



TESIS - SS 142501

**REGRESI KUANTIL BERBASIS MODEL REKURSIF
DAN ESTIMASI SPARSITY UNTUK ANALISIS
PUBLIKASI DOSEN ITS DI SCOPUS**

**ALFISYHRINA HAPSERY
NRP. 1315201023**

**DOSEN PEMBIMBING
Dr. Suhartono, M.Sc.
Dr.rer.pol. Dedy Dwi Prastyo, M.Si.**

**PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017**



TESIS - SS 142501

**QUANTILE REGRESSION BASED ON RECURSIVE
MODEL AND SPARSITY ESTIMATION TO
ANALYZE PUBLICATION OF ITS LECTURER IN
SCOPUS**

**ALFISYHRINA HAPSERY
NRP. 1315201023**

SUPERVISOR

Dr. Suhartono, M.Sc.

Dr.rer.pol. Dedy Dwi Prastyo, M.Si.

**PROGRAM OF MAGISTER
DEPARTMENT OF STATISTIKA
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017**

**REGRESI KUANTIL BERBASIS MODEL REKURSIF
DAN ESTIMASI SPARSITY UNTUK PUBLIKASI DOSEN ITS DI SCOPUS**

**Tesis ini disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si.)
di
Institut Teknologi Sepuluh Nopember
Oleh**

ALFISYAHIRINA HAPSERY

NRP : 1315 201 023

Tanggal Ujian : 9 Januari 2017

Periode Wisuda : Maret 2017

Disetujui Oleh:



**1. Dr. Suhartono, M.Sc.
NIP. 19710929 199512 1 001**

(Pembimbing I)



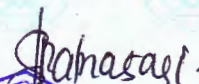
**2. Dr. rer.pol Dedy Dwi P, M.Si.
NIP. 19831204 200812 1 002**

(Pembimbing II)



**3. Dr. Santi Putri R, M.Si.
NIP. 19750115 199903 2 003**

(Penguji)



**4. Dr. Vita Ratnasari, M.Si.
NIP. 19700910 199702 2 001**

(Penguji)



**an. Direktur Program Pascasarjana
Asisten Direktur**

Direktur Program Pascasarjana,

**Prof. Dr. Ir. Tri Widjaja, M.Eng.
NIP. 19611021 198603 1 001**

**Prof. Ir. Djauhar Manfaat, M.Sc., Ph.D.
NIP. 19601202 198701 1 001**

Regresi Kuantil Berbasis Model Rekursif dan Estimasi Sparsity untuk Analisis Publikasi Dosen ITS Di Scopus

Nama Mahasiswa : Alfisyahrina Hapsery
NRP : 1315 201 023
Dosen Pembimbing : Dr. Suhartono, M.Sc.
Dr.rer.pol. Dedy Dwi Prastyo, M.Si.

ABSTRAK

Regresi kuantil adalah metode yang dikembangkan dari metode regresi untuk memodelkan data yang mempunyai sebaran tidak seragam. Dengan prinsip meminimumkan jumlah absolut residual yang dikenal dengan *Least Absolute Deviation* (LAD), regresi kuantil dapat mengetahui tingkat perubahan pada kuantil bersyarat yang dinyatakan dengan nilai koefisiennya. Tujuan dari penelitian ini yaitu mendapatkan estimasi titik regresi kuantil, taksiran interval dengan metode *direct* fungsi sparsity, dan menerapkan pada kasus riil. *Direct* fungsi sparsity adalah taksiran interval yang dapat menjelaskan penyebaran data dengan menggunakan ukuran kuantil sesuai sebaran data. Untuk mendapatkan estimator yang konsisten pada taksiran interval yang sempit perlu dilakukan algoritma yang sesuai yaitu simpleks. Algoritma simpleks merupakan algoritma yang paling stabil diantara algoritma lainnya. Regresi kuantil dalam penelitian ini diterapkan pada kasus publikasi dosen ITS di Scopus dimana variabel dependen yang digunakan adalah sitasi dan indeks h. Hasil studi simulasi menunjukkan bahwa residual yang homogen, pada regresi kuantil memiliki nilai *intercept* yang selalu naik, sedangkan nilai *slope* pada setiap kuantil sejajar. Untuk residual yang heterogen, pada regresi kuantil baik nilai *intercept* dan *slope* akan selalu naik. Hal ini sesuai dengan teori bahwa regresi kuantil dapat memodelkan sesuai dengan sebaran data. Hasil analisis regresi kuantil univariat diketahui bahwa Sitasi selalu berbanding lurus dengan jumlah dokumen dan jumlah *co-authors*. Selain itu, untuk regresi kuantil multivariabel jumlah dokumen berpengaruh terhadap sitasi dan indeks h hampir pada setiap kuantil. Untuk hasil secara keseluruhan diketahui bahwa regresi kuantil dapat memodelkan data publikasi dosen ITS di Scopus sesuai dengan sebaran data.

Kata Kunci : Regresi Kuantil, Sparsity, Publikasi Dosen, Indeks h, Sitasi, Scopus.

Quantile Regression Based On Recursive Model And Sparsity Estimation To Analyze Publication Of Its Lecturer In Scopus

Name : Alfisyahrina Hapsery
NRP : 1315 201 023
Supervisor : Dr. Suhartono, M.Sc.
Dr.rer.pol. Dedy Dwi Prastyo, M.Si.

ABSTRACT

Quantile regression is a method developed from regression to model the data that have ununiform distribution. The principle of this method is least absolute deviation (LAD). This method can determine the rate of change in the conditional on X indicated in the coefficient. The purposes of this study are to get the parameter of the quantile regression, estimate interval by the method of direct function of sparsity, and applying the real case on publication of ITS lecturers in Scopus. To obtain a consistent estimator on estimate interval, we have to use the appropriate algorithm that is simplex. Simplex algorithm is an algorithm which is the most stable among the others. The results of simulation studies indicate that homogen error in the quantile regression have the intercept values that always increase, where as the value of slope at each quantile is stable. For heterogen error, intercept and slope in quantile regression always increase in accordance with the theory. From the result of univariate quantile regression is known that the citation is always proportional to the number of documents and the number of co-authors. Beside, the result of multivariable quantile regression shows that number of documents affect to the citations and h index almost at each quantile. The result of applied studies on publication of ITS lecturers shows that this method can make the model relevant based on pattern of data.

Keywords: Quantile Regression, Sparsity, h index, Citation, Scopus.

KATA PENGANTAR

Segala puji bagi Allah SWT yang telah memberikan nikmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tesis yang berjudul “**REGRESI KUANTIL BERBASIS MODEL REKURSIF DAN ESTIMASI SPARSITY UNTUK ANALISIS PUBLIKASI DOSEN ITS DI SCOPUS**” dengan baik.

Dalam penulisan Tesis ini, tentunya banyak pihak yang telah memberikan bantuan dan dukungan. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Kedua orangtua tercinta ir. Hery Purwono dan Ibu Dra. R.Siti Hapsah, kakak tersayang Alladina Hapsery, ST dan belahan jiwaku Altasharina Hapsery, serta seluruh keluarga yang selalu memberikan doa, dukungan dan semangat. Terimakasih telah menemani disaat suka maupun duka, mengajarkan arti perjuangan, kebaikan dan kasih sayang kepada penulis.
2. Bapak Dr. Suhartono selaku dosen pembimbing, dosen wali dan Ketua Jurusan Statistika FMIPA ITS yang tidak kenal lelah meluruskan apa yang benar dan tidak, dan memberikan waktu disela-sela kesibukan sehingga Tesis ini bisa terselesaikan dengan baik.
3. Bapak Dr.rer.pol Dedy Dwi Prastyo selaku dosen pembimbing yang selalu mendengar, memberi masukan dan sabar menjawab setiap permasalahan dalam proses pengerjaan Tesis ini, sehingga dapat terselesaikan dengan baik.
4. Bapak Dr.rer.pol. Hery Kuswanto selaku Ketua Prodi Pasca Sarjana yang memberikan motivasi selama proses perkuliahan dan pengerjaan Tesis.
5. Ibu Dr. Santi Putri Rahayu dan Ibu Dr. Vita Ratnasari selaku dosen penguji atas kritik dan saran yang membangun.
6. Seluruh dosen dan karyawan jurusan Statistika ITS, atas ilmu yang telah diberikan selama penulis berada di bangku kuliah.
7. Sahabat-sahabat EKSAKTA 2010 terkhusus untuk Azifah Zaini. Sahabat yang selalu menyemangati dan ada disaat susah dan senang.
8. Sahabat-sahabat sehati Dessy, Dian, Farida, Fanny, dan Brina. Terimakasih telah berjuang sampai tahap ini. Serta Tim Kuantil, terutama Moh. Yahya,

teman seperjuangan dalam mengerjakan tesis. Teman-teman seperjuangan S2 Statistika 2015.

9. Semua pihak yang tidak dapat disebutkan satu persatu. Terimakasih telah memberikan pelajaran dan ilmu yang berharga bagi penulis.

Akhir kata, semoga Tesis ini bermanfaat untuk semua pihak, khususnya untuk kampus tercinta ITS.

Surabaya, Januari 2017

Penulis

DAFTAR ISTILAH

- T (e.g., \mathbf{x}^T) : Transpose
- \mathbf{y} : Vektor
- \mathbf{X} : Matrik
- \dagger : Kuantil
- p : Jumlah variabel independen
- k : Jumlah estimasi parameter
- n : Jumlah sampel
- $s(\dagger)$: Parameter regresi kuantil
- $\hat{s}(\dagger)$: Estimasi parameter regresi kuantil
- $F_Y(y)$: Fungsi distribusi kumulatif
- $Q_Y(\dagger)$: Fungsi kuantil
- $Q(\dagger)(\mathbf{y} | \mathbf{X})$: Fungsi kuantil ke $-\dagger$ dari variabel y dengan syarat \mathbf{X}
- $\dots(\dagger)(y)$: *Loss or check function*
- $Q(\dagger)(\mathbf{y} | \mathbf{X}) = \mathbf{X} s(\dagger) + v$: Model regresi kuantil

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISTILAH	xiii
DAFTAR ISI	xv
DAFTAR TABEL	xix
DAFTAR GAMBAR	xxi
DAFTAR LAMPIRAN	xxv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	5
1.4 Manfaat Penelitian	6
1.5 Batasan Penelitian	6
BAB II TINJAUAN PUSTAKA	8
2.1 Regresi Linier Berganda	8
2.1.1 Estimasi Parameter Regresi Linier	9
2.1.2 ANOVA.....	10
2.1.3 Uji Korelasi	10
2.2 Model Rekursif	11
2.3 Regresi Kuantil	12
2.3.1 Estimasi Parameter pada Regresi Kuantil.....	13
2.3.2 Optimasi dengan Algoritma Simpleks pada Model Regresi Kuantil	15
2.3.3 Konstruksi Selang Kepercayaan pada Regresi Kuantil	16
2.3.4 Pengujian Hipotesis	17
2.3.5 Kriteria Keباikan Model Regresi Kuantil	18

2.5 Institut Teknologi Sepuluh Nopember	19
2.6 Publikasi Scopus	19
2.7 Penelitian Tentang Regresi Kuantil Terdahulu	20
2.8 Penelitian Tentang Publikasi Terdahulu	21
BAB III METODE PENELITIAN	23
3.1 Sumber Data	23
3.2 Metode Penelitian	25
3.2.1 Kajian Teoritis	26
3.2.2 Kajian Empiris	27
BAB IV ANALISIS DAN PEMBAHASAN	33
4.1 Kajian Teoritis	33
4.1.1 Regresi Kuantil	33
4.1.2 Taksiran Interval <i>Direct</i> Fungsi Sparsity	40
4.2 Studi Simulasi	44
4.2.1 Studi Simulasi Residual Homogen	44
4.2.2 Studi Simulasi Residual Heterogen	45
4.2.3 Studi Simulasi Berdasarkan Ukuran Sampel	50
4.2.4 Studi Simulasi Berdasarkan Ukuran Varians	53
4.3 Karakteristik Dosen ITS di Scopus	57
4.4 Hubungan antar Variabel pada Analisis Regresi	66
4.5 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Linier Berdasarkan Model Rekursif	74
4.6 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Kuantil Berdasarkan Model Rekursif	80
4.6.1 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Kuantil Berdasarkan Model Rekursif Multivariabel	80
4.6.2 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Kuantil Berdasarkan Model Rekursif Univariat untuk Variabel Dependen Kontinu	87
BAB V KESIMPULAN DAN SARAN	91
5.1 Kesimpulan	91
5.2 Saran	92

DAFTAR PUSTAKA	93
LAMPIRAN	97
LAMPIRAN I Data Scopus	97
LAMPIRAN II Syntax Software R	98
LAMPIRAN III Output Software R	103
LAMPIRAN IV Statistika Deskriptif	106
LAMPIRAN V Regresi Kuantil	119
LAMPIRAN VI Langkah Analisis Menggunakan Stata 12	120

DAFTAR TABEL

		Halaman
Tabel 2.1	ANOVA	1
Tabel 2.2	Sitasi Tiap Dokumen Dosen “S”	19
Tabel 3.1	Stuktur Data	21
Tabel 3.2	Variabel Penelitian	21
Tabel 3.3	Keterangan Koding Variabel <i>Dummy</i>	22
Tabel 3.4	Rincian Jabatan Fungsional Dosen	23
Tabel 4.1	Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Linier Studi Simulasi <i>Error</i> Homogen	44
Tabel 4.2	Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Kuantil Studi Simulasi <i>Error</i> Homogen	44
Tabel 4.3	Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Linier Studi Simulasi <i>Error</i> Heterogen	47
Tabel 4.4	Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Kuantil Studi Simulasi <i>Error</i> Heterogen	48
Tabel 4.5	Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Linier Berdasarkan Ukuran Sampel	50
Tabel 4.6	Rata-rata Estimasi Parameter Hasil Regresi Kuantil pada Studi Simulasi Berdasarkan Perbedaan Ukuran Varians	56
Tabel 4.7	Statistika Deskriptif Sitasi Berdasarkan Jurusan	56
Tabel 4.8	Statistika Deskriptif Indeks h Setiap Jurusan	57
Tabel 4.9	Analisis Faktor	61
Tabel 4.10	Jumlah Masing-masing Anggota Cluster	63
Tabel 4.11	Karakteristik Dosen Masing-masing Cluster	64
Tabel 4.12	Nama Anggota pada Cluster Tiga	65
Tabel 4.13	Matriks Korelasi antara Variabel	68
Tabel 4.14	Hubungan Sitasi dengan Variabel Independen <i>Dummy</i>	72
Tabel 4.15	Uji Parsial pada Variabel Dependen Sitasi	74
Tabel 4.16	Pemilihan Model Terbaik pada Variabel Dependen Sitasi	74
Tabel 4.17	Uji Parsial pada Variabel Dependen Indeks h	75

Tabel 4.18	Pemilihan Model Terbaik pada Variabel Dependen Indeks h ...	75
Tabel 4.19	Estimasi Parameter regresi Kuantil untuk Variabel Dependen Sitasi	79
Tabel 4.20	Estimasi Parameter regresi Kuantil untuk Variabel Dependen Indeks h	81
Tabel 4.21	Hasil <i>Backtesting</i> untuk Model Regresi Kuantil Berbasis Model Rekurdif	82
Tabel 4.22	Hasil Dugaan Indeks h dengan Jumlah Dokumen Berdasarkan Jenis Kelamin	85
Tabel 4.23	Hasil Dugaan Indeks h dengan Jumlah Dokumen Berdasarkan Pendidikan Terakhir	85
Tabel 4.24	Hasil Dugaan Indeks h dengan Jumlah Dokumen Berdasarkan Tempat Pendidikan Terakhir	87

DAFTAR GAMBAR

		Halaman
Gambar 2.1	Model Rekursif	10
Gambar 2.2	<i>Loss Function</i>	13
Gambar 2.3	Perhitungan Indeks h	19
Gambar 2.4	Indeks h Dosen “S”	20
Gambar 3.1	Diagram Alir Kajian Teori	24
Gambar 3.2	Diagram Alir Pemodelan Dosen ITS di Scopus	26
Gambar 4.1	<i>Scatter Plot</i> Replikasi ke-1000 Hasil Studi Simulasi <i>Error</i> Homogen	45
Gambar 4.2	<i>Boxplot</i> \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Homogen	45
Gambar 4.3	<i>Boxplot</i> \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Homogen	46
Gambar 4.4	<i>Scatter Plot</i> Replikasi ke-1000 Hasil Studi Simulasi <i>Error</i> Heterogen	49
Gambar 4.5	<i>Boxplot</i> \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen	49
Gambar 4.6	<i>Boxplot</i> \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen	49
Gambar 4.7	<i>Boxplot</i> \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen dengan Ukuran Sampel 50	51
Gambar 4.8	<i>Boxplot</i> \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen dengan Ukuran Sampel 150	51
Gambar 4.9	<i>Boxplot</i> \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen dengan Ukuran Sampel 50	52
Gambar 4.10	<i>Boxplot</i> \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen dengan Ukuran Sampel 150	52

