tidak teramati, atau biasa disebut dengan *variable latent*, yaitu yang dipengaruhi oleh variabel bebas X_i, sedemikian sehingga semakin besar I_i, maka semakin besar probabilitas terjadinya suatu peristiwa. I_i dapat dituliskan :

$$\mathrm{I}_i = \mathbf{x}_i^T \boldsymbol{\beta} \quad (3.5)$$

Untuk setiap I_i memiliki nilai kritis (*threshold*) sebut I_i^*; $i = 1, 2, ..., n$ yang juga tidak teramati. Sehingga bila $\mathrm{I}_i \ge \mathrm{I}_i^*$ mengindikasikan terjadinya suatu peristiwa. Dan jika $\mathrm{I}_i < \mathrm{I}_i^*$ maka tidak terjadinya suatu peristiwa. Sehingga distribusi peluang y_i bisa dikatakan:

$$f\left(y_i \mid \mathbf{x}_i^T\right) = \begin{cases} 1 \; jika \, \mathrm{I}_i^* \le \mathrm{I}_i \\ 0 \; jika \, \mathrm{I}_i^* > \mathrm{I}_i \end{cases}$$

Karena I_i dan I_i^* tidak teramati, maka diasumsikan bahwa I_i^* berdistribusi normal dengan $\mathrm{I}_i^* \sim \mathrm{N}\left(\mu, \sigma^2\right)$. Sehingga mengakibatkan I_i juga berdistribusi normal dengan rata-rata μ dan variansi σ^2. Untuk memudahkan analisis, diasumsikan bahwa $\mu = 0$ dan $\sigma^2 = 1$, sehingga $\mathrm{I}_i^* \sim \mathrm{N}(0,1)$.

Dengan demikian, probabilitas untuk $y_i = 1$ adalah sebagai berikut:

$$\begin{aligned} \mathrm{P}\left(y_i = 1 \mid \mathbf{x}_i^T\right) &= \mathrm{P}\left(\mathrm{I}_i^* \le \mathrm{I}_i \mid \mathbf{x}_i^T\right) = \mathrm{P}\left(\mathrm{I}_i^* \le \boldsymbol{\beta}\mathbf{x}_i^T \mid \mathbf{x}_i^T\right) \\ &= \Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right) \\ &= p_i. \end{aligned} \tag{3.6}$$

Misalkan Z_i adalah *standart normal variable*, yaitu $\mathrm{Z} \sim \mathrm{N}(0,1)$, dan simbol $\Phi(.)$ untuk fungsi standar normal kumulatif atau *Cumulative Distribution Function* (CDF), maka berlaku :

$$\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}\, dz \tag{3.7}$$

Sedangkan probabilitas untuk $y_i = 0$ adalah

$$\begin{aligned} P\left(y_i = 0 \mid \mathbf{x}_i^T\right) &= 1 - \Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right) = 1 - p_i \\ &= 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}\, dz \end{aligned} \tag{3.8}$$

3.2 Estimasi Parameter Model Probit

Karena y berdistribusi Binomial maka fungsi probabilitas bersyarat untuk y_i adalah sebagai berikut:

$$
\begin{aligned}
P\left(y_i \mid \mathbf{x}_i^T\right) &= P\left(y_i = 1 \mid \mathbf{x}_i^T\right)^{y_i} P\left(y_i = 0 \mid \mathbf{x}_i^T\right)^{1-y_i} \\
&= p_i^{y_i}\left(1-p_i\right)^{1-y_i} \\
&= \left(\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right)^{y_i}\left(1-\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right)^{1-y_i} \\
&= \left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}\,dz\right)^{y_i}\left(1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}\,dz\right)^{1-y_i}
\end{aligned}
\tag{3.9}
$$

Sehingga fungsi probabilitas bersyarat gabungan untuk y_i yaitu sebagai berikut:

$$
\begin{aligned}
P\left(y_i \mid \mathbf{x}_i^T\right) &= P\left(y_1, y_2, \ldots, y_n \mid \mathbf{x}_i^T\right) \\
&= P\left(y_1 \mid \mathbf{x}_1^T\right)P\left(y_2 \mid \mathbf{x}_2^T\right)\ldots P\left(y_n \mid \mathbf{x}_n^T\right) \\
&= \prod_{i=1}^{n} P\left(y_i \mid \mathbf{x}_i^T\right) \\
&= \prod_{i=1}^{n} p_i^{y_i}\left(1-p_i\right)^{1-y_i}
\end{aligned}
\tag{3.10}
$$

Karena kejadian $y_i = 1$ dan $y_i = 0$ saling bebas, urutan masing-masing kejadian tidak saling mempengaruhi maka untuk peluang $y_i = 1$ nilai i berjalan dari 1 sampai m dan peluang $y_i = 0$ dapat dimisalkan i berjalan dari $m+1$ sampai ***n***. Sehingga persamaan 3.10 dapat dituliskan menjadi :

$$
P\left(y_i \mid \mathbf{x}_i^T\right) = p_i^{\sum_{i=1}^{m} y_i}\left(1-p_i\right)^{\sum_{i=m+1}^{n} 1-y_i} \tag{3.11}
$$

Fungsi ini dinamakan sebagai fungsi *likelihood*, dan fungsi *log-likelihood* untuk y_i yaitu sebagai berikut:

$$
\begin{aligned}
L(y_i) &= \ln P\left(y_i \mid \mathbf{x}_i^T\right) \\
&= \ln\left\{ p_i^{\sum_{i=1}^{m} y_i} (1-p_i)^{\sum_{i=m+1}^{n} 1-y_i} \right\} \\
&= \ln\left\{ p_i^{\sum_{i=1}^{m} y_i} (1-p_i)^{n-\sum_{i=m+1}^{n} y_i} \right\} \\
&= \ln\left\{ p_i^{n\bar{y}} (1-p_i)^{n-n\bar{y}} \right\} \qquad (3.12) \\
&= n\bar{y}\ln(p_i) + (n-n\bar{y})\ln(1-p_i) \\
&= n\bar{y}\ln(p_i) + n(1-\bar{y})\ln(1-p_i) \\
&= n\bar{y}\ln\left[\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right] + n\left(1-\bar{y}\right)\ln\left[1-\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right]
\end{aligned}
$$

Untuk memaksimumkan fungsi *log-likelihood* maka perlu mendapatkan turunan pertamanya terhadap parameter dan menyamakannya dengan nol, yaitu:

$$
\begin{aligned}
\frac{dL(Y)}{d\boldsymbol{\beta}} &= \frac{d}{d\beta}\left\{ \begin{array}{l} n\bar{y}\ln\left[\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right] \\ +n\left(1-\bar{y}\right)\ln\left[1-\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right] \end{array} \right\} \\
&= n\bar{y}\frac{1}{\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)}\phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\mathbf{x}_i \qquad (3.13) \\
&\quad -n\left(1-\bar{y}\right)\frac{1}{1-\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)}\phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\mathbf{x}_i
\end{aligned}
$$

$$0 = n\bar{y}\frac{1}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i$$

$$-n\left(1-\bar{y}\right)\frac{1}{1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i$$

$$n\bar{y}\frac{1}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i = n\left(1-\bar{y}\right)\frac{1}{1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i$$

$$\bar{y}\frac{1}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\mathbf{x}_i = \left(1-\bar{y}\right)\frac{1}{1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\mathbf{x}_i$$

$$\frac{1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}\mathbf{x}_i = \frac{1-\bar{y}}{\bar{y}}\mathbf{x}_i$$

$$\left[\frac{1}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}-1\right]\mathbf{x}_i = \frac{1-\bar{y}}{\bar{y}}\mathbf{x}_i$$

$$\frac{\mathbf{x}_i}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}-\mathbf{x}_i = \frac{1-\bar{y}}{\bar{y}}\mathbf{x}_i$$

$$\frac{\mathbf{x}_i}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)} = \mathbf{x}_i + \frac{1-\bar{y}}{\bar{y}}\mathbf{x}_i$$

$$= \left(1+\frac{1-\bar{y}}{\bar{y}}\right)\mathbf{x}_i$$

$$= \frac{1}{\bar{y}}\mathbf{x}_i$$

$$\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i = \bar{y}\mathbf{x}_i$$

$$\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right) = \bar{y} \tag{3.14}$$