Gambar 4.11	<i>Boxplot</i> \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen untuk $\dagger_{v_i}^2 = 0,25$	53
Gambar 4.12	<i>Boxplot</i> \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen untuk $\dagger_{v_i}^2 = 4$	53
Gambar 4.13	<i>Boxplot</i> \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen untuk $\dagger_{v_i}^2 = 0,25$	54
Gambar 4.14	<i>Boxplot</i> \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi <i>Error</i> Heterogen untuk $\dagger_{v_i}^2 = 4$	54
Gambar 4.15	Grafik Rata-rata Sitasi	55
Gambar 4.16	Grafik Rata-rata Indeks h	59
Gambar 4.17	<i>Bubble Chart</i> antara Indeks h dan Sitasi	60
Gambar 4.18	<i>Scatter Plot</i> (a) Variabel Usia dengan Sitasi, (b) Variabel Lama Bekerja dengan Sitasi, (c) Variabel Jumlah Dokumen dengan Sitasi, (d) Variabel Jumlah <i>co-authors</i> dengan Sitasi Simbol \square Merupakan Dosen Laki-laki, dan Simbol \square Merupakan Simbol untuk Dosen Perempuan	62
Gambar 4.19	Grafik Hubungan antara Sitasi dan Indeks h. Simbol \square Merupakan Dosen Laki-laki, dan Simbol \square Merupakan Simbol untuk Dosen Perempuan	63
Gambar 4.20	<i>Scatter Plot</i> (a) Variabel Usia dengan Indeks h, (b) Variabel Lama Bekerja dengan Indeks h, (c) Variabel Jumlah Dokumen dengan Indeks h, (d) Variabel Jumlah <i>co-authors</i> dengan Indeks h. Simbol \square Merupakan Dosen Laki-laki, dan Simbol \square Merupakan Simbol untuk Dosen Perempuan.....	69
Gambar 4.21	<i>Boxplot</i> (a) Variabel Jenis Kelamin dengan Sitasi, (b) Variabel Jenis Kelamin dengan Indeks h, (c) Variabel Jabatan Fungsional dengan Sitasi, (d) Variabel Jabatan Fungsional dengan Indeks h	70

Gambar 4.22	<i>Boxplot</i> (a) Variabel Pendidikan Terakhir dengan Sitasi, (b) Variabel Indeks h dengan Pendidikan Terakhir, (c) Variabel TempatPendidikan Terakhir dengan Sitasi, (d) Variabel Indeks h dengan TempatPendidikan Terakhir	71
Gambar 4.23	<i>Boxplot</i> Residual dari Model Rekursif Pertama	78
Gambar 4.24	<i>Boxplot</i> Residual dari Model Rekursif Kedua	78
Gambar 4.25	<i>Scatter Plot</i> Residual dan Prediksi dari Model Rekursif Pertama	79
Gambar 4.26	<i>Scatter Plot</i> Residual dan Prediksi dari Model Rekursif Kedua	79
Gambar 4.27	<i>Scatter Plot</i> Regresi Kuantil untuk Indeks h dengan Jumlah Dokumen <i>Conditional on X</i> dimana dibedakan berdasarkan Jenis Kelamin (a) Perempuan (b) Laki-laki	83
Gambar 4.28	<i>Scatter Plot</i> Regresi Kuantil untuk Indeks h dengan Jumlah Dokumen <i>Conditional on X</i> dimana dibedakan berdasarkan Pendidikan Terakhir (a) S3 (b) S2	85
Gambar 4.29	<i>Scatter Plot</i> Regresi Kuantil untuk Indeks h dengan Jumlah Dokumen <i>Conditional on X</i> dimana dibedakan berdasarkan Tempat Pendidikan Terakhir Dosen (a) DN (b) LN	86
Gambar 4.30	<i>Scatter Plot</i> (a) Variabel Jumlah Dokumen dengan Sitasi, (b) Variabel Usia dengan Sitasi (c) Variabel Lama bekerja dengan Sitasi, (d) Variabel Jumlah <i>co-authors</i> dengan Sitasi. Simbol \square Merupakan Dosen Laki-laki, dan Simbol \square Merupakan Simbol untuk Dosen Perempuan	83
Gambar 4.31	<i>Scatter Plot</i> (a) Variabel Jumlah Dokumen dengan Indeks h, (b) Variabel Usia dengan Indeks h, (c) Variabel Lama bekerja dengan Indeks h, (d) Variabel Jumlah <i>co-authors</i> dengan Indeks h. Simbol \square Merupakan Dosen Laki-laki, dan Simbol \square Merupakan Simbol untuk Dosen Perempuan.....	83
Gambar 4.32	<i>Scatter Plot</i> Variabel Indeks h dengan Sitasi. Simbol \square Merupakan Dosen Laki-laki, dan Simbol \square Merupakan Simbol untuk Dosen Perempuan.....	90

DAFTAR LAMPIRAN

	Halaman
Lampiran I. Data Publikasi Dosen ITS di Scopus	75
Lampiran II. Syntax Software R	76
1. Syntax Studi Simulasi untuk Data Homogen	77
2. Syntax Studi Simulasi untuk Data Heterogen	79
3. Syntax Regresi Linier Berbasis Model Rekursif	81
4. Syntax Regresi Kuantil Berbasis Model Rekursis Model Pertama ..	82
5. Syntax Regresi Kuantil Berbasis Model Rekursis Model Kedua	83
Lampiran III. Output Software R	85
1. Output Studi Simulasi Data Homogen	86
2. Output Studi Simulasi Data Heterogen	88
3. Output Model Regresi Linier Berbasis Model Rekursif	92
4. Output Model Regresi Linier Berbasis Model Rekursif Pemilihan Model Terbaik	94
5. Output Model Regresi Kuantil Berbasis Model Rekursif Model Pertama	96
6. Output Model Regresi Kuantil Berbasis Model Rekursif Model Kedua	98
Lampiran IV. Statistika Deskriptif	100
1. Output Statistika Deskriptif Indeks h Berdasarkan Jurusan	101
2. Output Statistika Deskriptif Sitasi Berdasarkan Jurusan	102
3. Output Analisis Faktor	103
4. Output Jumlah Masing-masing Anggota Cluster	103
5. Karakteristik Dosen per Cluster	104
6. Jumlah Dokumen Dosen ITS di Scopus	104
7. Persentase Kepemilikan Publikasi Berdasarkan Jenis Kelamin	105
8. Deskripsi Jabatan Dosen ITS per Jurusan	105
9. Kepemilikan Publikasi Berdasarkan Pendidikan terakhir	106
10. Persentase Kepemilikan Scopus Berdasarkan Tempat Pendidikan Terakhir	107

11. Deskripsi Jumlah <i>co-author</i>	108
12. Boxplot Variabel Independen dengan Variabel Dummy	109
13. Nilai Korelasi Antar Variabel	109
14. Hasil Uji Korelasi	110
Lampiran V. Regresi Kuantil	111
1. Estimasi Regresi Kuantil pada Setiap Kuantil untuk Variabel Dependen Sitasi	112
2. Estimasi Regresi Kuantil pada Setiap Kuantil untuk Variabel Dependen Indeks h.....	112
Lampiran VI. Langkah Analisis Menggunakan Stata 12	113

BAB I

PENDAHULUAN

1.1 Latar Belakang

Regresi kuantil adalah salah satu metode estimasi parameter yang dapat mengatasi sebaran data yang tidak seragam. Data yang memiliki pola tidak seragam disebut dengan heteroskedastisitas atau tidak terpenuhinya asumsi homoskedastisitas pada analisis regresi dengan metode *Ordinary Least Square* (OLS). Metode OLS kurang tepat digunakan untuk data yang teridentifikasi memiliki varians tidak konstan, sehingga terjadi kasus heteroskedastisitas. Sebaran data yang demikian dapat diatasi dengan metode regresi kuantil. Regresi kuantil pertama kali ditemukan oleh Koenker dan Baset (1978) yang merupakan perluasan dari metode regresi kuantil median, dimana distribusi kuantil bersyarat dari variabel respon dinyatakan sebagai fungsi dari kovariat yang diamati.

Koenker dan Hallock (2001) menyatakan bahwa regresi kuantil dapat digunakan untuk sebaran bersyarat yang asimetris, sebaran padat di bagian ekor, dan untuk sebaran yang terpotong. Chen dan Wei (2005) juga menjelaskan bahwa regresi kuantil sangat baik digunakan ketika ingin mengetahui tingkat perubahan pada kuantil bersyarat tertentu yang dinyatakan dengan nilai koefisiennya. Pada regresi kuantil prinsip yang digunakan adalah meminimumkan jumlah absolut residual yang dikenal dengan *Least Absolute Deviation* (LAD). Sedangkan estimasi parameter dapat dilakukan dengan tiga pendekatan, yaitu metode *direct*, *rank-score*, dan *resampling*. Pada metode *direct* taksiran interval yang digunakan adalah fungsi sparsity yang berbeda dengan pendekatan *rank-score* yang menggunakan *rank-score test*. Sedangkan pada metode *resampling* menggunakan taksiran interval dengan teknik bootstrap.

Zhou dan Portnoy (1996) melakukan penelitian mengenai sifat koefisien estimator pada regresi kuantil dengan tiga pendekatan melalui studi simulasi. Hasil dari penelitian tersebut menunjukkan bahwa *direct* menghasilkan penduga yang *robust* terhadap sebaran data yang tidak seragam dan memberikan hasil paling efisien untuk taksiran interval yang lebih sempit secara komputasi.

Kelebihan dari fungsi sparsity pada metode *direct* dapat menjelaskan penyebaran dari data dengan menggunakan ukuran kuantil sesuai sebaran data. Selain itu, taksiran interval pada metode *direct* berdasarkan kondisi asimptotik normal. Goh dan Knight (2009) juga melakukan penelitian mengenai regresi kuantil yang mengkaji mengenai batas asimptotik distribusi pada estimatornya.

Statistika inferensia pada regresi kuantil adalah hal yang penting bagi sebagian besar aplikasi, begitu juga dengan taksiran interval yang akan digunakan. Berbagai metode taksiran interval telah dikembangkan, seperti penelitian Chen dan Wei (2005) dan penelitian yang dilakukan Zhou dan Portnoy pada tahun 1996 diketahui bahwa metode *direct* hasil paling efisien untuk taksiran interval yang lebih sempit secara komputasi. Sedangkan untuk mendapatkan estimator yang konsisten pada taksiran interval yang sempit perlu dilakukan algoritma. Terdapat tiga algoritma yang dapat digunakan yaitu simpleks, *interior point*, dan *smoothing*. Dalam penelitian ini algoritma yang digunakan adalah simpleks, yang didasarkan pada penelitian Koenker dan D'Orey (1993) telah meneliti mengenai algoritma simpleks dan memberikan hasil bahwa algoritma simpleks merupakan algoritma yang paling stabil dan sesuai untuk jumlah observasi kurang dari 5000 dan 50 variabel.

Beberapa penelitian dengan menggunakan regresi kuantil pernah dilakukan oleh Utami (2013) yang mengaplikasikan pada data *Passenger Car Milage* yang terdapat pada buku Gujarti (2004), dimana pemodelan dilakukan pada kuantil 5% dan selanjutnya pada kuantil setiap kelipatan 5. Hasil yang diperoleh adalah semakin tinggi nilai kuantil yang digunakan model dikatakan lebih baik dikarenakan lebih banyak variabel yang signifikan dan menyesuaikan dengan sebaran data. Berikutnya dilakukan oleh Navianti (2014) mengenai regresi kuantil untuk pemodelan tingkat pengangguran terbuka di Indonesia, dan diperoleh hasil bahwa regresi kuantil terbukti dapat mengatasi kasus heteroskedastisitas. Penelitian mengenai penerapan regresi kuantil juga dilakukan oleh Wahyudi (2015) pada kasus IPM di pulau Jawa. Hasilnya menunjukkan bahwa regresi kuantil memiliki estimasi parameter yang bersifat *robust*, efisien dan konsisten.

Untuk penelitian regresi kuantil lainnya pernah dilakukan oleh Eide dan Showalter (1998) pada kasus pengaruh kualitas sekolah terhadap potensi siswa

berdasarkan nilai yang diperoleh dari hasil ujian. Hasil penelitian diketahui bahwa sebaran bersyarat dari hasil ujian siswa yang bersangkutan pada kuantil tertentu. Zhao, Yurgin, dan Winget (2009) menelitian mengenai perawatan yang sesuai untuk penderita jantung. Dengan membandingkan OLS dan regresi kuantil, hasil yang diperoleh regresi kuantil lebih baik dalam memodelkan pada kuantil tertentu sehingga sebaran pasien jantung dan perawatan yang sesuai dapat dijelaskan. Buchinsky (1994) mengaplikasikan regresi kuantil pada pemodelan upah di Amerika dengan distribusi bersyarat. Pada analisis tersebut dapat memodelkan upah dengan variabel yang berpengaruh yaitu pendidikan dan pengalaman kerja sesuai dengan sebaran data.

Berbagai bidang terapan yang telah menggunakan regresi kuantil, dalam penelitian ini menggunakan regresi kuantil berbasis model rekursif untuk memodelkan permasalahan pada publikasi dosen ITS di Scopus akibat adanya sebaran data yang tidak simetris. Model rekursif adalah suatu model yang dapat digunakan untuk model simultan, dimana variabel dependen pada model pertama menjadi variabel independen pada model kedua. Kondisi seperti ini lebih baik untuk memodelkan kedua variabel secara bersama dengan model yang disebut model rekursif. Model rekursif dapat digunakan dengan syarat bahwa antara residual pada model kedua dengan variabel dependen tidak memiliki hubungan (Gujarati, 2004). Seperti pada penelitian ini, variabel sitasi pada model kedua tidak memiliki hubungan dengan residual. Selain itu dalam model yang akan dibentuk, variabel sitasi mempengaruhi indeks h namun tidak sebaliknya. Keadaan inilah yang memperbolehkan OLS digunakan dalam model rekursif. Beberapa penelitian yang menggunakan regresi dengan model rekursif yaitu dilakukan oleh Rahmawati (2015) dan Sulistiyawati (2016).

Berkaitan dengan publikasi dan produktivitas, banyak penelitian yang dilakukan untuk mengetahui faktor-faktor yang mempengaruhi produktivitas penelitian dan publikasi. Penelitian tersebut diantaranya telah dilakukan oleh Pratt *et al.* (1999), Tien (2000), Sax *et al.* (2002), Hemmings dan Kay, (2010), Lei *et al.* (2014). Hemmings dan Kay (2010) meneliti tentang tingkat jurnal dan publikasi akademik di Australia, menyatakan kurang dari 50% dari universitas yang diteliti memiliki tingkat publikasi lebih dari 20% dari jurnal yang ada. Dari

penelitian tersebut diketahui bahwa seseorang yang memiliki jurnal lebih banyak cenderung memiliki tingkat publikasi yang lebih tinggi, senior cenderung memiliki tingkat publikasi lebih tinggi dibanding junior, dan seseorang dengan gelar doktoral memiliki tingkat publikasi paling tinggi dibandingkan yang bukan.

Penelitian lainnya yaitu Stack (2004) meneliti tentang akademisi yang mempunyai gelar doktoral menunjukkan bahwa publikasi akademisi perempuan lebih kecil dibandingkan akademisi laki-laki. Alasan yang memperkuat hasil penelitian tersebut adalah perempuan biasanya meluangkan banyak waktu dan energinya untuk mengurus rumah tangga dan anak (Sax *et al.*, 2002; Stack, 2004). Merujuk dari penelitian yang terkait dengan publikasi dosen ITS, Hal ini bertentangan dengan penelitian Suhartono *et al.* (2016) yang menyatakan bahwa dosen perempuan ITS mempunyai kecenderungan lebih besar menjadi penulis pertama dari publikasi di Scopus. Jenis kelamin dalam hal ini menjadi salah satu faktor dalam penelitian publikasi yang dilakukan akademisi.

Faktor tersebut semakin diperkuat dengan penelitian yang dilakukan oleh Rahmawati (2015) menyatakan faktor yang mempengaruhi publikasi dosen ITS adalah lama bekerja, jenis kelamin, pendidikan terakhir, dan jumlah dokumen. Model yang digunakan adalah model rekursif dengan kebaikan model sebesar 77,8%. Sedangkan penelitian lainnya dilakukan oleh Sulistiyawati (2016) tentang publikasi pada Scopus menyatakan bahwa faktor yang mempengaruhi sitasi adalah lama bekerja, jenis kelamin, pendidikan terakhir, dan jumlah dokumen. Model yang digunakan adalah regresi logistik dan model rekursif. Model yang dihasilkan penelitian tersebut masih mengandung pelanggaran asumsi yang disyaratkan. Penyebab terjadinya hal tersebut adalah karena sebaran data tidak simetris. Hal tersebut menjadi salah satu latar belakang dilakukan penelitian ini.

Dosen adalah pendidik profesional dengan tugas utama mentransformasikan, mengembangkan, dan menyebarluaskan ilmu pengetahuan, teknologi, dan seni melalui pendidikan, penelitian, dan pengabdian kepada masyarakat. Salah satu tugas utama dosen yaitu melakukan sebuah penelitian yang tertulis pada Peraturan Pemerintah Republik Indonesia nomor 37 tahun 2009 tentang dosen (Depkeu, 2009). Tugas utama Dosen untuk melakukan sebuah penelitian juga diatur dalam Tri Dharma Perguruan Tinggi yang menjadi salah satu visi dari seluruh perguruan

tinggi di Indonesia. Poin utama dalam Tri Dharma Perguruan Tinggi adalah pendidikan dan pengajaran, penelitian dan pengembangan, dan pengabdian kepada masyarakat. Penelitian adalah suatu kegiatan ilmiah yang didasarkan pada analisis dan dilakukan secara sistematis serta konsisten (Soekanto, 2014). Pada bulan April 2015, dikarenakan jumlah publikasi ilmiah perguruan tinggi di Indonesia lebih rendah dibandingkan dengan negara lain, maka Kemenristek Dikti memberikan tantangan kepada rektor untuk meningkatkan publikasi (Ristek, 2015).

Pada tanggal 1 Januari 2016, yang tertulis dalam peringkat *Webometrics* ITS adalah salah satu perguruan tinggi di Indonesia, yang menduduki peringkat 21 se-Indonesia dan peringkat 2939 di dunia. *Webometrics* adalah suatu sistem perankingan dunia yang berbasis web dengan menggunakan indikator *impact*, *presence*, *openess* dan *excellence*. Sehingga setiap dosen ITS harus meng-*update* indeks h sebagai salah satu syarat dalam Beban Kerja Dosen (BKD) yang terdapat pada sistem informasi kepegawaian tepatnya pada bagian rangkuman. Indeks h merupakan suatu indeks yang mulai sering digunakan untuk mengukur produktivitas seorang peneliti. Saat ini indeks h dihitung secara rinci dalam *database* khusus seperti Scopus dan *Web of Science* (WOS). Indeks h pertama kali diperkenalkan pada tahun 2005 oleh Jorge E. Hirsch. Selain itu, Archambault *et al.* (2009) pernah melakukan penelitian yang menunjukkan adanya korelasi yang kuat antara jumlah dokumen dan sitasi pada *Web of Science* (WoS) dan Scopus.

Scopus adalah bibliografi database yang berisi abstrak dan artikel, buku, dan makalah seminar. Informasi yang diperoleh pada akun Scopus diantaranya bibliografi, sitasi, dan indeks h. Indeks h merupakan indeks yang digunakan untuk mengukur dampak produktivitas dan sitasi dari pemilik akun. Indeks h adalah peneliti yang mempublikasikan paper sebanyak h, dengan sitasi untuk setiap paper tersebut minimal sama dengan h (Jenkins, 2015). Sedangkan sitasi adalah jumlah paper yang dikutip oleh peneliti lain. Perguruan tinggi yang sudah menggunakan indeks h selain ITS yaitu Institut Teknologi Bandung (ITB). ITB menggunakan indeks h sebagai salah satu faktor penting dalam perhitungan untuk mengajukan proposal penelitian dosen.

Berdasarkan uraian tersebut, maka pada penelitian ini terdapat dua kajian yang akan dibahas. Kajian pertama mengenai estimasi titik dan taksiran interval dengan metode *direct* fungsi sparsity pada regresi kuantil berbasis model rekursif. Kajian kedua akan membahas model untuk faktor-faktor yang mempengaruhi publikasi dosen ITS di Scopus. Penelitian ini diharapkan dapat memberi informasi berupa kajian estimasi parameter dan juga diharapkan dapat menjelaskan pola hubungan antara sitasi dan indeks h sehingga bisa menjadi bahan pertimbangan bagi ITS untuk meningkatkan produktivitas dosen ITS dalam publikasi internasional khususnya di Scopus.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan terkait dengan produktivitas dosen ITS yang dilihat berdasarkan publikasi pada Scopus. ITS adalah salah satu perguruan tinggi negeri yang menjalankan Tri Dharma Perguruan Tinggi berkaitan dengan penelitian dan karya ilmiah. Setiap dosen diwajibkan untuk mengisi indeks h pada system informasi kepegawaian ITS tepatnya pada bagian rangkuman. Hal ini dilakukan sebagai salah satu syarat beban kerja dosen (BKD) yang diukur melalui indeks h pada Scopus. Oleh karena itu, tujuan dari penelitian ini yang pertama adalah untuk mengetahui produktivitas dosen ITS di Scopus melalui model dari regresi kuantil. Tujuan kedua adalah untuk mengetahui estimasi titik serta taksiran interval metode *direct* fungsi sparsity. Sehingga, permasalahan yang diangkat dalam penelitian ini secara khusus sebagai berikut:

1. Bagaimana estimasi titik pada parameter regresi kuantil?
2. Bagaimana taksiran interval dengan metode *direct* fungsi sparsity pada parameter regresi kuantil?
3. Bagaimana aplikasi model regresi kuantil berbasis model rekursif dengan estimasi titik dan taksiran interval pada publikasi dosen ITS di Scopus?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dipaparkan, maka penelitian ini memiliki tujuan sebagai berikut.

1. Mengkaji estimasi titik pada parameter regresi kuantil.

2. Mengkaji taksiran interval dengan metode *direct* fungsi sparsity pada parameter regresi kuantil.
3. Mengaplikasikan model regresi kuantil berbasis model rekursif dengan estimasi titik dan taksiran interval pada publikasi dosen ITS di Scopus.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan informasi berupa kajian taksiran titik dan taksiran interval parameter regresi kuantil berbasis model rekursif, dan memberi informasi berupa model dari publikasi dosen ITS di Scopus. Selain itu, dari penelitian ini diharapkan dapat diketahui faktor-faktor yang berpengaruh terhadap sitasi dan indeks h. Dengan demikian, pengambilan kebijakan di ITS dapat menggunakan informasi tersebut sebagai dasar untuk meningkatkan publikasi dosen ITS dalam publikasi internasional khususnya di Scopus.

1.5 Batasan Penelitian

Pada penelitian ini, ruang lingkup permasalahan dibatasi dengan beberapa hal sebagai berikut.

1. Estimasi parameter pada regresi kuantil dengan model rekursif hanya mengkaji estimasi titik dan taksiran interval dengan metode *direct* fungsi sparsity.
2. Kuantil yang digunakan dibatasi yaitu 5%, 10%, 25%, 50%, 75%, 90%, dan 95%.
3. Data primer yang digunakan pada penelitian ini adalah data publikasi dosen ITS di Scopus pada bulan Juli 2016.
4. Objek penelitian adalah seluruh dosen ITS, kecuali dosen MKU dan UPT.
5. Penelitian hanya menggunakan variabel sitasi dan indeks h sebagai variabel dependen.

(Halaman ini sengaja dikosongkan)

BAB II

KAJIAN PUSTAKA

Pada bab ini akan diberikan penjelasan teori mengenai regresi linier berganda, model rekursif, dan regresi kuantil berbasis model rekursif serta beberapa teori yang berkaitan dengan penelitian ini.

2.1 Regresi Linier Berganda

Analisis regresi membahas mengenai hubungan antara variabel yang satu dengan yang lainnya. Analisis regresi berganda adalah metode statistik yang dapat digunakan untuk menganalisis hubungan antara satu variabel dependen dengan beberapa variabel independen. Dengan menggeneralisirkan model regresi linier dua dan tiga variabel, model regresi variabel dependen Y dengan p variabel independen X_1, X_2, \dots, X_p dapat ditulis

$$Y_i = S_0 + S_1 X_{1i} + S_2 X_{2i} + \dots + S_p X_{pi} + v_i, \quad (2.1)$$

dengan $i=1,2,\dots,n$ dengan S_0 adalah *intercept*, S_1 sampai dengan S_p menyatakan *slope*, v_i menyatakan nilai residual. Persamaan (2.1) merupakan persamaan sederhana dari n persamaan simultan

$$\begin{aligned} Y_1 &= S_0 + S_1 X_{11} + S_2 X_{21} + \dots + S_p X_{p1} + v_1 \\ Y_2 &= S_0 + S_1 X_{12} + S_2 X_{22} + \dots + S_p X_{p2} + v_2 \\ &\vdots \\ Y_n &= S_0 + S_1 X_{1n} + S_2 X_{2n} + \dots + S_p X_{pn} + v_n, \end{aligned} \quad (2.2)$$

dengan residual memenuhi asumsi :

1. $E(v_i) = 0$ untuk $i = 1, 2, \dots, n$.
2. $Var(v_i) = \sigma^2$ untuk $i = 1, 2, \dots, n$.
3. $Cov(v_i, v_j) = 0, i \neq j$.

Persamaan (2.2) dapat ditulis dalam bentuk matrik sesuai persamaan (2.3) berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ \vdots \\ S_p \end{bmatrix} + \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}. \quad (2.3)$$

Persamaan (2.3) dapat disederhanakan seperti persamaan (2.4) berikut (Gujarati, 2004: 926-928).

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{v}, \quad (2.4)$$

dengan

\mathbf{y} : vektor kolom $n \times 1$ dari observasi pada variabel dependen Y

\mathbf{X} : matrik $n \times (p+1)$ dengan n observasi pada p variabel X_1 sampai X_p

$\boldsymbol{\beta}$: vektor kolom $(p+1) \times 1$ parameter s_0, s_1, \dots, s_p

\mathbf{v} : vektor kolom $n \times 1$ residual v_i .

2.1.1 Estimasi Parameter Regresi Linier

Untuk mengestimasi OLS dari parameter $\boldsymbol{\beta}$, terlebih dahulu menuliskan sampel regresi

$$Y_i = \hat{S}_0 + \hat{S}_1 X_{1i} + \hat{S}_2 X_{2i} + \dots + \hat{S}_p X_{pi} + \hat{v}_i, \quad (2.5)$$

dengan notasi matriks

$$\mathbf{y} = \mathbf{X} \hat{\boldsymbol{\beta}} + \hat{\mathbf{v}}. \quad (2.6)$$

$\hat{\boldsymbol{\beta}}$ adalah vektor $(p+1)$ elemen dari estimator OLS dan $\hat{\mathbf{v}}$ merupakan vektor kolom residual $n \times 1$. Untuk mengetahui estimasi OLS dengan meminimalisir

$$\sum_{i=1}^n \hat{v}_i^2 = \sum_{i=1}^n (Y_i - \hat{S}_0 - \hat{S}_1 X_{1i} - \dots - \hat{S}_p X_{pi})^2, \quad (2.7)$$

dengan $\sum_{i=1}^n \hat{v}_i^2$ adalah *Sum of square residual*. Pada notasi matriks, persamaan

(2.7) sama dengan meminimalisir $\hat{\mathbf{v}}^T \hat{\mathbf{v}}$ karena

$$\hat{\mathbf{v}}^T \hat{\mathbf{v}} = [\hat{v}_1 \quad \hat{v}_2 \quad \cdots \quad \hat{v}_n] \begin{bmatrix} \hat{v}_1 \\ \hat{v}_2 \\ \vdots \\ \hat{v}_n \end{bmatrix} = \hat{v}_1^2 + \hat{v}_2^2 + \dots + \hat{v}_n^2 = \sum_{i=1}^n \hat{v}_i^2. \quad (2.8)$$

dari persamaan (2.6) didapatkan rumus sesuai dengan persamaan (2.9)

$$\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{X}}. \quad (2.9)$$

Oleh karena itu,

$$\begin{aligned} \hat{\mathbf{e}}^T \hat{\mathbf{e}} &= (\mathbf{y} - \hat{\mathbf{X}})^T (\mathbf{y} - \hat{\mathbf{X}}) \\ &= \mathbf{y}^T \mathbf{y} - 2 \hat{\mathbf{X}}^T \mathbf{y} + \hat{\mathbf{X}}^T \hat{\mathbf{X}}. \end{aligned}$$

Metode OLS mengestimasi s_0, s_1, \dots, s_p sehingga $\sum_{i=1}^n \hat{v}_i^2$ memiliki nilai seminimum mungkin dengan mendiferensialkan persamaan (2.7) secara parsial terhadap s_0, s_1, \dots, s_p dan menyamadengankan hasil yang diperoleh dengan nol (Gujarati, 2004: 931-933). Dalam notasi matriks dapat ditulis

$$\begin{bmatrix} n & \sum X_{li} & \cdots & \sum X_{pi} \\ \sum X_{li} & \sum X_{li}^2 & \cdots & \sum X_{pi} X_{li} \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{pi} & \sum X_{pi} X_{li} & \cdots & \sum X_{pi}^2 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_p \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & \cdots & X_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{pn} & \cdots & X_{pn} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

atau

$$(\mathbf{X}^T \mathbf{X}) \hat{\mathbf{e}} = \mathbf{X}^T \mathbf{y}. \quad (2.10)$$

Pada persamaan (2.10) diketahui $\mathbf{X}^T \mathbf{X}$ dan $\mathbf{X}^T \mathbf{y}$ merupakan perkalian antara variabel X dan Y dan yang belum diketahui adalah $\hat{\mathbf{e}}$. Invers dari $\mathbf{X}^T \mathbf{X}$ adalah $(\mathbf{X}^T \mathbf{X})^{-1}$, maka hasil perkalian kedua sisi persamaan (2.10) adalah

$$(\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X}) \hat{\mathbf{e}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (2.11)$$

Karena $(\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{X}) = \mathbf{I}$, suatu matriks identitas order $(p+1) \times (p+1)$ diperoleh

$$\hat{\mathbf{e}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y},$$

atau

$$\hat{\mathbf{e}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (2.12)$$

2.1.2 ANOVA

ANOVA digunakan untuk menguji signifikansi hasil estimasi secara serentak pada analisis regresi. Selain itu, ANOVA juga digunakan untuk menilai kontribusi variabel independen. Perhitungan ANOVA disajikan pada Tabel 2.1 berikut.

Tabel 2.1 ANOVA

<i>Source of variation</i>	<i>df</i>	<i>SS</i>	<i>MS</i>
<i>Regression</i>	p	$\hat{\mathbf{y}}^T \mathbf{X}^T \mathbf{y} - n\bar{Y}^2$	$\frac{\hat{\mathbf{y}}^T \mathbf{X}^T \mathbf{y} - n\bar{Y}^2}{p}$
<i>Error</i>	$n - p - 1$	$\mathbf{y}^T \mathbf{y} - \hat{\mathbf{y}}^T \mathbf{X}^T \mathbf{y}$	$\frac{\mathbf{y}^T \mathbf{y} - \hat{\mathbf{y}}^T \mathbf{X}^T \mathbf{y}}{n - p - 1}$
<i>Total</i>	$n - 1$	$\mathbf{y}^T \mathbf{y} - n\bar{Y}^2$	

Untuk analisis regresi berganda terdapat beberapa asumsi yang harus dipenuhi. Asumsi tersebut antara lain: (1) Model regresi linier (model regresi linier dalam parameter), (2) Prediktor dalam analisis regresi merupakan suatu nilai yang tetap (*fixed factor*), (3) Hasil ekspektasi dari residual bernilai nol, (4) Residual dari analisis regresi identik (variansi dari residual bersifat tetap), (5) Tidak terjadi *autocorrelation* (antara residual yang satu dengan yang lain saling bebas), (6) *Covariance* antara residual dengan prediktor bernilai nol, (7) banyaknya observasi (n) dalam analisis harus melebihi jumlah parameter (p) dari hasil estimasi, (8) Tidak terjadi *multicollinearity* (hubungan yang sangat kuat diantara prediktor) (Gujarati, 2004:65).

2.1.3 Uji Korelasi

Korelasi digunakan untuk mengetahui hubungan antara kedua variabel yang terdapat pada model regresi. Nilai korelasi adalah antara 1 dan -1, jika nilai korelasi sama dengan 0 maka tidak ada korelasi antara kedua variabel yang di uji. Hipotesis yang digunakan pada pengujian korelasi adalah (Walpole, 1995: 371).

$H_0 : \dots = 0$ (tidak ada hubungan antara variabel).

$H_1 : \dots \neq 0$ (ada hubungan antara variabel).