Perhatikan Persamaan (3.14) bahwa fungsi rasio probabilitas merupakan nilai rata-rata dari *y*. Dan untuk mendapatkan nilai estimasi parameter diperoleh dengan cara sebagai berikut:

$$\begin{aligned} \bar{y} &= \Phi\left(\mathbf{x}_i^T \hat{\boldsymbol{\beta}}\right) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\mathbf{x}_i^T \hat{\boldsymbol{\beta}}} e^{\frac{-Z^2}{2}} dz \\ &= -\frac{1}{2} erf\left(\frac{1}{2}\sqrt{2}\mathbf{x}_i^T \hat{\boldsymbol{\beta}}\right) + \frac{1}{2} \end{aligned} \tag{3.15}$$

Dimana

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \,,\; \mathbf{x}_i^T = \begin{bmatrix} 1 & x_{1i} & \ldots & x_{ki} \end{bmatrix}, \text{ dan } \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

Persamaan (3.15) adalah nonlinier sehingga solusi untuk parameter β dapat dilakukan secara iterasi numerik dengan menggunakan metode iterasi *Newton-Raphson Nonlinier Maksimum Likelihood*, dengan bentuk iterasi sebagai berikut :

$$\hat{\boldsymbol{\beta}}^{n+1} = \hat{\boldsymbol{\beta}}^{n} - t_n P_n \gamma_n \tag{3.16}$$

dimana

$$t_n = 1,\; P_n = \left(\frac{\partial^2 L}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} |_{\beta^n}\right)^{-1},\; \gamma_n = \frac{\partial L}{\partial \boldsymbol{\beta}} |_{\beta^n} \tag{3.17}$$

Untuk turunan pertama dari fungsi *log-likelihood* telah diperoleh (lihat persamaan 3.13), selanjutnya mencari turunan kedua dari fungsi *log-likelihood* sebagai berikut :

$$\frac{\partial^2 L}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T} = n\bar{y}\,\frac{\left(\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i^T\right)-\left(\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i^T\right)}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)}$$
$$-n\left(1-\bar{y}\right)\frac{\left(\left(1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i^T\right)-\left(\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\mathbf{x}_i^T\right)}{\left(1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)\left(1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)} \tag{3.18}$$

Kemudian dimasukkan kedalam rumus iterasi *Newton-Raphson Nonlinear Maksimum Likelihood* disekitar β^1 diperoleh:

$$\hat{\boldsymbol{\beta}}^2 = \hat{\boldsymbol{\beta}}^1 - t_1 P_1 \gamma_1$$
$$= \hat{\boldsymbol{\beta}}^1 - 1\left(\begin{matrix} n\bar{y}\mathbf{x}_i^T\,\dfrac{\mathbf{x}_i^T\left(\left(\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)-\left(\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)\right)}{\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)} \\ -n\left(1-\bar{y}\right)\mathbf{x}_i^T\,\dfrac{\mathbf{x}_i^T\left(\left(\left(1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)-\left(\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)\right)}{\left(1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)\left(1-\Phi\left(\mathbf{x}_i^T\hat{\boldsymbol{\beta}}\right)\right)}\end{matrix}\right)^{-1}$$
$$\left(\begin{matrix} n\bar{y}\dfrac{1}{\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)}\phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\mathbf{x}_i \\ -n\left(1-\bar{y}\right)\dfrac{1}{1-\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)}\phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\mathbf{x}_i\end{matrix}\right)$$

$$\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_k\end{bmatrix}^2 = \begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_k\end{bmatrix}^1 - 1\left(n\bar{y}\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i^T \frac{\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i^T\left(\left(\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)\left(\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)\left(\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i^T\begin{bmatrix}\beta_0\\ \beta_1\\ \vdots\\ \beta_k\end{bmatrix}\right)\right)-\left(\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)^2}{\left(1-\left(\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)\right)^2}\right)^{-1}$$
$$\left(\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i\left(\left(n\bar{y}\frac{\left(\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)}{\left(\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)}\right)-\left(n\left(1-\bar{y}\right)\frac{\left(\dfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)}{\left(1-\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)}\right)\right)\right)$$

(3.19)

Matrik $\begin{bmatrix}1\\x_1\\\vdots\\x_k\end{bmatrix}_i^T \begin{bmatrix}1\\x_1\\\vdots\\x_k\end{bmatrix}_i^T = \left(\begin{bmatrix}1\\x_1\\\vdots\\x_k\end{bmatrix}_i^T\right)^T \begin{bmatrix}1\\x_1\\\vdots\\x_k\end{bmatrix}_i^T = \begin{bmatrix}1\\x_1\\\vdots\\x_k\end{bmatrix}_i \begin{bmatrix}1\\x_1\\\vdots\\x_k\end{bmatrix}_i^T$, sehingga ordo dari P_n pada Persamaan (3.19) adalah kxk kemudian dikalikan dengan γ_n yang mempunyai ordo $kx1$, diperoleh matrik yang berukuran $kx1$. Untuk mempermudah hitungan, nilai $\hat{\boldsymbol{\beta}}^1$ dapat didekati disekitar 0. Setelah diperoleh nilai $\hat{\boldsymbol{\beta}}^2$, dapat digunakan untuk mencari $\hat{\boldsymbol{\beta}}^3$ dan seterusnya sampai nilai $\hat{\boldsymbol{\beta}}$ konvergen.

3.3 Aplikasi Data

Model probit digunakan untuk menganalisis variabel terikat yang bersifat kategorik dan variabel bebas yang bersifat non kategorik. Dalam penelitian ini, model probit digunakan untuk mengetahui kemungkinan terjadinya hujan berdasarkan temperatur dan kelembaban udara. Dimana curah hujan sebagai variabel terikat, temperatur dan kelembaban udara sebagai variabel bebas.

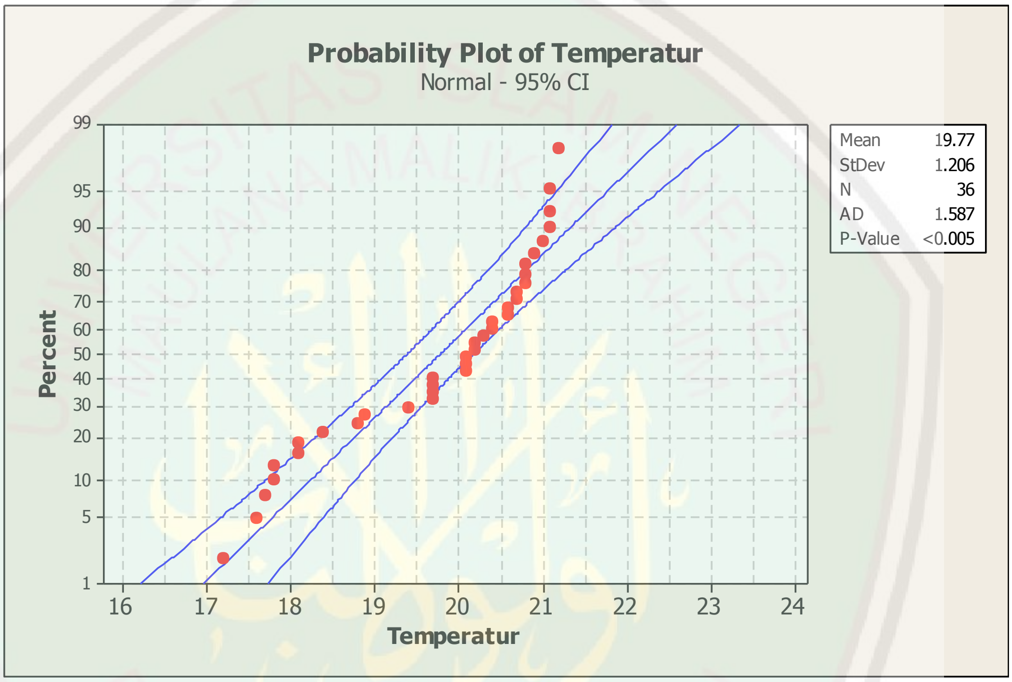
Data yang di pakai adalah data bulanan pada tahun 2007 sampai 2009 di Karangploso, Malang. Dengan ketentuan : pada variabel terikat, disimbolkan dengan angka 1 untuk bulan yang sering terjadi hujan, dengan kriteria curah hujan lebih dari 150 milimeter dan disimbolkan dengan angka 0 untuk bulan yang jarang terjadi hujan, dengan kriteria curah hujan kurang dari 150 milimeter. Data dapat dilihat pada tabel berikut :

Tabel 3.1 Data Curah Hujan di Karangploso Malang.