Statistik uji korelasi adalah

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}, \quad (2.13)$$

dengan

$$r = \frac{n \sum_{i=1}^n X_i Y_i - (\sum_{i=1}^n X_i)(\sum_{i=1}^n Y_i)}{\sqrt{[n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2][n \sum_{i=1}^n Y_i^2 - (\sum_{i=1}^n Y_i)^2]}}$$

Daerah Kritis : Tolak H_0 apabila nilai $|t|$ lebih dari $t_{(\alpha/2, n-2)}$.

2.2 Model Rekursif

Metode OLS tidak dapat digunakan untuk mengestimasi model regresi dengan persamaan simultan sebagai berikut:

$$\begin{aligned} Y_{1i} &= S_{10} + S_{12}Y_{2i} + X_{11}X_{1i} + V_{1i} \\ Y_{2i} &= S_{20} + S_{21}Y_{1i} + X_{21}X_{1i} + V_{2i} \end{aligned} \quad (2.14)$$

Pada persamaan simultan hubungan sebab akibat antara variabel independen dan dependen menghasilkan nilai yang berbeda. Dalam hal ini, lebih baik untuk memodelkan kedua variabel secara bersama. Jika persamaan simultan digunakan, maka akan menghasilkan estimasi yang bias dan tidak konsisten sehingga seberapa besar pun sampel yang digunakan, bias yang terjadi tidak akan hilang. Namun terdapat keadaan dimana OLS dapat diaplikasikan secara benar, yaitu dalam pengaplikasian model rekursif. Model simultan dapat diaplikasikan dalam model rekursif ketika hubungan dari variabel independen pertama dan kedua adalah serarah, yaitu Y_1 mempengaruhi Y_2 dan tidak sebaliknya. Persamaan model rekursif sebagai berikut:

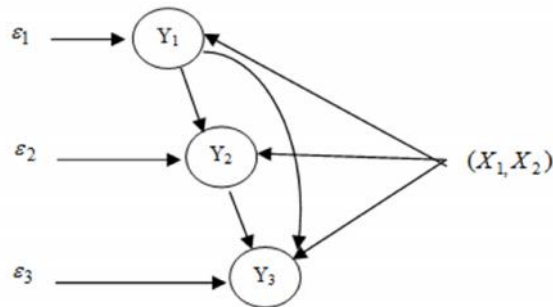
$$\begin{aligned} Y_{1i} &= S_{10} + X_{11}X_{1i} + X_{12}X_{2i} + V_{1i} \\ Y_{2i} &= S_{20} + S_{21}Y_{1i} + X_{21}X_{1i} + X_{22}X_{2i} + V_{2i} \\ Y_{3i} &= S_{30} + S_{31}Y_{1i} + S_{32}Y_{2i} + X_{31}X_{1i} + X_{32}X_{2i} + V_{3i}, \end{aligned} \quad (2.15)$$

dengan Y adalah variabel dependen dan X adalah variabel independen. Kovarian dari residual adalah sebagai berikut:

$$\text{cov}(V_{1i}, V_{2i}) = \text{cov}(V_{1i}, V_{3i}) = \text{cov}(V_{2i}, V_{3i}) = 0. \quad (2.16)$$

Persamaan (2.16) menunjukkan bahwa residual periode yang sama pada persamaan yang berbeda tidak memiliki hubungan. Berdasarkan persamaan pertama dari model persamaan (2.15), persamaan tersebut hanya mengandung variabel independen dan tidak berkorelasi dengan residual V_{1i} , maka persamaan pertama memenuhi asumsi klasik dari OLS sehingga OLS dapat diaplikasikan

langsung pada persamaan ini. Untuk persamaan kedua dari persamaan (2.15) yang mengandung Y_1 sebagai variabel independen juga dapat menggunakan OLS, dengan syarat Y_{1i} dan v_{2i} tidak berkorelasi. Dengan mengembangkan syarat tersebut, estimasi OLS dapat diaplikasikan terhadap persamaan ketiga dari persamaan (2.15), karena baik Y_{1i} dan Y_{2i} tidak berkorelasi dengan v_{3i} . Oleh karena itu pada model rekursif OLS dapat digunakan dalam masing-masing persamaan secara terpisah seperti terlihat pada Gambar 2.1 (Gujarati, 2004: 764-766).



Gambar 2.1 Model Rekursif.

2.3 Regresi Kuantil

Regresi kuantil pertama kali diperkenalkan oleh Koenker dan Basset pada tahun 1978. Metode ini merupakan perluasan dari model regresi pada kuantil bersyarat dimana distribusi kuantil bersyarat dari variabel dependen yang dinyatakan sebagai fungsi dari kovariat yang diamati. Dengan pendekatan ini, dapat memungkinkan menduga fungsi kuantil dari sebaran bersyarat variabel dependen pada setiap nilai kuantil sesuai dengan kuantil yang diinginkan (Chen, 2005). Karena sifatnya yang *robust* terhadap data *outlier* maka regresi kuantil sangat dianjurkan untuk menganalisis sejumlah data yang tidak simetris serta memiliki distribusi yang tidak homogen. Taksiran interval pada regresi kuantil dapat dilakukan dengan beberapa metode, yaitu dengan pendekatan *direct*, *rank-score*, dan *resampling* (Koenker, 2005:14, 88-105).

Misalkan diberikan data $\{X_{1i}, X_{2i}, \dots, X_{pi}, Y_i\}, i = 1, 2, \dots, n$ merupakan himpunan berpasangan dari variabel random yang berdistribusi secara independen

dan tidak identik dengan kuantil $\dagger \in (0,1)$. Data ini memiliki fungsi sebaran peluang bersyarat $F(Y | X_i) = P(Y \leq y | X_i)$ dan fungsi invers $F^{-1}(\dagger) = \inf \{y : F(y) \geq \dagger\}$ yang merupakan kuantil ke $-\dagger$ didefinisikan sebagai $Q(\dagger) = \inf \{y : F(y) \geq \dagger\} = F^{-1}(\dagger)$ yang merupakan fungsi kuantil ke $-\dagger$ dari variabel dependen (Y) (Goh dan Knight, 2009). Persamaan umum regresi kuantil linier khusus untuk kuantil bersyarat $Q_{Y_i}(\dagger | X_{1i}, X_{2i}, \dots, X_{pi})$ dari variabel dependen Y_i yaitu:

$$Y_i = s_0(\dagger) + s_1(\dagger)X_{1i} + \dots + s_p(\dagger)X_{pi} + v_i(\dagger), \quad (2.17)$$

dengan $i = 1, 2, \dots, n$.

Apabila model regresi kuantil disajikan dalam bentuk matriks, persamaan (2.17) dapat ditulis sebagai berikut:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix} \begin{bmatrix} s_0(\dagger) \\ s_1(\dagger) \\ \vdots \\ s_p(\dagger) \end{bmatrix} + \begin{bmatrix} v_1(\dagger) \\ v_2(\dagger) \\ \vdots \\ v_n(\dagger) \end{bmatrix}, \quad (2.18)$$

Selanjutnya persamaan (2.18), dapat ditulis dalam bentuk model linier berikut:

$$\mathbf{y} = \mathbf{X}(\dagger) + \mathbf{v}. \quad (2.19)$$

Jika fungsi bersyarat dari kuantil ke $-\dagger$ dengan variabel independen X tertentu, maka fungsi bersyarat tersebut didefinisikan sebagai berikut :

$$\begin{aligned} Q_{Y_i}(\dagger | X_{1i}, X_{2i}, \dots, X_{pi}) &= Q_Y(\dagger | X) \\ &= \mathbf{X}^T(\dagger), i = 1, 2, \dots, n. \end{aligned} \quad (2.20)$$

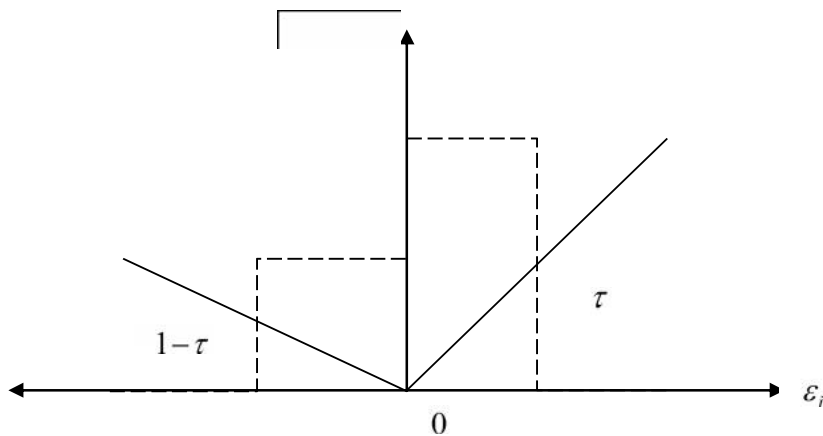
Maka solusi optimasi pada regresi kuantil adalah sebagai berikut :

$$\min_{\mathbf{s} \in \mathbb{R}^p} \sum_{i=1}^n \dots(\dagger)(\mathbf{y} - \mathbf{X}^T(\dagger)), \quad i = 1, 2, \dots, n \quad \dagger \in (0,1) \quad (2.21)$$

dimana $y_i = \{y_1, y_2, \dots, y_n\}$ merupakan sampel random dengan variabel dependen (Y) dan $x_i \in \mathbb{R}^p$ merupakan vektor kovariat, sedangkan $\dots(\dagger)(v) = v(\dagger - I(v < 0))$ $0 < \dagger < 1$ yang merupakan *loss function* yang asimetrik dari v yang merupakan residual dari estimasi parameter (Koenker, 2005:5).

2.3.1 Estimasi Parameter pada Regresi Kuantil

Estimasi parameter dalam regresi OLS, hanya dapat digunakan untuk memberi solusi permasalahan *mean*, sehingga Koenker dan Basset (1978) mengembangkan metode alternatif yaitu regresi kuantil. Regresi kuantil merupakan pengembangan dari regresi kuantil median. Regresi OLS diestimasi dengan meminimumkan jumlah kuadrat residual, sedangkan regresi kuantil akan meminimumkan jumlah absolut residual yang lebih dikenal dengan *Least Absolute Deviation* (LAD).



Gambar 2.2 Loss Function.

Pada regresi kuantil median, residual diberi bobot yang sama. Sedangkan pada regresi kuantil residual diberi bobot yang berbeda seperti yang terlihat pada Gambar 2.2. Bobot yang digunakan yaitu τ untuk nilai residual yang lebih besar atau sama dengan nol, dan $(1-\tau)$ untuk residual yang kurang dari nol (Koenker and Hallock, 2001: 3). Perkalian antara residual dengan bobot yang diberikan membentuk *loss function* ($\dots(\tau)$) yaitu:

$$\dots(\tau)(v) = \sum_{i=1, v_i \geq 0}^n \tau |v_i| + \sum_{i=1, v_i < 0}^n (1-\tau) |v_i|. \quad (2.22)$$

dengan demikian, dalam regresi kuantil terdapat fungsi kuantil bersyarat ke- τ yang mempertimbangkan penduga $\hat{S}(\tau)$, sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan sebagai berikut:

$$\min_{S \in \mathbb{R}^p} \sum_{i=1}^n \dots(\tau)(\mathbf{y} - Q(\tau)(y | X)); \tau \in (0,1) \quad (2.23)$$

dengan :

$\rho_{\tau}(\cdot)$: *loss function* yang asimetris

τ : Indeks kuantil dengan $\tau \in (0,1)$

$Q(\tau)(y | X)$: fungsi kuantil ke τ dari variabel Y dengan syarat X .

Dalam regresi kuantil, pada kuantil ke τ dari F_y meminimumkan *loss function* dari persamaan (2.22) adalah

$$\begin{aligned}\hat{S}(\tau) &= \min_{S \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau}(v_i) \\ &= \min_{S \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau}(\mathbf{y} - \mathbf{X}^T(S)),\end{aligned}\tag{2.24}$$

dimana $\tau \in (0,1)$ dan $\rho_{\tau}(v_i)$ pada persamaan (2.24) didefinisikan

$$\rho_{\tau}(v) = \begin{cases} \tau v & , \text{ jika } v \geq 0 \\ (\tau - 1)v & , \text{ jika } v < 0. \end{cases}\tag{2.25}$$

Jika Y merupakan fungsi X yang diketahui dan memiliki fungsi probabilitas $F_{Y|X}(y)$, maka kuantil ke τ dari fungsi tersebut dapat dituliskan seperti pada persamaan berikut :

$$\min_S \tau \int_{i=1; v \geq 0}^n |v_i| dF_Y(y) + (1-\tau) \int_{i=1; v < 0}^n |v_i| dF_Y(y),\tag{2.26}$$

dengan mempertimbangkan $\hat{S}(\tau)$, sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan

$$\hat{S}(\tau) = \min_{S \in \mathbb{R}^p} \left\{ \tau \sum_{i=1; v \geq 0}^n |\mathbf{y} - \mathbf{X}^T(S)| + (1-\tau) \sum_{i=1; v < 0}^n |\mathbf{y} - \mathbf{X}^T(S)| \right\}.\tag{2.27}$$

2.3.2 Optimasi dengan Algoritma Simpleks pada Model Regresi Kuantil

Algoritma simpleks adalah salah satu metode pendugaan parameter secara numerik untuk regresi kuantil yang telah dikembangkan oleh Barrodam dan Robert pada tahun 1974. Metode simpleks merupakan cara untuk menentukan kombinasi optimal dari tiga variabel atau lebih, dimana algoritma ini dapat memberikan solusi permasalahan program linier dengan komputasi serta melibatkan banyak variabel (Davino, Furno, dan Vistocco, 2014: 23-31). Bentuk

umum formulasi metode primal dari permasalahan regresi kuantil seperti pada persamaan (2.21)

$$\min_{\mathbf{s} \in \mathbb{R}^p} \sum_{i=1}^n \dots (\ddagger) (\mathbf{y} - \mathbf{X}^T \mathbf{s}).$$

Diberikan $[\mathbf{x}]_+$, non negatif bagian dari \mathcal{X} , yaitu $s_1 = [\mathbf{y} - \mathbf{X}]_+$ dan $s_2 = [\mathbf{X} - \mathbf{y}]_+$ maka masalah linier dapat ditulis

$$\min_{\mathbf{s} \in \mathbb{R}^p} \{1^T \mathbf{s}_1 + 1^T \mathbf{s}_2 \mid \mathbf{y} = \mathbf{X} \mathbf{s}_1 - \mathbf{X} \mathbf{s}_2, (\mathbf{s}_1, \mathbf{s}_2) \in \mathbb{R}_+^n\},$$

selanjutnya diberikan $\mathbf{B} = [\mathbf{X} \quad -\mathbf{X} \quad -\mathbf{I}]$, dan

$$\mathbf{C} = \begin{bmatrix} [] \\ [-] \\ [\mathbf{y} - \mathbf{X}] \\ [\mathbf{X} - \mathbf{y}] \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 0_{[p]} \\ 0_{[p]} \\ 1_{[n]} \\ 1_{[n]} \end{bmatrix}.$$

Seperti reformulasi dari permasalahan pemrograman linier standar, maka formulasinya dapat ditulis $\min_{\mathbf{z}} \mathbf{y}^T \mathbf{z}$ dimana untuk $\mathbf{B}^T \mathbf{z} \leq \mathbf{d}$. Oleh karena itu parameter dual dapat ditulis $\max_{\mathbf{C}} \mathbf{C}^T \mathbf{d}$ dimana untuk $\mathbf{B} \mathbf{C} = \mathbf{y}, \ddagger \geq 0$. Secara sederhana masalah diatas dapat dirumuskan seperti pada persamaan (2.28) berikut:

$$\max_{\mathbf{z}} \left\{ \mathbf{y}^T \mathbf{z} \mid \mathbf{X}^T \mathbf{z} = \mathbf{0}, \mathbf{z} \in [-1, +1]^n \right\}, \quad (2.28)$$

yang ditransformasikan seperti pada persamaan (2.29) berikut:

$$\mathbf{X}^T \left(\frac{1}{2} \mathbf{z} + \frac{1}{2} \mathbf{1} \right) = \frac{1}{2} \mathbf{X}^T \mathbf{1} \mathbf{X}^T. \quad (2.29)$$

Jika $\frac{1}{2} \mathbf{z} + \frac{1}{2} \mathbf{1} = \mathbf{y}$ dan $\frac{1}{2} \mathbf{X}^T \mathbf{1} = \mathbf{b}$, maka persamaan dapat ditulis $\mathbf{X}^T \mathbf{y} = \mathbf{b}$.