No	Bulan dan Tahun		Curah Hujan	Temperatur	Kelembaban
1	Januari		0	20.4	48
2	Februari		1	20.9	52
3	Maret		1	21.2	42
4	April		1	20.8	53
5	Mei	2	0	20.1	43
6	Juni	0	0	19.7	60
7	Juli	0	0	18.8	48
8	Agustus	7	0	17.8	44
9	September		0	17.7	27
10	Oktober		0	19.7	40
11	November		1	20.3	47
12	Desember		1	20.6	55
13	Januari		1	20.4	55
14	Februari		1	21.1	58
15	Maret		1	20.2	61
26	April		0	20.1	53
17	Mei	2	0	19.7	41
18	Juni	0	0	18.4	43
19	Juli	0	0	17.2	40
20	Agustus	8	0	18.1	39
21	September		0	18.1	28
22	Oktober		0	21	33
23	November		1	21.1	59
24	Desember		1	20.7	56
25	Januari		1	20.8	54
26	Februari		1	21.1	56
27	Maret		0	20.2	46
28	April		0	20.8	46
29	Mei	2	0	19.7	41
30	Juni	0	0	18.9	41
31	Juli	0	0	17.8	41
32	Agustus	9	0	17.6	38
33	September		0	19.4	33
34	Oktober		0	20.1	29
35	November		1	20.7	35
36	Desember		1	20.6	37

3.3.1 Uji Normalitas Data

Langkah awal dalam aplikasi data adalah menguji kenormalan data. Dalam hal ini digunakan *software* Minitab untuk membuat plot data. Diperoleh hasil sebagai berikut :



Gambar 3.1. Plot Temperatur

Pada Gambar 3.1 dapat diketahui bahwa plot probabilitas data temperatur tampak tersebar tidak mengikuti garis lurus, yang berarti data temperatur dapat dikatakan tidak mengikuti distribusi normal. Untuk mengetahui lebih jelas apakah data berdistribusi normal atau tidak digunakan uji hipotesis dengan nilai toleransi (α) yang ditetapkan adalah 5% (0,05).

<u>Hipotesis</u> :

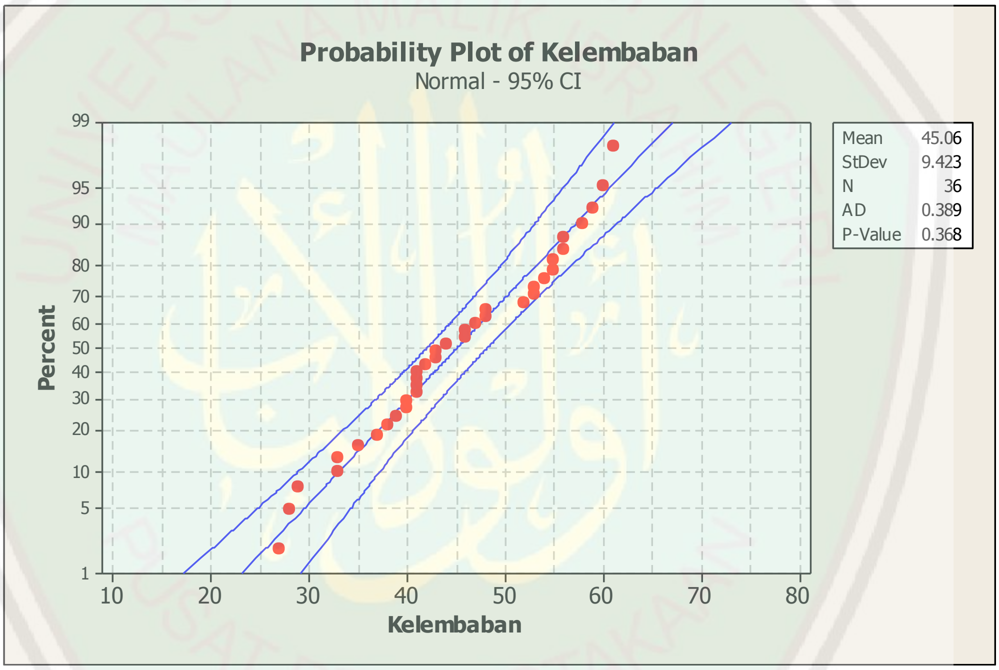
H_0 = data mengikuti distribusi normal

H_1 = data tidak mengikuti distribusi normal

Daerah penolakan :

Jika p - value < α menolak H_0

Pada Gambar 3.1, p - value < α yaitu menolak H_0. Artinya data temperatur tidak mengikuti distribusi normal. Sedangkan untuk data kelembaban dapat dilihat pada gambar berikut :



Gambar 3.2. Plot Kelembapan

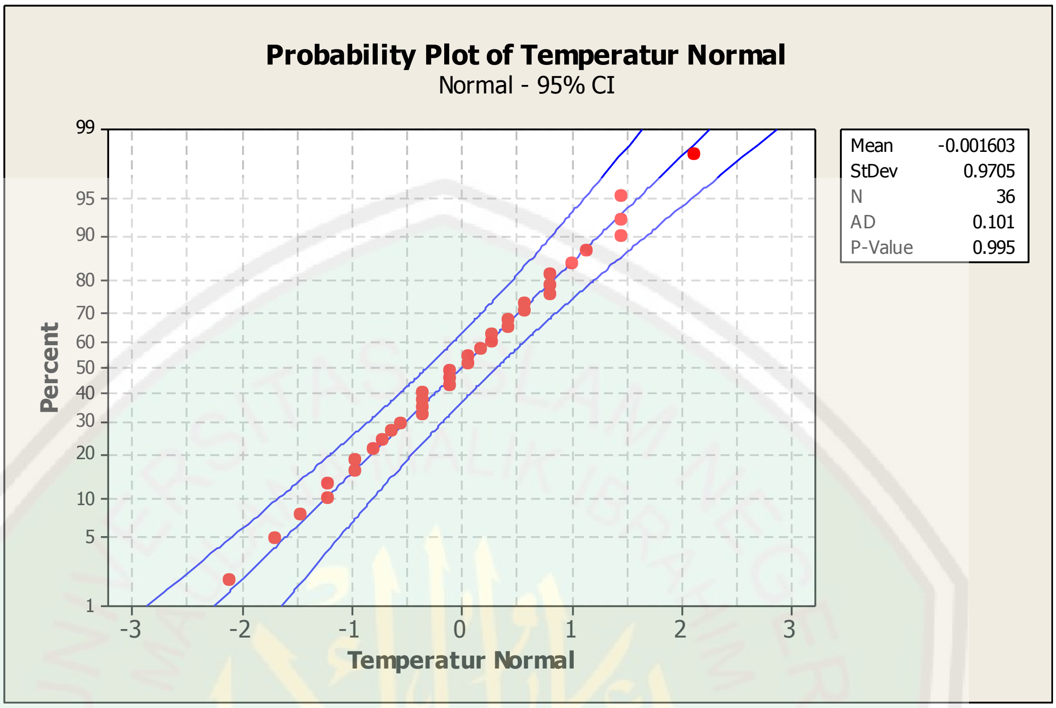
Dari Gambar 3.2 dapat dilihat bahwa plot probabilitas data kelembaban tampak mengikuti garis lurus yang berarti data kelembaban mengikuti distribusi normal. Atau dapat dilihat dari nilai p - value sebesar 0,368 yang berarti > α, menerima H_0.

Untuk data temperatur karena tidak mengikuti distribusi normal, perlu dinormalkan. Untuk menormalkan data temperatur digunakan metode *normal scores* dengan bantuan minitab, diperoleh data berikut :

Tabel 3.2 Data Temperatur yang Dinormalkan.