Nilai 1/2 pada persamaan (2.29) merupakan kunci utama generalisasi untuk kondisi kuantil ke $-\ddagger$ sehingga diperoleh

$$\min_{\mathbf{s}} \sum_{i=1}^n \dots \ddagger (y_i - x_i^T \mathbf{s} (\ddagger)). \quad (4.30)$$

Dengan cara yang sama diperoleh formulasi dari persamaan dual

$$\max_{\mathbf{z}} \left\{ \mathbf{y}^T \mathbf{z} \mid \mathbf{X}^T \mathbf{z} = (1-\dagger) \mathbf{X}^T \mathbf{1}, \mathbf{z} \in [0,1]^n \right\},$$

dimana $(1-\dagger)$ memiliki peran yang sama seperti $\frac{1}{2}$ pada rumus di median.

2.3.3 Konstruksi Taksiran Interval pada Regresi Kuantil

Beberapa metode dalam regresi kuantil yang digunakan untuk menghitung selang kepercayaan S salah satunya yaitu fungsi *sparsity*. Pada setiap \dagger yang sudah ditentukan, taksiran interval dengan formula sebagai berikut (Davino, Furno, dan Vistocco, 2014: 132-134) :

$$P\left(\hat{S}_j(\dagger) - t_{(r/2, df)} se(\hat{S}_j(\dagger)) \leq S_j(\dagger) \leq \hat{S}_j(\dagger) + t_{(r/2, df)} se(\hat{S}_j(\dagger))\right) = 1-r. \quad (2.31)$$

Nilai $se(\hat{S}_j(\dagger))$ adalah nilai diagonal yang diperoleh dari perhitungan matrik $\tilde{S}^2(\dagger)\mathbf{D}^{-1}$, dimana $\mathbf{D}^{-1} = (\mathbf{X}^T \mathbf{X})^{-1}$ dan $\tilde{S}^2(\dagger) = \dagger(1-\dagger) / f(F^{-1}(\dagger))^2$. Untuk F adalah fungsi distribusi kumulatif dan $f = F'$ adalah fungsi densitas, maka fungsi *sparsity* dinotasikan sebagai berikut:

$$S(\dagger) = f(F^{-1}(\dagger))^{-1}. \quad (2.32)$$

2.3.4 Pengujian Hipotesis

Karena parameter pada kuantil ke $-\dagger$ dimungkinkan sangat banyak pada $\dagger \in (0,1)$, sehingga perlu diuji signifikansi pada parameter untuk mengetahui pengaruh antara variabel independen dengan variabel dependen. Diberikan model regresi kuantil sebagai berikut :

$$\mathbf{y} = \mathbf{X}^T(\dagger) + \epsilon, \quad (2.33)$$

dengan hipotesis:

$$H_0 : S_j(\dagger) = 0$$

$$H_1 : S_j(\dagger) \neq 0, j = 1, 2, \dots, p.$$

Statistik uji yang digunakan yaitu uji t (Davino, Furno, dan Vistocco, 2014: 134)

$$t = \frac{\hat{S}_j(\dagger)}{se(\hat{S}_j(\dagger))}. \quad (2.34)$$

H_0 ditolak apabila $|t| > t_{(r/2, df)}$ dengan derajat bebas (*degree of freedom*) tertentu. *Standart error* pada uji t merupakan salah satu komponen untuk membentuk konfidensi interval dengan formula seperti pada persamaan (2.31).

2.3.5 Kriteria Keباikan Model Regresi Kuantil

Kriteria kebaikan model regresi kuantil menggunakan *backtesting procedure*. *Backtesting* biasa digunakan sebagai pendekatan untuk mengetahui kesesuaian model pada regresi kuantil. Tujuan dari *backtesting procedure* adalah untuk mengukur akurasi dari estimator kuantil ($\hat{y}_\dagger | \mathbf{x}$). Model memiliki akurasi yang baik apabila $\dagger^* = \dagger$, dimana \dagger adalah kuantil dari variabel respon dan $\dagger^* = P((y | X) \leq (\hat{y}_\dagger | X))$.

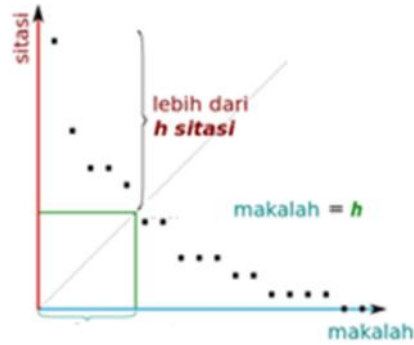
2.4 Institut Teknologi Sepuluh Nopember

Institut Teknologi Sepuluh Nopember (ITS) adalah salah satu perguruan tinggi negeri yang terletak di Surabaya. Pendirian ITS dipelopori oleh para pejuang kemerdekaan Republik Indonesia pada tahun 1957, melibatkan dr. Angka Nitisastro, Soedjasmono, K.H. Yahya Hasyim, dan didukung oleh Roeslan Abdulgani. ITS didirikan oleh Yayasan Perguruan Tinggi Teknik (YPTT) yang diketuai oleh dr. Angka Nitisastro tepatnya pada tanggal 10 November 1957. Yayasan tersebut dibentuk sebagai wadah untuk memikirkan tindakan-tindakan lebih lanjut dan memperbincangkan segala konsekuensi yang berkaitan dengan pengambilan keputusan dalam rangka mendirikan sebuah Perguruan Tinggi Teknik di Kota Surabaya.

ITS diresmikan oleh Presiden Soekarno, dan memiliki dua jurusan yaitu, Jurusan Teknik Sipil dan Teknik Mesin. Beberapa tahun kemudian melalui usaha-usaha yang dirintis oleh para tokoh YPTT, Perguruan Tinggi Teknik 10 Nopember dirubah statusnya menjadi Perguruan Tinggi Negeri dengan nama: “Institut Teknologi Sepuluh Nopember”. ITS yang semula memiliki 2 (dua) jurusan berubah menjadi lima yaitu: Teknik Sipil, Teknik Elektro, Teknik Mesin, Teknik Perkapalan, dan Teknik Kimia. ITS memiliki staf pengajar profesional di bidangnya yang terdiri dari profesor, doktor dan master yang berasal dari lulusan perguruan tinggi terkemuka di luar maupun di dalam negeri. Visi ITS adalah mendidik para pemimpin yang unggul di bidang sains dan teknologi. Pada tahun 2013 ITS telah berkembang pesat hingga memiliki 26 jurusan/program studi tingkat sarjana dari 5 fakultas. Selain itu, ITS juga memiliki 12 program doktor, 18 program magister, 6 program studi diploma, dan 1 program studi D-4.

2.5 Publikasi Scopus

Scopus adalah bibliografi database yang berisi abstrak dan artikel, buku, dan makalah seminar. Scopus dimiliki oleh Elsevier yang tersedia secara *online* dengan berlangganan, yang terdiri dari empat *subject areas*, yaitu *life sciences*, *health sciences*, *physical sciences* dan *social sciences & humanities*. Pada scopus terdapat informasi diantaranya bibliografi, sitasi, dan indeks h. Indeks h merupakan indeks yang digunakan untuk mengukur dampak produktivitas dan sitasi dari pemilik akun. Indeks h diperkenalkan oleh Jorge E. Hirsch pada tahun 2005. Menurut Hirsch (2007) menyatakan bahwa indeks h merupakan indikator yang lebih baik dari indikator lainnya (sitasi tiap dokumen dan jumlah dokumen) dalam pemberian penghargaan kepada ilmuwan/akademisi. Gambar 2.3 dapat memberikan informasi mengenai nilai indeks h seorang penulis. Apabila penulis tersebut memiliki h dokumen dan setidaknya memiliki h sitasi, maka seorang penulis memiliki indeks h sebesar h dengan menerbitkan h dokumen dengan masing-masing dokumen telah di-sitasi dalam karya-karya lain setidaknya sebanyak h kali (Hirsch, 2005).



Gambar 2.3 Perhitungan Indeks h.

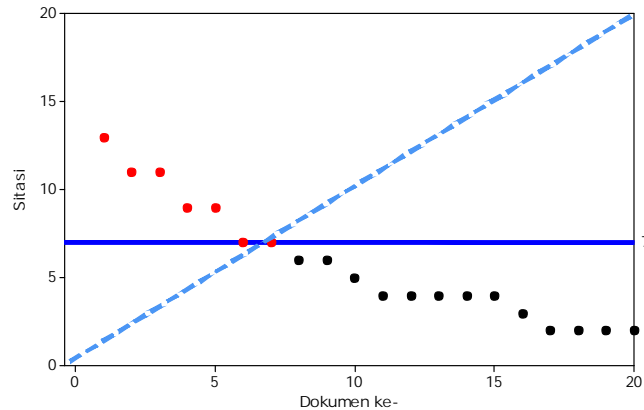
Indeks h mencerminkan jumlah publikasi dan sitasi per publikasi. Sebagai contoh, dosen “S” memiliki publikasi di scopus dengan jumlah dokumen sebanyak 38 dan jumlah kutipan sebesar 119. Informasi mengenai sitasi dan dokumen secara rinci dapat dilihat pada Tabel 2.2.

Tabel 2.2 Sitasi Tiap Dokumen Dosen “S”

Dokumen ke-	Sitasi	Dokumen ke-	Sitasi	Dokumen ke-	Sitasi
1	13	7	7	13	4
2	11	8	6	14	4
3	11	9	6	15	4
4	9	10	5	16	3
5	9	11	4	17	2
6	7	12	4	18	2

Tabel 2.2 merupakan data sitasi dari masing-masing dokumen milik dosen “S” yang kemudian akan dibuat grafik untuk mengetahui indeks h dosen tersebut.

Data tersebut diperoleh dengan mengurutkan 38 dokumen dari yang terbesar atau memiliki sitasi tertinggi. Kemudian dibuat plot dimana sumbu x menunjukkan dokumen dan sumbu y menunjukkan sitasi. Gambar 2.4 berikut terdapat lima titik berwarna merah yang berada diatas garis biru dan dua titik berada tepat digaris biru. Tujuh titik berwarna merah tersebut menunjukkan indeks h dari dosen “S”, yang artinya dosen tersebut telah menerbitkan tujuh dokumen dengan masing-masing dokumen telah dikutip dalam karya-karya lain setidaknya tujuh kali.



Gambar 2.4 Indeks h Dosen “S”.

2.6 Penelitian Tentang Regresi Kuantil Terdahulu

Beberapa penelitian yang pernah dilakukan dari berbagai bidang terapan yaitu pendidikan, kesehatan, ekonomi, dan bidang terapan ekologi dan lingkungan dengan menggunakan regresi kuantil adalah sebagai berikut :

1. Wahyudi (2015) menggunakan regresi kuantil untuk memodelkan IPM di pulau jawa. Hasilnya menunjukkan bahwa regresi kuantil memiliki estimasi parameter yang bersifat *robust*, efisien dan konsisten.
2. Navianti (2014) memodelkan tingkat pengangguran terbuka di Indonesia, dan diperoleh hasil bahwa regresi kuantil terbukti dapat mengatasi kasus heteroskedastisitas.
3. Uthami (2013) memodelkan data *Passenger Car Milage* pada kuantil setiap kelipatan 5. Navianti mengatakan semakin tinggi nilai model dikatakan lebih baik karena menyesuaikan dengan sebaran data.
4. Zhao *et al.* (2009) melakukan penelitian mengenai perawatan yang sesuai untuk penderita jantung. Penelitian ini membandingkan hasil antara regresi OLS dengan regresi kuantil. Hasil yang diperoleh adalah pada regresi OLS beberapa asumsi tidak dapat dipenuhi, sedangkan pada regresi kuantil dapat diatasi dengan memodelkan pada kuantil tertentu sehingga sebaran pasien jantung dan perawatan yang sesuai dapat dijelaskan.
5. Zhou dan Portnoy (1996) diketahui bahwa metode *direct* hasil paling efisien untuk taksiran interval yang lebih sempit secara komputasi.

Sedangkan untuk mendapatkan estimator yang konsisten pada taksiran interval yang sempit perlu dilakukan algoritma.

6. Koenker dan D'Orey (1993) meneliti mengenai algoritma simplek pada regresi kuantil menunjukkan bahwa algoritma simplek adalah algoritma yang memiliki proses lambat namun stabil. Algoritma ini sesuai untuk jumlah observasi yang kecil.

2.7 Penelitian Tentang Publikasi Terdahulu

Penelitian mengenai Publikasi yang pernah dilakukan adalah sebagai berikut:

1. Rahmawati (2015) menyelidiki tentang publikasi pada Scopus dan faktor yang mempengaruhi kepemilikan publikasi dosen ITS menggunakan regresi logistik. Hasil dari penelitian menunjukkan bahwa faktor yang berpengaruh adalah usia, jabatan fungsional, pendidikan terakhir, dan tempat pendidikan dengan ketepatan klasifikasi 73,7%.
2. Hapsery (2015) meneliti publikasi dosen di *Google Scholar* menunjukkan bahwa jumlah kutipan dipengaruhi oleh jumlah paper yang ditulis dalam bahasa Inggris, lulusan dan jabatan sebagai guru besar. Dengan model rekursif indeks h dipengaruhi oleh jumlah paper dalam bahasa Inggris, pendidikan, usia, fakultas (FMIPA) dan lulusan.
3. Margaretha dan Saragih (2012) melakukan penelitian pada 3 Provinsi yaitu DKI Jakarta, Jawa Barat dan Jawa Tengah, menunjukkan bahwa masa kerja dan investasi berpengaruh terhadap produktivitas penelitian dosen. Ekspektasi dosen senior dalam melakukan sebuah penelitian berbeda dengan dosen junior.

BAB III

METODOLOGI PENELITIAN

Pada penelitian ini analisis regresi berbasis model rekursif dan regresi kuantil akan diterapkan pada data kepegawaian dosen ITS dan publikasi dosen ITS di Scopus.

3.1 Sumber Data

Sumber data pada penelitian ini berasal dari data primer dan sekunder. Data sekunder yaitu data kepegawaian ITS didapat dari Sulistiyawati (2016). Sedangkan untuk data primer diperoleh dari website scopus di www.scopus.com yang diakses pada bulan Juli 2016. Unit observasi yang digunakan dalam penelitian ini sebanyak 457 Dosen.

Tabel 3.1 Stuktur Data.

Dosen ke- <i>i</i>	Y_1	Y_2	X_1	X_2	...	X_8
1	$Y_{1,1}$	$Y_{2,1}$	$X_{1,1}$	$X_{2,1}$...	$X_{8,1}$
2	$Y_{1,2}$	$Y_{2,2}$	$X_{1,2}$	$X_{2,2}$...	$X_{8,2}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	$Y_{1,n}$	$Y_{2,n}$	$X_{1,n}$	$X_{2,n}$...	$X_{8,n}$

Adapun variabel penelitian yang digunakan terdiri dari variabel dependen dan variabel independen yang ditunjukkan pada Tabel 3.2 dan Tabel 3.3 dimana $i = \text{dosen ke } 1, 2, 3, \dots, n$.

Defisini oprasional

- Sitasi

Sitasi adalah jumlah dokumen yang dikutip oleh peneliti lain, Sitasi yang digunakan adalah sitasi keseluruhan dari dokumen yang telah dipublikasikan oleh dosen tersebut.

- Indeks h

Indeks h menunjukkan bahwa seorang dosen mempublikasikan dokumen sebanyak h, dengan sitasi untuk setiap dokumen tersebut minimal sama dengan h. Apabila sitasi meningkat maka indeks h juga akan meningkat, hal ini menunjukkan bahwa sitasi berpengaruh terhadap indeks h.

- **Usia**
Usia adalah lama waktu hidup yang dihitung sejak dilahirkan hingga waktu usia tersebut dihitung.
- **Lama bekerja**
Lama bekerja seorang dosen dihitung sejak dosen tersebut diterima menjadi CPNS. Apabila dosen tersebut ditugaskan untuk belajar, maka masa studi tersebut tetap dihitung sebagai masa bekerja di ITS.
- **Jumlah dokumen**
Jumlah dokumen adalah banyaknya dokumen yang telah dipublikasikan oleh dosen tersebut, perhitungan dimulai dari dokumen awal hingga dokumen akhir yang dipublikasikan.
- **Jumlah co-authors**
Suatu penelitian umumnya dilakukan secara bersama-sama dengan peneliti lainnya.
- **Jenis kelamin**
Jenis kelamin adalah perbedaan bentuk, sifat dan fungsi biologis antara laki-laki dan perempuan yang membedakan peran dalam penyelenggaraan upaya meneruskan garis keturunan.
- **Jabatan fungsional**
Dosen ITS mempunyai jabatan mulai dari asisten ahli hingga profesor. Setiap dosen akan mengalami kenaikan jabatan fungsional, ketika dapat memenuhi angka kredit yang telah ditetapkan seperti yang disajikan pada Tabel 3.4 berikut.

Tabel 3.4 Rincian Jabatan Fungsional Dosen.

Jabatan	Angka Kredit	Pangkat
Guru Besar	1050	IV/e
	850	IV/d
Lektor Kepala	700	IV/c
	550	IV/b
Lektor	400	IV/a
	300	III/d
	200	III/c
Asisten Ahli	150	III/b
	100	III/a

- Pendidikan terakhir

Pendidikan terakhir adalah pendidikan yang telah ditempuh seorang dosen baik didalam negeri maupun pendidikan diluar negeri. Kategori pendidikan terakhir dosen ITS adalah S2 dan S3, karena dosen ITS diwajibkan minimal mempunyai pendidikan S2. Seiring berjalannya waktu, tidak sedikit para dosen ITS yang mendapatkan tugas belajar pendidikan S3.

- Tempat pendidikan terakhir

Tempat pendidikan bisa dilakukan didalam maupun diluar negeri. Tempat pendidikan yang digunakan pada penelitian ini adalah tempat pendidikan terakhir dosen didalam negeri atau diluar negeri.

Tabel 3.2 Variabel Penelitian

Nama Variabel		Skala
Dependen		
$y_{1,i}$	Sitasi oleh dosen ke-i	Rasio
$y_{2,i}$	Indeks h oleh dosen ke-i	Rasio
Independen		
$X_{1,i}$	Usia	Rasio
$X_{2,i}$	Lama bekerja	Rasio
$X_{3,i}$	Jumlah dokumen	Rasio
$X_{4,i}$	Jumlah <i>co-authors</i>	Rasio
$X_{5,i}$	Jenis kelamin	Nominal
$X_{6,i}$	Jabatan Fungsional	Ordinal
$X_{7,i}$	Pendidikan terakhir	Ordinal
$X_{8,i}$	Tempat pendidikan terakhir	Nominal

Tabel 3.3 Keterangan Koding Variabel Dummy

Nama Variabel	Kategori
Jenis kelamin	1= Perempuan 2= Laki-laki
Jabatan Fungsional	1=Guru Besar 2= Lektor Kepala 3=Lektor 4=Asisten Ahli
Pendidikan terakhir	1= S2 2= S3
Tempat pendidikan terakhir	1= Dalam Negeri 2= Luar Negeri

3.2 Metode Penelitian

Secara umum langkah yang dilakukan dalam penelitian ini yaitu langkah dalam kajian teoritis dan kajian empiris yaitu:

3.2.1 Kajian Teoritis

Kajian estimasi model regresi kuantil dilakukan dengan tahapan sebagai berikut :

1. Mengestimasi parameter model regresi kuantil, yaitu meminimumkan :

$$\min_s \tau \int_{i=1;v \geq 0}^n |v_i| dF_Y(y) + (1-\tau) \int_{i=1;v < 0}^n |v_i| dF_Y(y).$$

2. Mengoptimasi simpleks.
3. Mendapatkan estimator \hat{S} yaitu:

$$\hat{S}(\tau) = \min_{s \in \mathbb{R}^p} \left\{ \tau \sum_{i=1;v \geq 0}^n |y - \mathbf{X}^T(\tau)| + (1-\tau) \sum_{i=1;v < 0}^n |y - \mathbf{X}^T(\tau)| \right\}.$$

4. Membentuk taksiran interval dengan metode *direct* fungsi sparsity

$$P\left(\hat{S}_j(\tau) - t_{(r/2, df)} se(\hat{S}_j(\tau)) \leq s_j(\tau) \leq \hat{S}_j(\tau) + t_{(r/2, df)} se(\hat{S}_j(\tau))\right) = 1 - r.$$

dimana fungsi sparsity yang digunakan yaitu :

$$S(\tau) = f\left(F^{-1}(\tau)\right)^{-1}.$$

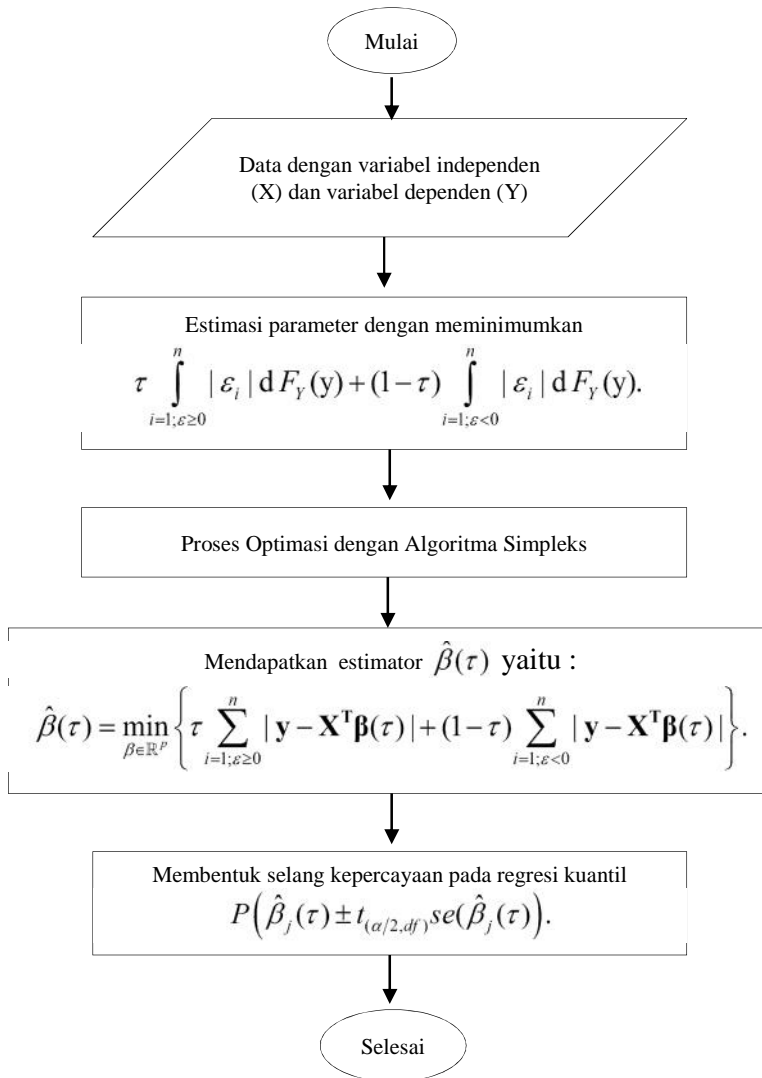
Langkah kajian teoritis disajikan dalam bentuk diagram seperti pada Gambar 3.1. Selain kajian teoritis juga dilakukan studi simulasi dengan membangkitkan *error* homogen dan heterogen dengan ukuran sampel kecil (50), sampel sedang (150), dan sampel besar (300), dengan tiga macam varians pada *error* yaitu 0.25, 1, dan 4. Adapun langkah-langkah yang dilakukan sebagai berikut:

- a. Menentukan besarnya nilai koefisien parameter yang digunakan pada model.
- b. Membangkitkan *error* homogen dan heterogen yang berdistribusi normal.
- c. Membangkitkan variabel independen untuk *error* homogen yang berdistribusi normal dan *error* heterogen yang berdistribusi uniform. Model 1 adalah model untuk *error* homogen dan model 2 untuk *error* heterogen.

Model 1.
$$Y_i = 10 + 5X_{li} + v_i, \quad (3.1)$$

Model 2.
$$Y_i = 10 + 5X_{1i} + \exp^{0.5X_i} v_i, \quad (3.2)$$

- d. Mendapatkan data Y untuk *error* heterogen dan *error* homogen dari perhitungan berdasarkan variabel independen yang telah dibangkitkan dan nilai koefisien parameter yang telah ditentukan.
- e. Memodelkan Y dengan variabel independen menggunakan regresi OLS dan regresi kuantil. Pemodelan dilakukan sebanyak 1000 kali. Dari replikasi yang dilakukan, kemudian akan diperoleh rata-rata nilai parameter untuk regresi OLS dan rata-rata nilai parameter untuk regresi kuantil disetiap nilai kuantil.



Gambar 3.1 Diagram Alir Kajian Teori

3.2.2 Kajian Empiris

Kajian empiris yang dilakukan adalah dengan menerapkan metode regresi kuantil pada kasus riil yaitu publikasi dosen ITS di Scopus dengan langkah sebagai berikut :

1. Mendeskripsikan karakteristik dosen ITS menggunakan statistika deskriptif untuk mengetahui produktifitas dosen ITS yang dilihat berdasarkan publikasi pada Scopus.
2. Memeriksa hubungan antar variabel untuk variabel independen maupun variabel dependen. Tahap pemeriksaan data dilakukan dengan cara mendeteksi adanya kasus heteroskedastisitas. Cara untuk mendeteksi dilakukan dengan membuat *scatter plot* antara variabel independen dengan variabel dependen. Data dikatakan teridentifikasi kasus heteroskedastisitas dengan melihat pola yang dihasilkan membentuk suatu pola tertentu seperti bentuk corong atau lainnya
3. Memodelkan faktor-faktor yang mempengaruhi sitasi dan indeks h dosen ITS menggunakan regresi linier berganda berbasis model rekursif. Langkah analisis yang digunakan sebagai berikut.
 - a. Mengestimasi parameter model regresi linier berganda.
 - b. Melakukan pengujian parameter secara serentak maupun secara parsial.
 - c. Menentukan model regresi linier berganda antara variabel dependen dengan variabel independen yang signifikan.
 - d. Menginterpretasi model regresi linier berganda.
 - e. Mengukur kebaikan model dengan menggunakan koefisien determinasi.
 - f. Melakukan pemeriksaan asumsi pada residual berdasarkan model yang didapat.
4. Memodelkan publikasi dosen ITS menggunakan regresi kuantil berbasis model rekursif pada kuantil 5%, 10%, 25%, 50%, 75%, 90 dan 95% dengan model sebagai berikut :

$$Y_{1i}(\tau) = S_{10}(\tau) + x_{11}(\tau)X_{1i} + x_{12}(\tau)X_{2i} + v_{1i}(\tau)$$

$$Y_{2i}(\tau) = S_{20}(\tau) + S_{21}(\tau)Y_{1i} + x_{21}(\tau)X_{1i} + x_{22}(\tau)X_{2i} + v_{2i}(\tau).$$

a. Mengestimasi parameter model regresi kuantil, yaitu meminimumkan :

$$\min_{S \in \mathbb{R}^p} \tau \int_{i=1; v \geq 0}^n |v_i| dF_Y(y) + (1-\tau) \int_{i=1; v < 0}^n |v_i| dF_Y(y).$$

b. Mengoptimasi simpleks.

c. Mendapatkan estimator $\hat{S}(\tau)$ untuk kuantil 5%, 10%, 25%, 50%, 75%, 90% dan 95%.

$$\hat{S}(\tau) = \min_{S \in \mathbb{R}^p} \left\{ \tau \sum_{i=1; v \geq 0}^n |y - \mathbf{X}^T(\tau)| + (1-\tau) \sum_{i=1; v < 0}^n |y - \mathbf{X}^T(\tau)| \right\}$$

d. Membentuk selang kepercayaan regresi kuantil

$$P\left(\hat{S}_j(\tau) - t_{(r/2, df)} se(\hat{S}_j(\tau)) \leq S_j(\tau) \leq \hat{S}_j(\tau) + t_{(r/2, df)} se(\hat{S}_j(\tau))\right) = 1 - r.$$

e. Melakukan uji signifikansi parameter dengan statistik uji:

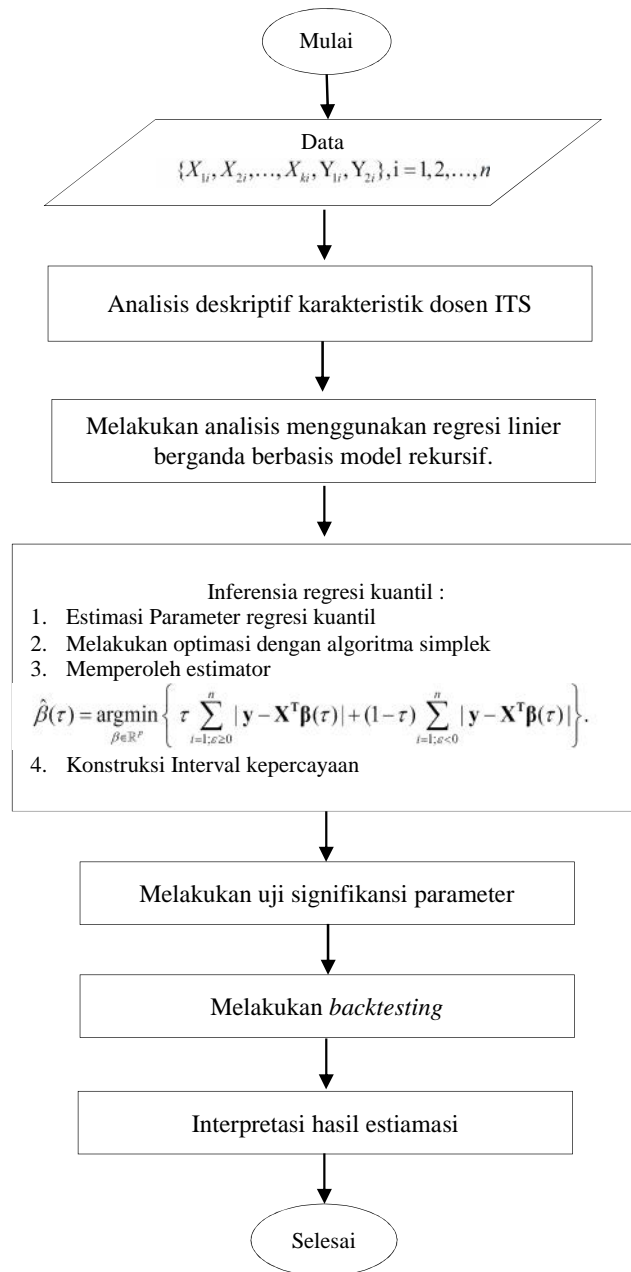
$$t = \frac{\hat{S}_j(\tau)}{se(\hat{S}_j(\tau))}.$$

f. Melakukan uji kesesuaian model menggunakan *backtesting procedure* dengan melihat akurasi dari model. Model dikatakan baik apabila

$$\tau^* = \tau \text{ dan } \tau^* = P((y | X) \leq (\hat{y}_t | X)).$$

g. Menginterpretasikan model yang diperoleh dari hasil analisis.

Diagram alir penelitian ini dapat dilihat pada Gambar 3.2 berikut:



Gambar 3.2 Diagram Alir Pemodelan Publikasi Dosen ITS

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bagian ini akan dilakukan analisis dan pembahasan dari estimasi parameter, studi simulasi serta aplikasi pada data publikasi dosen ITS di Scopus. Pemodelan terhadap publikasi dosen ITS di Scopus akan dilakukan dengan menggunakan regresi kuantil berbasis model rekursif. Secara lengkap hasil analisis dan pembahasan dapat dijelaskan sebagai berikut.

4.1 Kajian Teoritis

Terdapat dua kajian teoritis yang akan dibahas, yang pertama adalah taksiran titik regresi kuantil dan selanjutnya adalah taksiran interval metode *direct* fungsi sparsity.

4.1.1 Regresi Kuantil

Taksiran titik pada regresi kuantil bermula dari suatu himpunan data berpasangan $\{X_{1i}, X_{2i}, \dots, X_{pi}, Y_i\}, i = 1, 2, \dots, n$ yang berdistribusi independen dan identik dengan kuantil $\dagger \in (0,1)$. Data tersebut memiliki fungsi distribusi bersyarat :

$$F(y|X_i) = P(Y \leq y|X_i), \quad (4.1)$$

dan fungsi invers dari

$$\begin{aligned} F^{-1}(\dagger) &= \inf \{y : F(y) \geq \dagger\} \\ Q(\dagger) &= \inf \{y : F(y) \geq \dagger\} \\ Q_{Y_i}(\dagger | X_{1i}, X_{2i}, \dots, X_{pi}) &= \inf \{y : F(y) \geq \dagger\}, \end{aligned} \quad (4.2)$$

yang merupakan kuantil ke- \dagger dari variabel dependen (Y) (Koenker, 2005: 59). Jadi, model fungsi kuantil bersyarat yaitu :

$$Q_{Y_i}(\dagger | X_{1i}, X_{2i}, \dots, X_{pi}) = Q_Y(\dagger | \mathbf{X}) = \mathbf{X}^T(\dagger). \quad (4.3)$$

Definisi fungsi distribusi peluang bersyarat:

$$Q_Y(\dagger | \mathbf{X}) = F_{Y_i}^{-1}(y|X_i) = \inf \{y : F(y|X) \geq \dagger\}. \quad (4.4)$$

Persamaan (4.5) adalah persamaan umum regresi kuantil yaitu:

$$Y_i = S_0(\dagger) + S_1(\dagger)X_{1i} + \dots + S_p(\dagger)X_{pi} + v_i(\dagger), \quad (4.5)$$

dengan $i = 1, 2, \dots, n$.