No	Bulan dan Tahun		Temperatur Setelah Dinormalkan
1	Januari		0.28
2	Februari		1.02
3	Maret		2.11
4	April		0.81
5	Mei	2	-0.1
6	Juni	0	-0.35
7	Juli	0	-0.71
8	Agustus	7	-1.21
9	September		-1.46
10	Oktober		-0.35
11	November		0.17
12	Desember		0.43
13	Januari		0.28
14	Februari		1.46
15	Maret		0.07
26	April		-0.1
17	Mei	2	-0.35
18	Juni	0	-0.81
19	Juli	0	-2.11
20	Agustus	8	-0.96
21	September		-0.96
22	Oktober		1.14
23	November		1.46
24	Desember		0.58
25	Januari		0.81
26	Februari		1.46
27	Maret		0.07
28	April		0.81
29	Mei	2	-0.35
30	Juni	0	-0.63
31	Juli	0	-1.21
32	Agustus	9	-1.7
33	September		-0.54
34	Oktober		-0.1
35	November		0.58
36	Desember		0.43

Dapat dilihat gambar plotnya sebagai berikut :



Gambar 3.3. Plot Temperatur yang Telah Dinormalkan

Pada Gambar 3.3 menunjukkan nilai p - value adalah 0,995, yang berarti lebih besar dari nilai toleransi (α). Sehingga menerima H_0, data temperatur mengikuti distribusi normal.

3.3.2 Regresi Probit dari Data dan Estimasi Parameter

Dalam memprediksi peluang terjadinya hujan yang dipengaruhi oleh temperatur minimum dan kelembaban minimum, dapat diregresikan sebagai berikut :

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \tag{3.19}$$

dimana :

Y_i = curah hujan
X_{1i} = temperatur
X_{2i} = kelembapan

$\beta_0 \beta_1 \beta_2$ = parameter

Karena data Y_i adalah kategorik maka persamaan 3.19 dapat dituliskan kembali menjadi :

$$DY_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \tag{3.21}$$

dengan D adalah *dummy* dari data Y yang hanya bernilai 0 dan 1.

Kemudian dilakukan pendugaan parameter temperatur dan parameter kelembaban. Dengan bantuan E-views diperoleh *output* sebagai berikut:

Dependent Variable: CURAHHUJAN
Method: ML - Binary Probit
Date: 01/09/11 Time: 00:38
Sample: 2007:01 2009:12
Included observations: 36
Convergence achieved after 6 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-4.506676	2.040618	-2.208486	0.0272
TEMPERATUR	2.000155	0.770506	2.595899	0.0094
KELEMBABAN	0.081135	0.040648	1.996049	0.0459

Mean dependent var	0.388889	S.D. dependent var	0.494413
S.E. of regression	0.317308	Akaike info criterion	0.698438
Sum squared resid	3.322574	Schwarz criterion	0.830398
Log likelihood	-9.571877	Hannan-Quinn criter.	0.744495
Restr. log likelihood	-24.05695	Avg. log likelihood	-0.265885
LR statistic (2 df)	28.97014	McFadden R-squared	0.602116
Probability(LR stat)	5.12E-07		
Obs with Dep=0	22	Total obs	36
Obs with Dep=1	14		

Gambar 3.4 Hasil Analisis Probit

Interpretasi Output:

Pada hasil analisis Probit (lihat Gambar 3.4) dapat diketahui bahwa nilai $\hat{\beta}_0$ adalah -4.506676, $\hat{\beta}_1$ adalah 2.000155 dan $\hat{\beta}_3$ adalah 0.081135, atau dapat dituliskan kedalam regresi sebagai berikut :

$$\text{CurahHujan} = -4,51 + 2.00(\text{Temperatur}) + 0.08(\text{Kelembapan}) \qquad (3.22)$$

Misalnya akan diestimasi seberapa besarkah kemungkinan musim hujan pada bulan Januari tahun 2010, dengan temperatur 0,28 dan kelembaban 48. Dengan memasukkan ke dalam regresi probit, diperoleh :

$$\begin{aligned}\text{Probit} &= -4,51 + 2.00(\text{Temperatur}) + 0.08(\text{Kelembapan})\\ &= -4,51 + 2.00(0.28) + 0.08(48)\\ &= -4,51 + 0.56 + 3.84\\ &= -0.11\end{aligned}$$

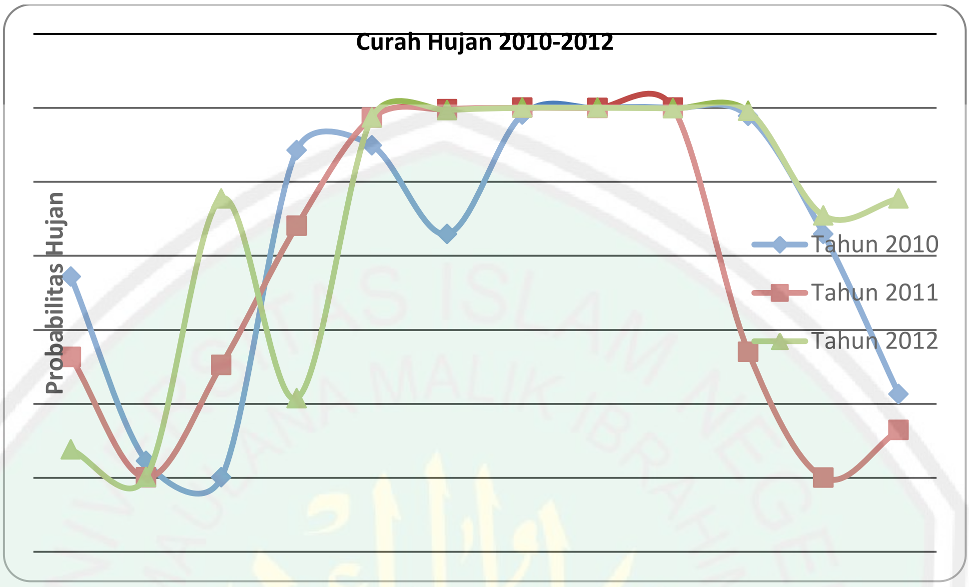
Kemudian di cari pada tabel statistik Z, diperoleh 0,4562. Selanjutnya mengurangkan 1 dengan nilai tersebut, sehingga diperoleh 1 - 0,4562 = 0,5438 atau 54,38%. Dengan demikian, kemungkinan musim hujan pada bulan Januari 2010 adalah 54,38%.

Perhitungan ini dilakukan pada setiap data bulanan pada tahun 2007 sampai 2009. Sehingga dapat memprediksi tiga tahun ke depan, yaitu tahun 2010 sampai 2012. Hasilnya dapat dilihat pada tabel berikut :

Tabel 3.3 Data Probabilitas Curah Hujan.

No	Bulan dan Tahun		Probabilitas Curah Hujan
1	Januari		0.5438
2	Februari		0.0455
3	Maret		0.0011
4	April		0.885
5	Mei	2	0.898
6	Juni	0	0.6591
7	Juli	1	0.9817
8	Agustus	0	0.9997
9	September		1
10	Oktober		0.9778
11	November		0.6591
12	Desember		0.2266
13	Januari		0.3264
14	Februari		0.0011
15	Maret		0.305
26	April		0.6808
17	Mei	2	0.9732
18	Juni	0	0.9964
19	Juli	1	1
20	Agustus	1	0.9995
21	September		1
22	Oktober		0.3409
23	November		0.0009
24	Desember		0.1292
25	Januari		0.0764
26	Februari		0.0019
27	Maret		0.7549
28	April		0.2148
29	Mei	2	0.9732
30	Juni	0	0.9936
31	Juli	1	1
32	Agustus	2	1
33	September		0.9984
34	Oktober		0.9916
35	November		0.7088
36	Desember		0.7549

Dapat ditampilkan dalam grafik sebagai berikut :



Gambar 3.5. Grafik Probabilitas Curah Hujan untuk tahun 2010-2012

Nilai probabilitas yang lebih dari 0.5 menunjukkan pada bulan tersebut termasuk dalam musim penghujan. Pada grafik curah hujan (Gambar 3.5) menunjukkan pada tahun 2010, bulan yang termasuk musim penghujan dimulai bulan April sampai dengan bulan November. Hal ini menunjukkan pada tahun 2010 intensitas musim hujan sangat banyak, karena musim hujan terjadi selama 8 bulan.

Pada tahun 2011, bulan yang termasuk musim hujan dimulai pada bulan April sampai dengan bulan September. Pada tahun ini, musim hujan terjadi selama 6 bulan. Sedangkan pada tahun 2012, bulan yang termasuk musim hujan terjadi pada bulan Maret, Mei sampai dengan bulan Desember. Intensitas musim hujan lebih banyak dibandingkan dengan tahun-tahun sebelumnya, musim hujan terjadi selama 9 bulan.