Apabila persamaan (4.5) dijabarkan, maka diperoleh persamaan sebagai berikut:

$$\begin{aligned} i = 1 \quad Y_1 &= S_0(\dagger) + S_1(\dagger)X_{11} + S_2(\dagger)X_{21} + \dots + S_p(\dagger)X_{p1} + v_1(\dagger) \\ i = 2 \quad Y_2 &= S_0(\dagger) + S_1(\dagger)X_{12} + S_2(\dagger)X_{22} + \dots + S_p(\dagger)X_{p2} + v_2(\dagger) \\ &\vdots \\ i = n \quad Y_n &= S_0(\dagger) + S_1(\dagger)X_{1n} + S_2(\dagger)X_{2n} + \dots + S_p(\dagger)X_{pn} + v_n(\dagger). \end{aligned} \quad (4.6)$$

Apabila model regresi kuantil pada persamaan (4.6) disajikan dalam bentuk matriks yaitu :

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix} \begin{bmatrix} S_0(\dagger) \\ S_1(\dagger) \\ \vdots \\ S_p(\dagger) \end{bmatrix} + \begin{bmatrix} v_1(\dagger) \\ v_2(\dagger) \\ \vdots \\ v_n(\dagger) \end{bmatrix}. \quad (4.7)$$

Selanjutnya persamaan (4.7) dapat ditulis dalam bentuk model linier berikut:

$$\mathbf{y} = \mathbf{X}(\dagger) + \mathbf{v}. \quad (4.8)$$

dimana :

\mathbf{y} : vektor kolom berukuran $n \times 1$ dari variabel dependen (Y).

\mathbf{X} : matrik berukuran $n \times (p+1)$ dengan n observasi pada p variabel X_j

dimana $j = 1, 2, \dots, p$.

: vektor kolom berukuran $(p+1) \times 1$ parameter s_j dimana $j = 1, 2, \dots, p$.

: vektor kolom berukuran $n \times 1$ residual v_i .

Regresi kuantil merupakan salah satu metode regresi yang diperoleh dari metode regresi median yang diperumum. Regresi kuantil dapat digunakan dalam memecahkan permasalahan dalam rataan. Basset dan koenker (1978) dalam makalahnya membahas mengenai regresi linier, yang kemudian berkembang menjadi regresi median yang dinyatakan :

$$\min_{\mathbf{s} \in \mathbb{R}^p} \sum_{i=1}^n |\mathbf{y}_i - \mathbf{X}_i^T \boldsymbol{\beta}|. \quad (4.9)$$

Secara umum dispesifikasikan dalam fungsi kuantil bersyarat ke- τ dengan mempertimbangkan estimasi bagi $(\hat{\mathbf{s}}(\tau))$, sehingga diperoleh ide untuk permasalahan tersebut yang dinyatakan sebagai berikut:

$$\min_{\mathbf{s} \in \mathbb{R}^p} \sum_{i=1}^n \dots(\tau)(\mathbf{y}_i - Q(\tau)(y_i | X_i)), \quad (4.10)$$

dengan

$\dots(\tau)(\cdot)$: *loss function* yang asimetrik.

$Q(\tau)(y | X)$: fungsi kuantil ke- τ dari Y dengan syarat X .

Berdasarkan asumsi integrabilitas, fungsi kuantil bersyarat dapat diperoleh dari meminimumkan:

$$Q(\tau)(y | X) = \underset{q(X)}{\operatorname{argmin}} E[\dots(\tau)(\mathbf{y} - q(\mathbf{X}))]. \quad (4.11)$$

Regresi kuantil dapat terpenuhi dengan mengganti suatu model linier $q(X)$ pada persamaan (4.11), dan diperoleh hasil

$$\hat{\mathbf{s}}(\tau) = \underset{\mathbf{s} \in \mathbb{R}^p}{\operatorname{argmin}} E[\dots(\tau)(\mathbf{y} - \mathbf{X} \boldsymbol{\beta})]. \quad (4.12)$$

Dalam regresi kuantil ke- τ dari F_Y dapat diperoleh dari persamaan (4.13) berikut:

$$\frac{\partial}{\partial \hat{y}} E[\dots(\tau)] | Y = \hat{y} = \frac{\partial}{\partial \hat{y}} \int_{-\infty}^{\hat{y}} (1-\tau) |y - \hat{y}| f(y) dy + \int_{\hat{y}}^{\infty} \tau |y - \hat{y}| f(y) dy. \quad (4.13)$$

Selanjutnya pada persamaan (4.13) diminimumkan menjadi nol

$$\frac{\partial}{\partial \hat{y}} E[\dots(\tau) | Y = \hat{y}] = 0$$

$$\frac{\partial}{\partial \hat{y}} \int_{-\infty}^{\hat{y}} (1-\tau) |y - \hat{y}| f(y) dy + \int_{\hat{y}}^{\infty} \tau |y - \hat{y}| f(y) dy = 0$$

$$\frac{\partial}{\partial \hat{y}} \left[(1-\tau) \int_{-\infty}^{\hat{y}} (y - \hat{y}) f(y) dy + \tau \int_{\hat{y}}^{\infty} (y - \hat{y}) f(y) dy \right] = 0$$

$$\begin{aligned}
& (1-\tau) \left[(y-\hat{y})f(y) \Big|_{-\infty}^{\hat{y}} + \int_{-\infty}^{\hat{y}} \frac{\partial}{\partial \hat{y}} (y-\hat{y})f(y) dy \right] + \\
& \tau \left[(y-\hat{y})f(y) \Big|_{\hat{y}}^{-\infty} + \int_{\hat{y}}^{\infty} \frac{\partial}{\partial \hat{y}} (y-\hat{y})f(y) dy \right] = 0 \\
& (1-\tau) \left[(y-\hat{y})f(y) \Big|_{y=\hat{y}} + \int_{-\infty}^{\hat{y}} f(y) dy \right] + \\
& \tau \left[(y-\hat{y})f(y) \Big|_{y=\hat{y}} + \int_{\hat{y}}^{\infty} f(y) dy \right] = 0 \\
& (1-\tau)[0 + F_Y(\hat{y})] + \tau[0 - (1 - F_Y(\hat{y}))] = 0 \\
& (1-\tau)F_Y(\hat{y}) - \tau(1 - F_Y(\hat{y})) = 0 \\
& (1-\tau)F_Y(\hat{y}) - \tau + \tau F_Y(\hat{y}) = 0 \\
& F_Y(\hat{y}) - \tau = 0
\end{aligned}$$

dan diperoleh

$$F_Y(\hat{y}) = \tau \quad (4.14)$$

(Davino *et al*, 2014: 2-8).

Apabila dalam estimasi OLS dari suatu model linier terhadap Y diperoleh dengan meminimumkan jumlah kuadrat residual. Sedangkan dalam estimasi regresi kuantil dari suatu model linier terhadap Y diperoleh dengan meminimumkan nilai *loss function* yang tidak simetris. Meminimumkan nilai tersebut yaitu dengan meminimumkan nilai estimasi $\dots(\tau)(v)$.

$$\hat{S}(\tau) = \min_{s \in \mathbb{R}^p} \left\{ \tau \sum_{i=1; v \geq 0}^n |\mathbf{y} - \mathbf{X}^T(\tau)| + (1-\tau) \sum_{i=1; v < 0}^n |\mathbf{y} - \mathbf{X}^T(\tau)| \right\}, \quad (4.15)$$

atau

$$\hat{S}(\tau) = \min_{s \in \mathbb{R}^p} \sum_{i=1}^n \dots(\tau)(v_i). \quad (4.16)$$

dimana:

$$\dots(\ddagger)(v_i) = \begin{cases} \ddagger v_i & , v_i \geq 0 \\ (\ddagger - 1)v_i & , v_i < 0. \end{cases}$$

dengan ;

$\hat{S}(\ddagger)$: Penduga parameter.

\ddagger : Indeks kuantil dengan $\ddagger \in (0,1)$.

$\dots(\ddagger)(v_i)$: *Loss function*.

v_i : Residual dari estimasi parameter.

Selanjutnya akan dibuktikan bahwa *loss function* berbentuk asimetris dengan penjelasan berikut :

$$\begin{aligned} \dots(\ddagger) &= [\ddagger I(v \geq 0) + (1 - \ddagger) I((v < 0))] |v|. \\ &= [\ddagger - I(v < 0)] v. \end{aligned} \quad (4.17)$$

dengan

v : residual dari estimasi parameter.

$I(v)$: fungsi indikator yang telah didenifisikan.

dimana

$$I(v \geq 0) = \begin{cases} 1 & , v \geq 0 \\ 0 & , v < 0. \end{cases} \quad (4.18)$$

Sehingga dapat dibuktikan

a. untuk $v \geq 0$

$$\begin{aligned} \dots(\ddagger) &= [\ddagger I(v \geq 0) + (1 - \ddagger) I(v < 0)] |v| \\ &= [\ddagger I(v \geq 0) + (1 - \ddagger) I(v < 0)] v \\ &= [\ddagger 1 + (1 - \ddagger) I(v < 0)] v \\ &= [\ddagger + I(v < 0) - \ddagger I(v < 0)] v \\ &= [\ddagger + (1 - I(v \geq 0)) - \ddagger (1 - I(v \geq 0))] v \\ &= [\ddagger + (1 - 1) - \ddagger (1 - 1)] v = \ddagger v. \end{aligned}$$

atau

$$\begin{aligned}
\dots(\dagger) &= [\dagger - I(v < 0)]v \\
&= [\dagger - (1 - I(v \geq 0))]v \\
&= [\dagger - (1 - 1)]v = \dagger v.
\end{aligned}$$

b. untuk $v < 0$

$$\begin{aligned}
\dots(\dagger) &= [\dagger I(v \geq 0) + (1 - \dagger)I(v < 0)]|v| \\
&= [\dagger I(v \geq 0) + (1 - \dagger)I(v < 0)](-v) \\
&= [\dagger 0 + (1 - \dagger)I(v < 0)](-v) \\
&= [(\dagger - 1)I(v < 0)]v \\
&= [(\dagger - 1)(1 - I(v < 0))]v \\
&= [(\dagger - 1)(1 - 0)]v = (\dagger - 1)v.
\end{aligned}$$

atau

$$\begin{aligned}
\dots(\dagger) &= [\dagger - I(v < 0)]v \\
&= [\dagger - (1 - I(v \geq 0))]v \\
&= [\dagger - (1 - 0)]v \\
&= (\dagger - 1)v.
\end{aligned}$$

Sehingga menjadi

$$\begin{aligned}
\dots(\dagger) &= [\dagger - I(v \geq 0) + (1 - \dagger)I(v < 0)]|v| \\
&= [\dagger - I(v < 0)]v, \forall v
\end{aligned}$$

Solusi dari persamaan (4.16) tidak dapat diperoleh secara analitik, tetapi secara numerik. Metode numerik yang digunakan adalah algoritma simpleks yang dikembangkan oleh Barrodale dan Robert pada tahun 1974. Metode simpleks merupakan salah satu metode untuk menentukan kombinasi optimal dari tiga variabel atau lebih. Algoritma ini memberikan solusi permasalahan program linier yang melibatkan beberapa variabel keputusan dengan bantuan komputasi. Bentuk umum formulasi metode primal dari permasalahan regresi kuantil linier menggunakan kuantil $\dagger = 0.05$ yaitu:

$$\min_{\mathbf{s} \in \mathbb{R}^p} \sum_{i=1}^n \dots(\dagger)(\mathbf{y} - \mathbf{X}^T \mathbf{s}). \quad (4.19)$$

Diberikan $[x]_+$, non negatif bagian dari x , yaitu $s_1 = [\mathbf{y} - \mathbf{X}]_+$ dan $s_2 = [\mathbf{X} - \mathbf{y}]_+$ maka masalah linier dapat ditulis

$$\min_{\mathbf{s} \in \mathbb{R}^p} \{1^T \mathbf{s}_1 + 1^T \mathbf{s}_2 \mid \mathbf{y} = \mathbf{X} \mathbf{s}_1 - \mathbf{X} \mathbf{s}_2, (\mathbf{s}_1, \mathbf{s}_2) \in \mathbb{R}_+^n\},$$

selanjutnya diberikan

$$\mathbf{B} = [\mathbf{X} \quad -\mathbf{X} \quad -\mathbf{I}],$$

dan

$$\mathbf{c} = \begin{bmatrix} [] \\ [] \\ [\mathbf{y} - \mathbf{X}] \\ [\mathbf{X} - \mathbf{y}] \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 0_{[p]} \\ 0_{[p]} \\ 1_{[n]} \\ 1_{[n]} \end{bmatrix}.$$

seperti reformulasi dari permasalahan pemrograman linier standar, maka formulasinya dapat ditulis $\min_{\mathbf{z}} \mathbf{c}^T \mathbf{z}$ dimana untuk $\mathbf{B}^T \mathbf{z} \leq \mathbf{d}$. Oleh karena itu parameter dual dapat ditulis $\max_{\mathbf{t}} \mathbf{t}^T \mathbf{d}$ dimana untuk $\mathbf{B} \mathbf{t} = \mathbf{c}, \mathbf{t} \geq 0$.

Mengingat hasil utama program linier yaitu solusi dari masalah minimasi tersebut diselesaikan dengan metode simpleks. Secara sederhana masalah diatas dapat dirumuskan seperti pada persamaan (4.20) berikut:

$$\max_{\mathbf{z}} \left\{ \mathbf{y}^T \mathbf{z} \mid \mathbf{X}^T \mathbf{z} = \mathbf{0}, \mathbf{z} \in [-1, +1]^n \right\}, \quad (4.20)$$

untuk $\mathbf{X}^T \mathbf{z} = 0$ yang ditransformasikan dengan cara dikalikan dengan $\frac{1}{2}$, dan menambah $\frac{1}{2} \mathbf{X}^T \mathbf{1}$ sehingga menjadi $\frac{1}{2} \mathbf{X}^T \mathbf{z} + \frac{1}{2} \mathbf{X}^T \mathbf{1} = \frac{1}{2} \mathbf{X}^T \mathbf{1}$. Hasil dari transformasi dapat ditulis seperti pada persamaan (4.21) berikut:

$$\mathbf{X}^T \left(\frac{1}{2} \mathbf{z} + \frac{1}{2} \mathbf{1} \right) = \frac{1}{2} \mathbf{X}^T \mathbf{1}. \quad (4.21)$$

Jika $\frac{1}{2} \mathbf{z} + \frac{1}{2} \mathbf{1} = \mathbf{y}$ dan $\frac{1}{2} \mathbf{X}^T \mathbf{1} = \mathbf{b}$, maka persamaan dapat ditulis $\mathbf{X}^T \mathbf{y} = \mathbf{b}$.

Nilai $1/2$ pada persamaan (4.21) merupakan kunci utama generalisasi untuk kondisi kuantil yang lain. Bahkan masalah minimalisasi untuk median

bersyarat dapat dipakai untuk kuantil bersyarat pada ke- (\dagger) sehingga diperoleh hasil seperti pada persamaan (4.22) berikut:

$$\min_{\mathbf{z}} \sum_{i=1}^n \dots \dagger (y_i - x_i^T \mathbf{z}(\dagger)). \quad (4.22)$$

Dengan cara yang sama diperoleh formulasi dari persamaan dual

$$\max_{\mathbf{z}} \left\{ \mathbf{y}^T \mathbf{z} \mid \mathbf{X}^T \mathbf{z} = (1 - \dagger) \mathbf{X}^T \mathbf{1}, \mathbf{z} \in [0, 1]^n \right\},$$

dimana $(1 - \dagger)$ memiliki peran yang sama seperti $\frac{1}{2}$ pada rumus di median (Davino *et al.*, 2014: 23-31).

4.1.4 Taksiran Interval *Direct* Fungsi Sparsity

Terdapat tiga metode untuk mendapatkan taksiran interval pada regresi kuantil, yaitu *direct* fungsi sparsity, rank-score, re-sampling. Pada penelitian ini estimasi yang digunakan adalah *direct* fungsi sparsity (Chen dan Wei, 2005:409-411). Kelebihan metode ini adalah dapat menjelaskan penyebaran data dengan menggunakan ukuran kuantil sesuai sebaran data. Selain itu, metode ini dapat menghasilkan hasil paling efisien untuk taksiran interval yang lebih sempit secara komputasi (Zhou dan Portnoy, 1996).