Perkiraan musim hujan di Karangploso Malang pada tahun 2010-2012 dapat dirinci sebagai berikut :

Tabel 3.4 Data Curah Hujan.

No	Bulan dan Tahun		Musim Hujan
1	Januari		**Hujan**
2	Februari		Tidak Hujan
3	Maret		Tidak Hujan
4	April		**Hujan**
5	Mei	2	**Hujan**
6	Juni	0	**Hujan**
7	Juli	1	**Hujan**
8	Agustus	0	**Hujan**
9	September		**Hujan**
10	Oktober		**Hujan**
11	November		**Hujan**
12	Desember		Tidak Hujan
13	Januari		Tidak Hujan
14	Februari		Tidak Hujan
15	Maret		Tidak Hujan
26	April		**Hujan**
17	Mei	2	**Hujan**
18	Juni	0	**Hujan**
19	Juli	1	**Hujan**
20	Agustus	1	**Hujan**
21	September		**Hujan**
22	Oktober		Tidak Hujan
23	November		Tidak Hujan
24	Desember		Tidak Hujan
25	Januari		Tidak Hujan
26	Februari		Tidak Hujan
27	Maret		**Hujan**
28	April		Tidak Hujan
29	Mei	2	**Hujan**
30	Juni	0	**Hujan**
31	Juli	1	**Hujan**
32	Agustus	2	**Hujan**
33	September		**Hujan**
34	Oktober		**Hujan**
35	November		**Hujan**
36	Desember		**Hujan**

3.3.3 Statistik Uji dari Parameter β

Uji parameter yang digunakan adalah uji tiap-tiap parameter (uji Wald). Dengan nilai α yang ditetapkan adalah 5% (0,05).

Hipotesis :

$H_0 : \beta_j = 0$ untuk suatu j = 0,1,2.

$H_1 : \beta_j \neq 0$

Statistik uji yang digunakan adalah

$$W_j = \left[\frac{\tilde{\beta}_1}{SE(\beta_1)} \right] \text{; j = 0,1,2.}$$

Daerah penolakan :

jika $W_j > \chi^2_{\alpha,1}$; menolak H_0.

Untuk $\chi^2_{0.05,1}$ dapat dilihat pada tabel sebaran khi-kuadrat di peroleh 3,841. Bila H_0 ditolak, artinya parameter tersebut signifikan secara statistik pada tingkat signifikansi α.

Untuk β_0

$$\begin{aligned} W_0 &= \left[\frac{\hat{\beta}_0}{SE\left(\hat{\beta}_0\right)} \right]^2 \\ &= \left[\frac{-4.506676}{2.040618} \right]^2 \\ &= (-2.208486)^2 \\ &= 4.87741 \end{aligned}$$

Diperoleh $W_0 > \chi^2_{0.05,1}$ yaitu $4.87741 > 3,841$, sehingga menolak H_0. Artinya β_0 signifikan secara statistik pada tingkat signifikansi 0,05%.

Untuk β_1

$$W_1 = \left[\frac{\hat{\beta}_1}{SE\left(\hat{\beta}_1\right)}\right]^2$$

$$= \left[\frac{2.000155}{0.770506}\right]^2$$

$$= (2.595898)^2$$

$$= 6.738687$$

Diperoleh $W_1 > \chi^2_{0.05,1}$ yaitu $6.738687 > 3,841$, sehingga menolak H_0. Artinya β_1 signifikan secara statistik pada tingkat signifikansi 0,05%.

Untuk β_2

$$W_2 = \left[\frac{\hat{\beta}_2}{SE\left(\hat{\beta}_2\right)}\right]^2$$

$$= \left[\frac{0.081135}{0.040648}\right]^2$$

$$= (1.996039)^2$$

$$= 3.984172$$

Diperoleh $W_1 < \chi^2_{0.05,1}$ yaitu $3.984172 > 3,841$, sehingga menerima H_0. Artinya β_2 signifikan secara statistik pada tingkat signifikansi 0,05%.

3.3.4 Uji Kebaikan Model

Untuk uji model dilihat nilai dari AIC dan SC, semakin kecil AIC dan SC maka model akan semakin bagus. Pada hasil analisis model Probit dengan

E-views (lihat Gambar 3.4), nilai *Akaike Information Criterion* (AIC) adalah 0,698438 dan nilai *Schawarz Criterion* (SC) adalah 0,830398. Nilai AIC dan SC dengan model Probit ternyata cukup besar, hampir mendekati satu. Akan tetapi bisa jadi nilai AIC dan SC ini cukup kecil jika dibandingkan dengan model lain. Jika memang demikian, maka model Probit cukup bagus untuk diterapkan pada kasus curah hujan.

BAB IV

PENUTUP

4.1 Kesimpulan

Berdasarkan hasil-hasil analisis dan pembahasan pada Bab III, dapat diambil kesimpulan sebagai berikut:

1. Fungsi yang dipakai dalam model Probit adalah *Cumulative Distribution Function* (CDF) normal baku. Karena nilai variabel bebas hanya 1 atau 0, maka distribusi probabilitas model Probit mengikuti distribusi binomial. Sehingga diperoleh model berikut :

$$\left(\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right)^{y_i}\left(1-\Phi\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right)^{1-y_i}$$

2. Estimasi model Probit tidak cukup hanya menggunakan *maxsimum likelihood*, karena persamaannya nonlinier maka untuk estimasi di bantu dengan iterasi *Newton-Raphson* menghasilkan persamaan berikut :

$$\begin{aligned}\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_k\end{bmatrix}^{n+1} &= \begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_k\end{bmatrix}^{n} - 1\left(n\bar{y}\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i^T \frac{\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i^T\left(\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)\left(\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i^T\begin{bmatrix}\beta_0\\ \beta_1\\ \vdots\\ \beta_k\end{bmatrix}\right)\right)-\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)^2}{\left(1-\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)\right)^2}\right)^{-1} \\ &\quad \left(\begin{bmatrix}1\\ x_1\\ \vdots\\ x_k\end{bmatrix}_i\left(\left(n\bar{y}\frac{\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)}{\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)}\right)-\left(n\left(1-\bar{y}\right)\frac{\left(\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{1}{2}\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)^2}\right)}{\left(1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\mathbf{x}_i^T\boldsymbol{\beta}} e^{\frac{-Z^2}{2}}dz\right)}\right)\right)\right)\end{aligned}$$

3. Hasil perolehan perkiraan hujan di daerah Karangploso Malang pada tahun 2010-2012 adalah : pada tahun 2010 musim hujan terjadi pada

bulan Januari, April-November, pada tahun 2011 musim hujan terjadi pada bulan April-September dan pada tahun 2012 musim hujan terjadi pada bulan Maret, Mei-Desember.

4.2 Saran

Penulis menerapkan model Probit pada curah hujan, dengan data bulanan. Penulis menyarankan agar dilakukan penelitian lebih lanjut dengan menggunakan data harian, agar hasil perkiraan lebih akurat.

DAFTAR PUSTAKA

Al-Bassam, Abdullah bin Abdurrahman. 2006. *Syarah Bulughul Maram, Jilid 4,* Jakarta Selatan : Pustaka Azzam.

Bain, Lee J dan Engelhard. Max. 1991. *Introduction to Probability and Mathematical Statistic*. California: Duxbury Press.

Damodar N, Gujarati. 2003. *Basic Econometric, Fourth Edition*. North Amerika: Mc Graw Hill.

Damodar N, Gujarati. 2007. *Dasar-dasar Ekonometri Edisi Ketiga, Jilid I dan II*. Terjemahan M. Jullius A. Jakarta: Erlangga.

Davidson, Russel dan Mackinnon. James G. 1999. *Econometric Theory and Methods.*

Dudewicz, Edward J dan Mishra. Satya N. 1995. *Statistika Matematika Modern*. Terjemahan Sembiring. RK. Bandung: ITB Bandung.

Grenee, William.H. 2003. *Econometric Analysis, Fifth Edition*. New Jersey: Prentice Hall.

Harini, Sri & Kusumawati. Ririen. 2007. *Metode Statistik*, Jakarta: Prestasi Pustaka.

Judge, G.G. 1985. *The Theory and Practice of Ekonometrics*. John Wiley & Sons, Inc.

Irianto, Agus. 2006. *Statistik Konsep Dasar dan Aplikasinya*. Jakarta: Kencana Prenada Media.

Iriawan, Nur. 2006. *Mengolah Data Statistik dengan Mudah Menggunakan MINITAB 14*. Yogyakarta : ANDI.

Nachrowi, 2008. *Pengggunaan Teknik Ekonometri*. Jakarta: PT Raja Grafindo Persada.

Puji B, Siwi Tri. 2010. Waspadai Curah Hujan Tinggi Sepanjang Oktober-Desember. *http://www.republika.co.id/berita/breaking-news/nasional/10/10/04/138022-waspadai-curah-hujan-tinggi-sepanjang-oktoberdesember*

(diakses pada tanggal 27 Januari 2011)

Schimmel, Annemarie. 2006. *The Mistery of Number*. Terjemahan Prihantoro. Agung. Bandung: Pustaka Hidayah.

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

Shihab, M. Quraish. 2003. *Tafsir Al- Mishbah, Volume 7,10,11 dan 13*. Jakarta : Lentera Hati.

Soko . 2009. Cuaca dan Iklim.

http://thehyposentrum.blogspot.com/2009/12/cuaca-dan-iklim-apakah-yang-dimaksud.html
(diakses pada tanggal 27 Januari 2011)

Supranto, J. 2004. *Ekonometri, Jilid I dan II*. Jakarta: Ghalia Indonesia.

Supangat, Andi. 2008. *Statistika dalam Kajian Deskriptif, Inferensi dan Nonparametrik*. Jakarta: Kencana Prenada Media Group.

Walpole, Ronald E & Myers. Raymond H. 1995. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan, Edisi 4*. Terjemahan Sembiring. RK. Bandung: ITB Bandung.

Waryono, dkk. 1987. *Pengantar Meterologi dan Klimatologi*. Surabaya: PT Bina Ilmu

Winarno, Wing Wahyu. 2007. *Analisis Ekonometrika dan Statistik dengan EVIEWS*. Yogyakarta : UPP STIM YKPN.

Lampiran 1

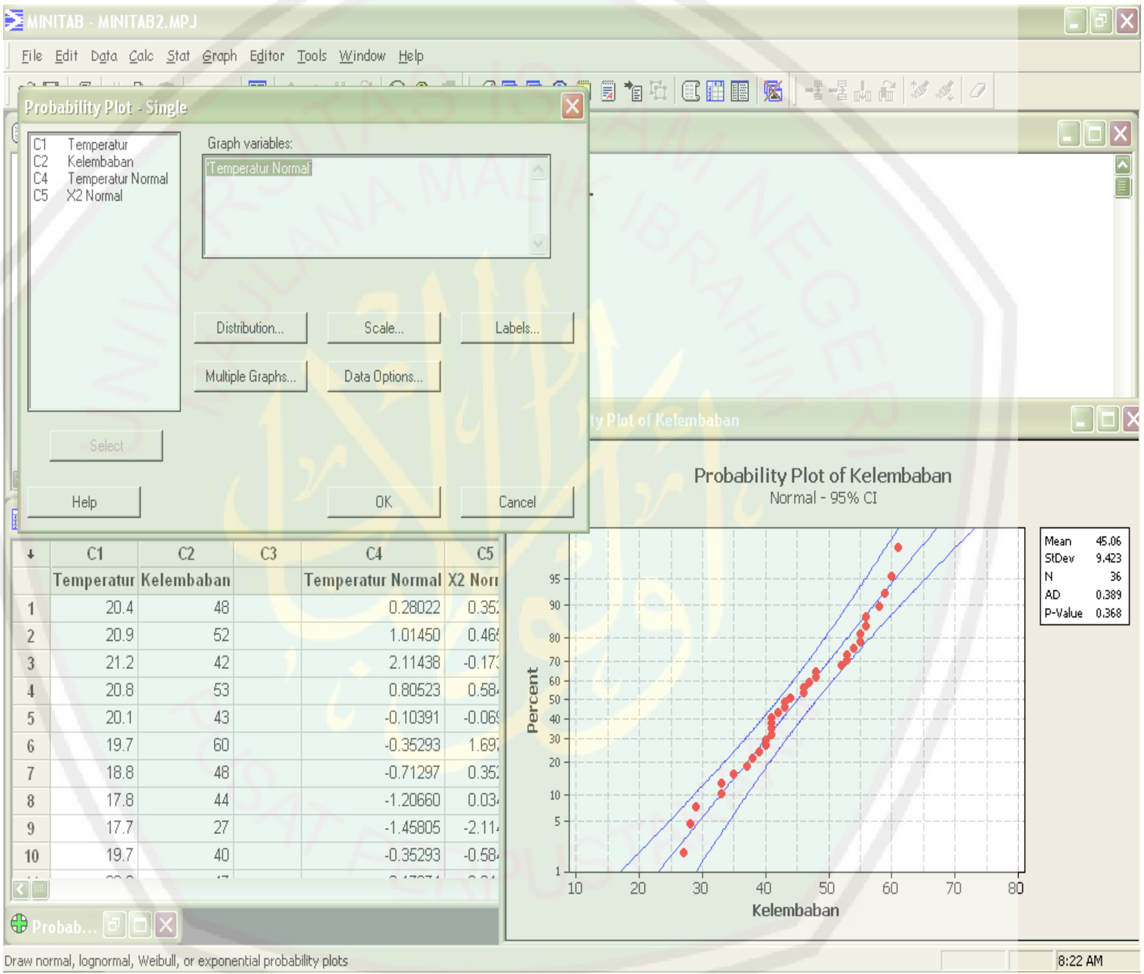
Data Klimatologi Tahun 2007-2009 Karangploso Malang

NO.	BULAN DAN TAHUN		CURAH HUJAN	TEMPERATUR MINIMUM	KELEMBABAN MINIMUM
1	Januari	2007	129	20.4	48
2	Februari		182	20.9	52
3	Maret		173	21.2	42
4	April		235	20.8	53
5	Mei		6	20.1	43
6	Juni		15	19.7	60
7	Juli		7	18.8	48
8	Agustus		1	17.8	44
9	September		10	17.7	27
10	Oktober		61	19.7	40
11	November		272	20.3	47
12	Desember		423	20.6	55
13	Januari	2008	206	20.4	55
14	Februari		315	21.1	58
15	Maret		460	20.2	61
16	April		66	20.1	53
17	Mei		61	19.7	41
18	Juni		2	18.4	43
19	Juli		0	17.2	40
20	Agustus		47	18.1	39
21	September		8	18.1	28
22	Oktober		92	21	33
23	November		174	21.1	59
24	Desember		241	20.7	56
25	Januari	2009	258	20.8	54
26	Februari		435	21.1	56
27	Maret		81	20.2	46
28	April		67	20.8	46
29	Mei		62	19.7	41
30	Juni		70	18.9	41
31	Juli		39	17.8	41
32	Agustus		0	17.6	38
33	September		4	19.4	33
34	Oktober		35	20.1	29
35	November		200	20.7	35
36	Desember		224	20.6	37

Lampiran 2

Cara Membuat Plot pada MINITAB 14 :

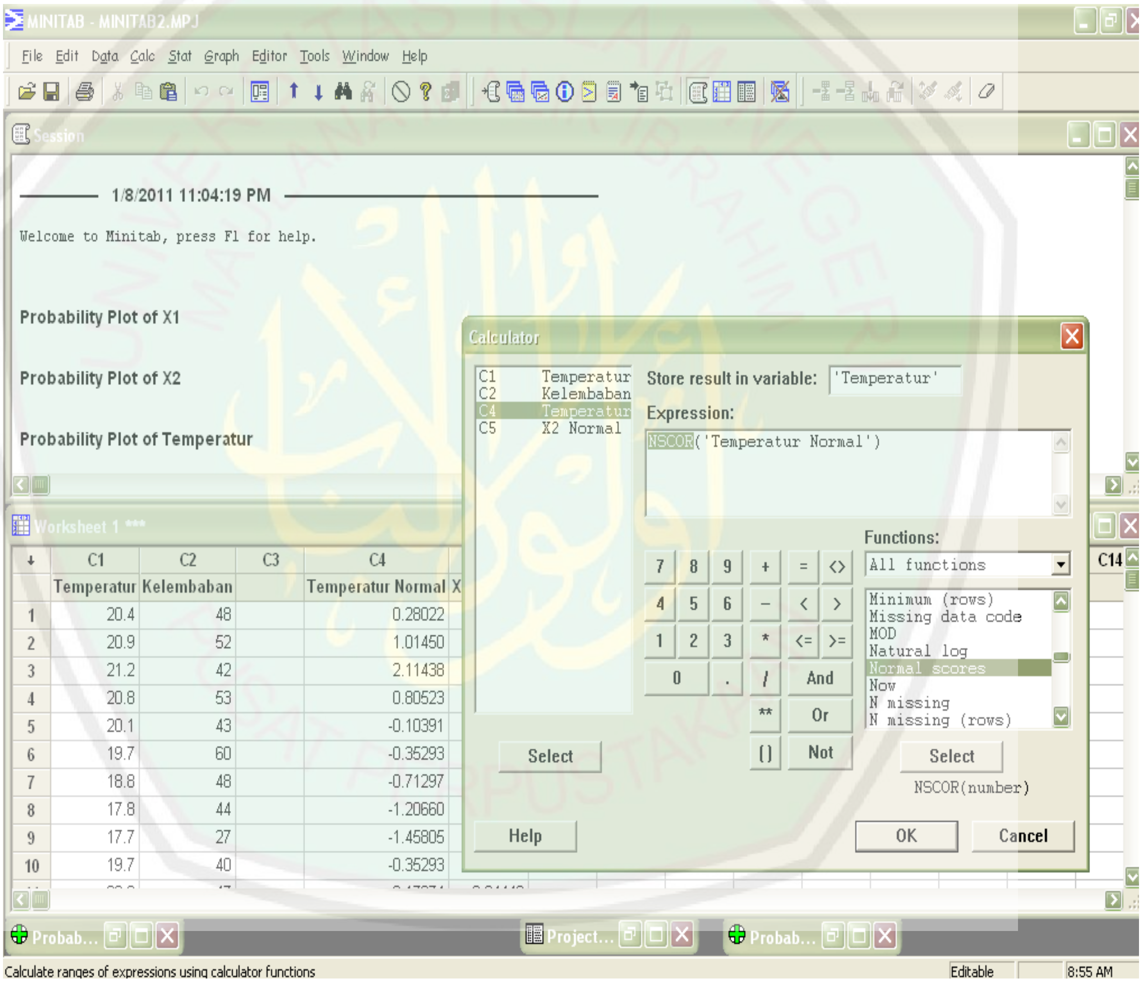
Graph>>Probability Plot>>Single>>Select Data>>OK



Lampiran 3

Cara Menormalkan Data pada MINITAB 14 :

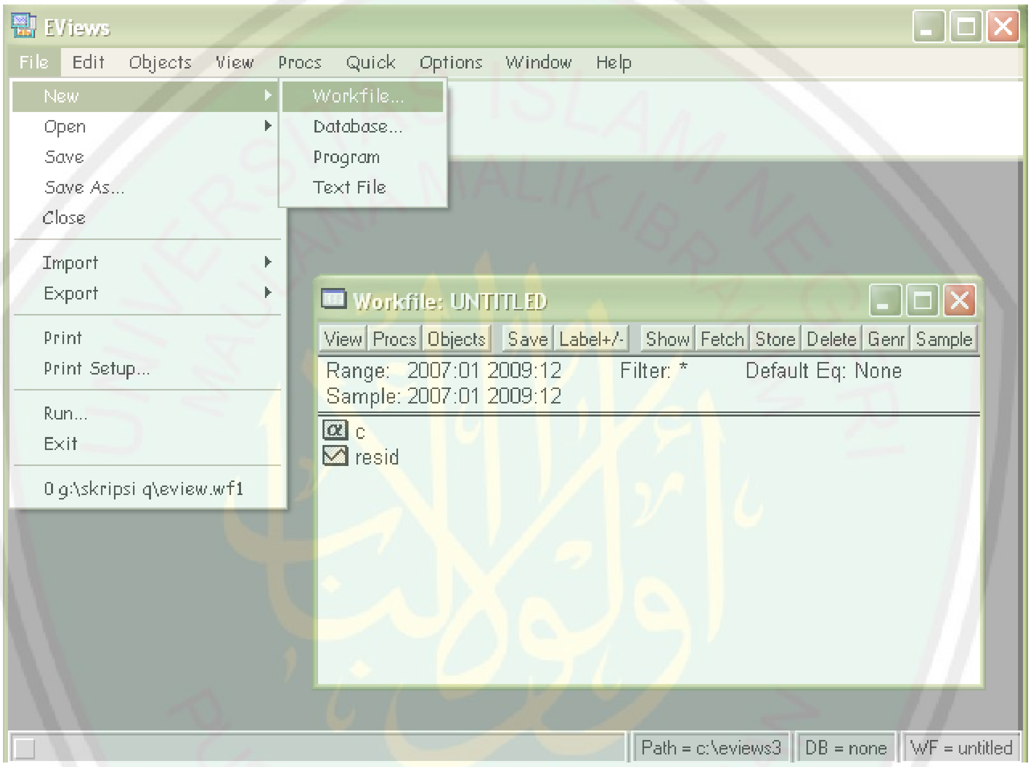
Calc>>Calculator>>Store Result in Variable : pilih variabel yang akan di normalkan>>Function : pilih *Normal Scores*>>Expression : masukkan tenpat untuk variabel yang dinormalkan pada *NSCOR* .



Lampiran 4

Cara Membuka Jendela pada E-VIEWS 3 :

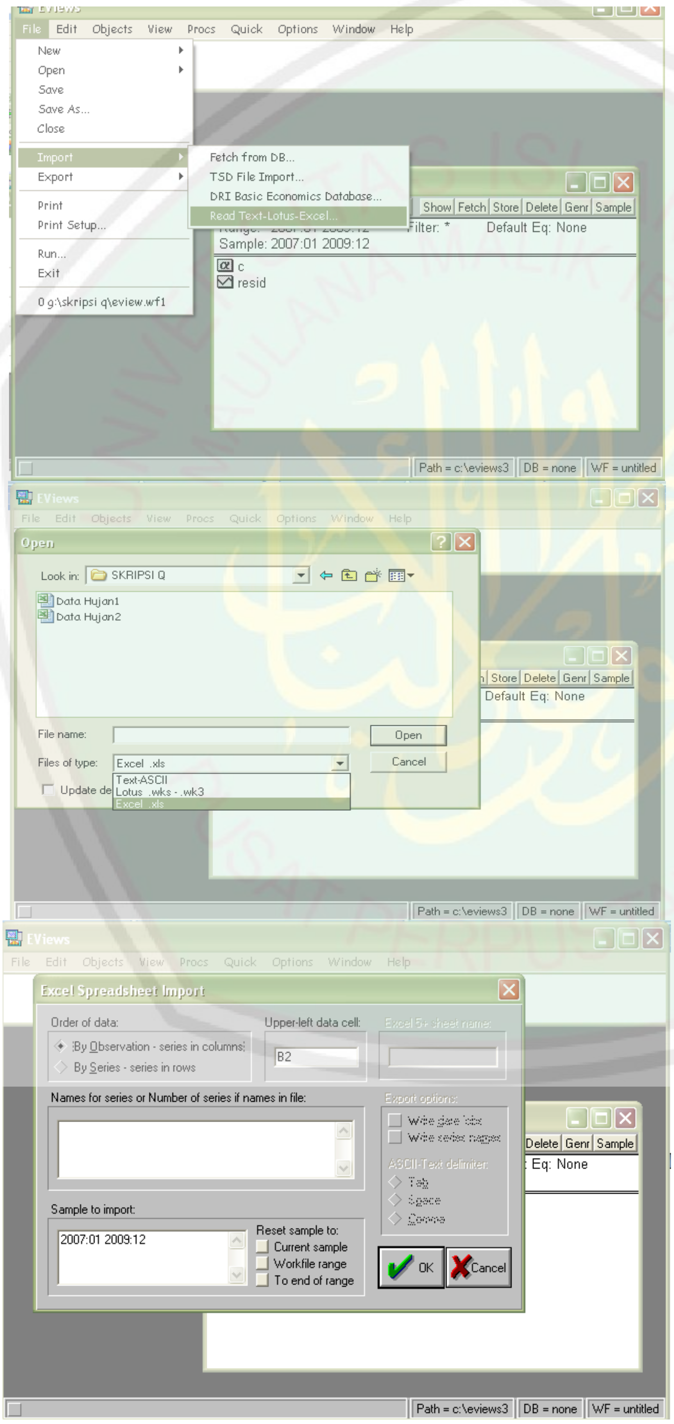
File>>New>>Workfile>> Workfile Range : pilih jenis data *Monthly*>>OK



Lampiran 5

Cara Import Data dari Excel pada E-VIEWS 3 :

File>>Import>>Read Text Lotus Excel>>File Type : Excel.xls>>Open



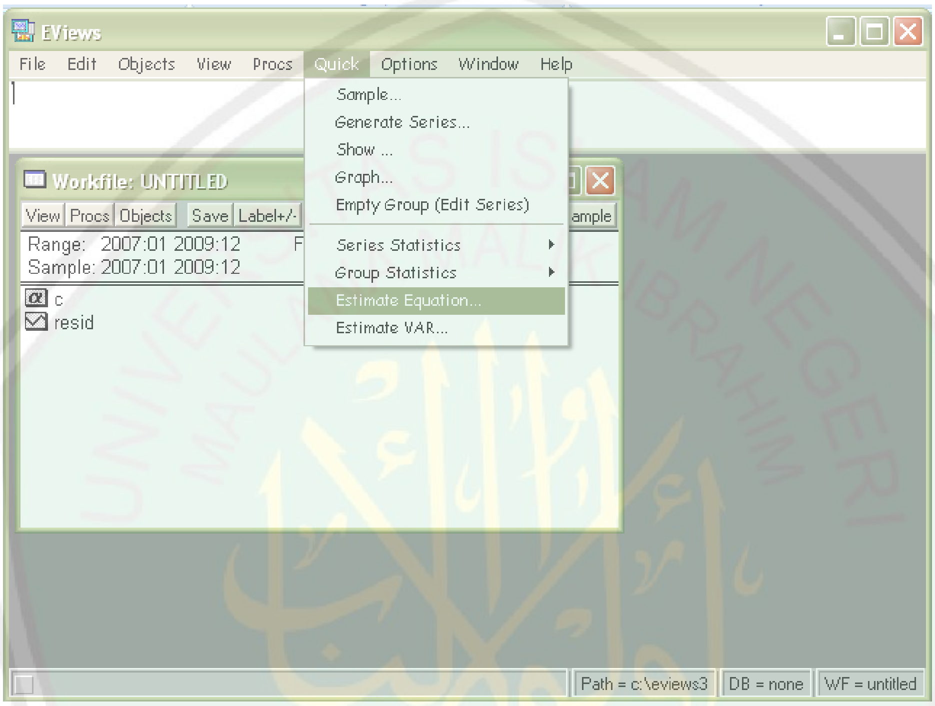
Pada Upper-left data celt : diisi sesuai data di taruh pada kolom berapa di Excel.

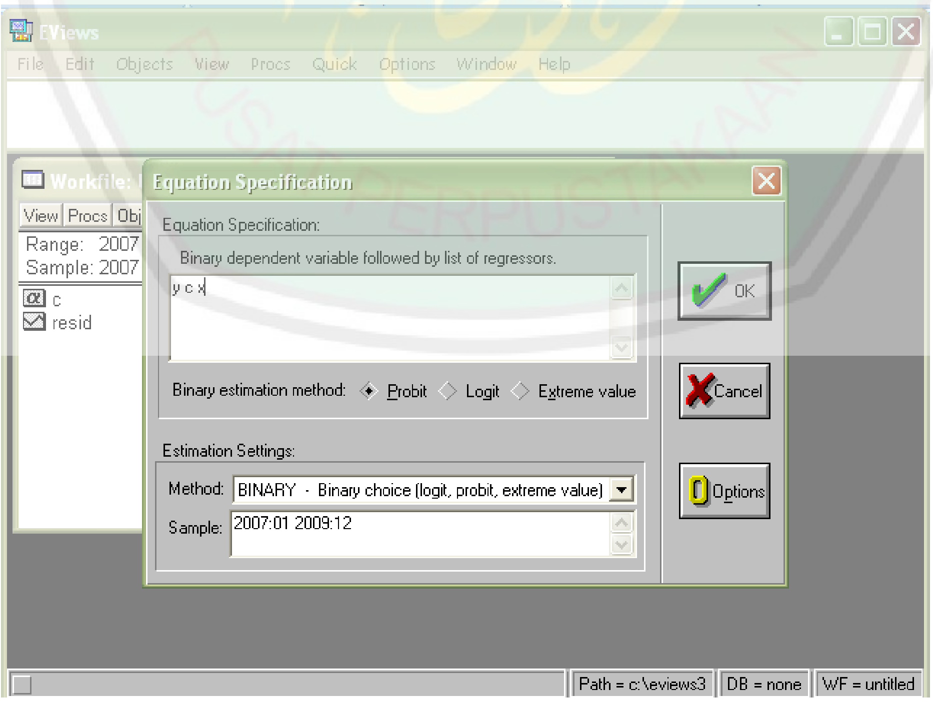
PERINGATAN : Saat membuka Eviews pastikan file Excel dalam keadaan menutup.

Lampiran 6

Cara Estimasi Data pada E-VIEWS 3 :

Quick>>Estimate Equation>>Method: pilih *BINARY PROBIT*>>OK





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 : Fita Zunaidatus Sa'adah
NIM : 06510066
Fakultas/ Jurusan : Sains dan Teknologi/ Matematika
Judul Skripsi : Analisis Regresi *Dummy Varible* Model Probit (Kasus pada Estimasi Hujan di Karangploso Malang)
Pembimbing I : Abdul Aziz, M.Si
Pembimbing II : Dr. H. Ahmad Barizi, M.A

No	Tanggal	HAL	Tanda Tangan	
1	05 Oktober 2010	BAB I & II	1.	
2	06 Oktober 2010	BAB I & II Agama		2.
3	13 Oktober 2010	Revisi BAB I & II	3.	
4	22 Oktober 2010	BAB III		4.
5	01 November 2010	Presentasi I BAB III	5.	
6	11 November 2010	Revisi BAB III		6.
7	15 November 2010	Presentasi II BAB III	7.	
8	14 Desember 2010	ACC BAB I, II Agama		8.
9	18 Desember 2010	BAB IV	9.	
10	09 Januari 2011	Presentasi Keseluruhan		10.
11	12 Januari 2011	ACC Keseluruhan	11.	

Malang,13 Januari 2011

Mengetahui,
Ketua Jurusan Matematika

Abdussakir, M.Pd
NIP. 19751006 200312 1 001

DAFTAR RIWAYAT HIDUP

Nama : Fita Zunaidatus Sa'adah
Tempat/Tgl. Lahir : Tulungagung, 11 September 1988
Agama : Islam
Tempat Tinggal : Dsn. Blimbing RT/RW: 009/002
Ds. Jeli Kec. Karangrejo
Kab. Tulungagung 66253
E-mail : vta_potter@yahoo.com

PENDIDIKAN FORMAL:

1. Madrasah Ibtidaiyah Negeri (MIN) Jeli Karangrejo Tulungagung , Tamat Tahun 2000.
2. Madrasah Tsanawiyah (MTS) TRIBAKTI Lirboyo Kediri, Tamat Tahun 2003.
3. Sekolah Menengah Atas (SMA) Islam SGJ, Ngunut Tulungagung, Tamat Tahun 2006.
4. S1 Fakultas Sains dan Teknologi (Jurusan Matematika), Universitas Islam Negeri Maulana Malik Ibrahim Malang.

PENGALAMAN ORGANISASI

1. Devisi Penerbitan dan Pengembangan Wacana IPPNU PKPT UIN Maulana Malik Ibrahim Malang, tahun 2009.

2. Wakil Ketua I IPPNU PKPT UIN Maulana Malik Ibrahim Malang, tahun 2010.

Demikian daftar riwayat hidup ini saya buat dengan sebenar-benarnya.

Malang, 11 Januari 2011

Fita Zunaidatus Sa'adah