Dengan menggeneralisirkan model regresi linier dua dan tiga variabel, model regresi variabel dependen (Y) dengan p variabel independen X_1, X_2, \dots, X_p dapat ditulis :

$$Y_i = S_0 + S_1 X_{1i} + S_2 X_{2i} + \dots + S_p X_{pi} + v_i, \quad (4.23)$$

dimana

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{p1} \\ 1 & X_{12} & X_{22} & \dots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{pn} \end{bmatrix}, \quad = \begin{bmatrix} S_0 \\ S_1 \\ \vdots \\ S_p \end{bmatrix}, \quad = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}.$$

Dari persamaan (4.23) kemudian mengacu pada persamaan umum regresi kuantil seperti pada persamaan (4.5) berikut:

$$Y_i = S_0(\dagger) + S_1(\dagger) X_{1i} + \dots + S_p(\dagger) X_{pi} + v_i(\dagger),$$

dengan $i = 1, 2, \dots, n$.

Residual berdistribusi IIDN dengan cdf f pada kuantil ke- (\dagger) , diperoleh $f(F^{-1}(\dagger) = \mathbf{X}^T > 0)$. Bentuk kuadrat dari matriks \mathbf{X} diasumsikan konvergen menuju definit positif matriks \mathbf{D} sebagai berikut:

$$\mathbf{D} = \lim_{n \rightarrow \infty} \frac{1}{n} (\mathbf{X}^T \mathbf{X})^{-1}. \quad (4.24)$$

Adapun matriks \mathbf{D} pada persamaan (4.24) diperoleh dari langkah-langkah berikut:

$$\begin{aligned} \mathbf{D} &= \lim_{n \rightarrow \infty} \frac{1}{n} (\mathbf{X}^T \mathbf{X})^{-1} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & \dots & X_{pn} \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{21} & \dots & X_{p1} \\ 1 & X_{12} & X_{22} & \dots & X_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & X_{2n} & \dots & X_{pn} \end{bmatrix} \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \begin{bmatrix} n & \sum X_{li} & \dots & \sum X_{pi} \\ \sum X_{li} & \sum X_{li}^2 & \dots & \sum X_{pi} X_{li} \\ \vdots & \vdots & \ddots & \vdots \\ \sum X_{pi} & \sum X_{pi} X_{li} & \dots & \frac{\sum X_{pi}^2}{n} \end{bmatrix} \\ &= \lim_{n \rightarrow \infty} \begin{bmatrix} 1 & \frac{\sum X_{li}}{n} & \dots & \frac{\sum X_{pi}}{n} \\ \frac{\sum X_{li}}{n} & \frac{\sum X_{li}^2}{n} & \dots & \frac{\sum X_{pi} X_{li}}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sum X_{pi}}{n} & \frac{\sum X_{pi} X_{li}}{n} & \dots & \frac{\sum X_{pi}^2}{n} \end{bmatrix}. \end{aligned}$$

Langkah selanjutnya adalah mendapatkan estimator kuantil regresi dari vektor koefisien $\hat{\mathbf{s}}(\dagger) = [s_0(\dagger) \quad s_1(\dagger) \quad \dots \quad s_p(\dagger)]$. Sehingga diperoleh:

$$\sqrt{n} [\hat{\mathbf{s}}(\dagger) - \mathbf{s}(\dagger)] \rightarrow N(0, w^2(\dagger) \mathbf{D}^{-1}). \quad (4.25)$$

Sebelumnya telah didapatkan estimator dari $\hat{\mathbf{s}}(\dagger)$ pada kajian estimasi parameter regresi kuantil yaitu persamaan (4.16) sebagai berikut:

$$\hat{S}(\dagger) = \min_{S \in \mathbb{R}^p} \sum_{i=1}^n \dots(\dagger)(v_i).$$

Dimana $\check{S}^2(\dagger)$ untuk setiap kuantil $ke-(\dagger)$ didefinisikan sebagai berikut (Davino *et al*, 2014: 131-134):

$$\check{S}^2(\dagger) = \frac{\dagger(1-\dagger)}{f(F^{-1}(\dagger))^2}. \quad (4.26)$$

Untuk F adalah fungsi distribusi kumulatif dan $f = F'$ adalah fungsi densitas, maka fungsi sparsity dinotasikan sebagai berikut:

$$S(\dagger) = f(F^{-1}(\dagger))^{-1}. \quad (4.27)$$

Pada setiap \dagger yang sudah ditentukan, konfidensi interval dengan formula seperti pada persamaan (4.28) berikut:

$$P(\hat{S}_j(\dagger) - t_{(r/2, df)} se(\hat{S}_j(\dagger)) \leq S_j(\dagger) \leq \hat{S}_j(\dagger) + t_{(r/2, df)} se(\hat{S}_j(\dagger))) = 1 - r. \quad (4.28)$$

Nilai $se(\hat{S}_j(\dagger))$ adalah nilai diagonal yang diperoleh dari perhitungan matrik $\check{S}^2(\dagger)\mathbf{D}^{-1}$.

4.2 Studi Simulasi

Studi simulasi dalam penelitian ini menggunakan model regresi linier dengan satu variabel independen. Simulasi dilakukan replikasi sebanyak 1000 dengan membangkitkan dua jenis *error* yaitu *error* homogen dimana ukuran sampel sebanyak 300 pengamatan, dan heterogen dengan ukuran sampel kecil (50), sampel sedang (150), dan sampel besar (300), dengan tiga macam varians pada *error*. Bentuk persamaan umum data simulasi adalah sebagai berikut:

$$Y_i = S_0 + S_1 X_{1i} + v_i, \quad (4.29)$$

dimana

X : variabel independen random.

v_i : *error* random $v_i \sim N(0, \dagger^2)$, $i = 1, 2, \dots, n$.

S : parameter regresi.

4.2.1 Studi Simulasi *Error* Homogen

Studi simulasi pertama adalah membangkitkan *error* homogen untuk variabel independen dengan kriteria residual berdistribusi normal, rata-rata sama dengan nol dan varians sama dengan dua puluh lima. Model ini disebut dengan model regresi parametrik linear sederhana untuk jenis *error* yang homogen. Model yang digunakan mengacu pada persamaan (4.29) dimana $S_0 = 10$ dan $S_1 = 5$ sehingga persamaan menjadi:

$$Y_i = 10 + 5X_{ii} + v_i, \quad (4.30)$$

dimana

X : variabel independen random $X_i \sim N(0,100)$, $i = 1, 2, \dots, n$.

v_i : *error* random $v_i \sim N(0,25)$, $i = 1, 2, \dots, n$.

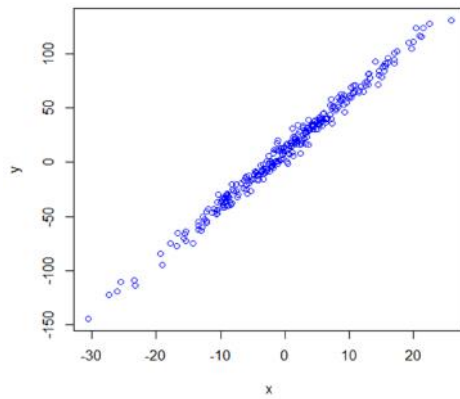
S : parameter regresi.

Berdasarkan hasil simulasi untuk model regresi linier dengan jumlah data sebanyak 300, diperoleh hasil seperti pada Tabel 4.1

Tabel 4.1 Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Linier Studi Simulasi *Error* Homogen

Sampel	\hat{S}_p	Nilai	Batas Bawah	Batas Atas	Akurasi
300	\hat{S}_0	9,988	9,788	10,179	100%
	\hat{S}_1	4,998	4,978	5,018	100%

Replikasi pada studi simulasi *error* homogen dilakukan sebanyak 1000, kemudian didapatkan nilai rata-rata dari seluruh nilai parameter sebesar $\bar{\hat{S}}_0 = 9,987$ dan $\bar{\hat{S}}_1 = 4,998$. Nilai tersebut berada diantara batas bawah dan batas atas, menunjukkan kebenaran dari hasil studi simulasi. Plot antara variabel dependen dan independen untuk replikasi ke-1000 yang diperoleh dari hasil simulasi *error* homogen sesuai model pada persamaan (4.30), dapat dilihat pada Gambar 4.1 dan estimasi parameter pada Tabel 4.2.



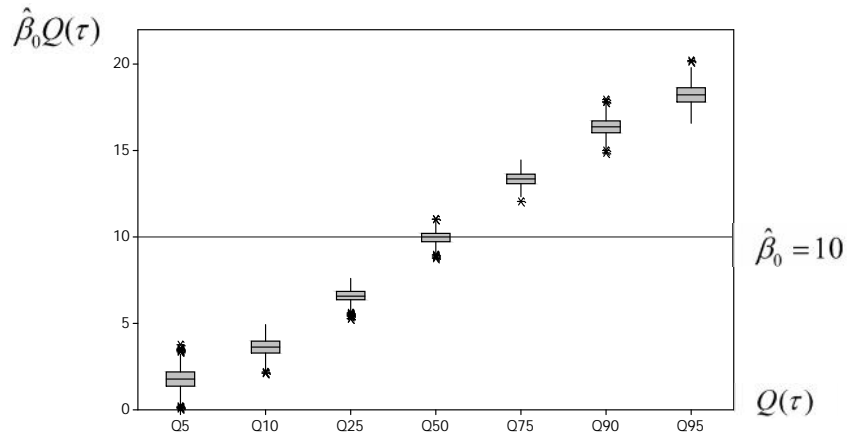
Gambar 4.1 Scatter Plot Replikasi ke-1000 Hasil Studi Simulasi *Error* Homogen

Tabel 4.2 Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Kuantil
Studi Simulasi *Error* Homogen

Sampel	‡	\widehat{S}_p	Nilai	Batas Bawah	Batas Atas	Akurasi
5%		\widehat{S}_0	1,782	1,392	4,972	100%
		\widehat{S}_1	4,997	2,177	5,023	100%
10%		\widehat{S}_0	3,598	4,956	13,069	100%
		\widehat{S}_1	4,997	5,04	13,634	100%
25%		\widehat{S}_0	6,621	3,259	4,972	100%
		\widehat{S}_1	4,997	3,945	5,027	100%
50%		\widehat{S}_0	9,992	4,964	16,024	100%
		\widehat{S}_1	4,998	5,031	16,712	100%
75%		\widehat{S}_0	13,35	6,371	4,965	100%
		\widehat{S}_1	4,999	6,878	5,033	100%
90%		\widehat{S}_0	16,37	4,97	17,789	100%
		\widehat{S}_1	4,999	5,023	18,613	100%
95%		\widehat{S}_p	18,19	9,746	4,958	100%
		\widehat{S}_0	4,999	10,24	5,040	100%

Identifikasi yang dilakukan untuk mengetahui apakah sebaran data teridentifikasi heterogen atau tidak seragam. Berdasarkan Gambar 4.1 diatas menunjukkan bahwa varians *error* konstan, yang artinya sebaran data dari hasil studi simulasi adalah homogen. Langkah selanjutnya yaitu memodelkan data dengan menggunakan regresi kuantil seperti yang disajikan pada Tabel 4.2.

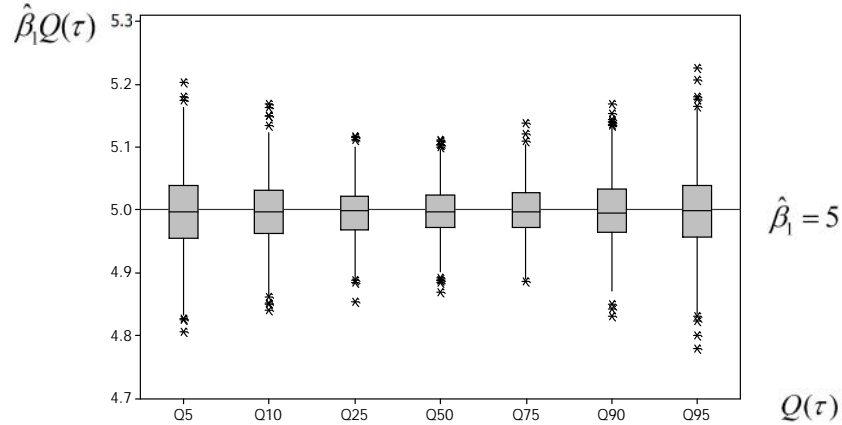
Tabel 4.2 menunjukkan nilai rata-rata dari estimasi parameter regresi kuantil dengan replikasi sebanyak 1000 kali. Terlihat bahwa seluruh nilai estimasi berada diantara batas atas dan batas bawah dengan tingkat akurasi 100%. Hal ini menunjukkan bahwa hasil simulasi sesuai dengan hasil yang diharapkan. Selain itu, dapat diketahui juga nilai *intercept* semakin lama semakin besar sejalan dengan kenaikan nilai kuantil. Pada saat kuanti 50%, nilai rata-rata *intercept* pada regresi kuantil identik atau sama dengan nilai *intercept* pada regresi linier. Seperti yang ditunjukkan pada Gambar 4.2.



Gambar 4.2 *Boxplot* \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi *Error* Homogen.

Berbeda dengan hasil estimasi untuk \hat{S}_1 pada Gambar 4.3. Dapat dilihat bahwa nilai *slope* sejajar untuk semua kuantil ke- \dagger . Informasi lain yaitu untuk kuantil 50% identik dengan *slope* pada regresi linier. Berdasarkan Tabel 4.2 diketahui nilai *slope* rata-rata dari 1000 replikasi untuk masing-masing kuantil tidak jauh berbeda. Oleh karena itu, apabila digambarkan seperti pada Gambar 4.3 terlihat sama. Artinya, apabila *error* homogen,

untuk hasil estimasi *slope* masing-masing kuantil akan sama dengan estimasi *slope* pada regresi linier.



Gambar 4.3 Boxplot Regresi Kuantil \hat{S}_1 Studi Simulasi *Error* Homogen

Model pada kuantil ke 5%, 50%, dan 95% dapat dituliskan sebagai berikut:

Model untuk kuantil 5% : $Q_{0,05}(y | X) = 1,782 + 4,997 X_i$.

Model untuk kuantil 50% : $Q_{0,50}(y | X) = 9,992 + 4,998 X_i$.

Model untuk kuantil 95% : $Q_{0,95}(y | X) = 18,19 + 4,999 X_i$.

4.2.2 Studi Simulasi *Error* Heterogen

Studi simulasi kedua adalah membangkitkan data untuk sebaran data yang tidak seragam atau heterogen. Variabel independen berdistribusi uniform, sedangkan nilai residual diperoleh dengan membangkitkan data berdistribusi normal dengan rata-rata nol dan varians satu.

$$Y_i = 10 + 5X_{ii} + \exp^{0,5X_i} v_i, \quad (4.31)$$

dimana

X : variabel independen random $X_i \sim U(0,5)$, $i = 1, 2, \dots, n$

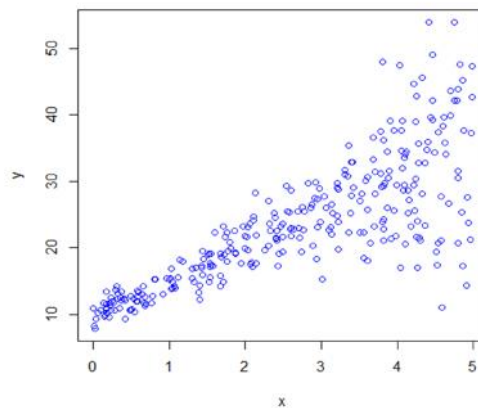
v_i : *error* random $v_i \sim N(0,1)$, $i = 1, 2, \dots, n$

s : parameter regresi dengan $S_0 = 10$ dan $S_1 = 5$.

Tabel 4.3 Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Linier
Studi Simulasi *Error* Heterogen

Sampel	\hat{S}_p	Nilai	Batas Bawah	Batas Atas	Akurasi
300	\hat{S}_0	10,011	9,703	10,30	100%
	\hat{S}_1	4,994	4,824	5,18	100%

Simulasi pada kasus heteroskedastisitas ini dilakukan pengulangan sebanyak 1000 kali. Untuk memperoleh nilai rata-rata dari parameter, dilakukan dengan menjumlahkan seluruh nilai parameter kemudian membagi nilai parameter tersebut dengan banyaknya replikasi. Sehingga didapatkan nilai rata-rata sebesar $S_0 = 10,011$ dan $S_1 = 4,994$. Identifikasi sebaran data dari hasil simulasi dapat dilihat pada Gambar 4.4.



Gambar 4.4 Scatter Plot Replikasi ke-1000 Hasil Studi Simulasi *Error* Heterogen

Berdasarkan Gambar 4.4 menunjukkan bahwa varians *error* tidak konstan. Hal ini dapat diartikan sebaran data dari hasil simulasi adalah heterogen atau sebaran data tidak seragam. Tabel 4.4 menunjukkan nilai estimasi parameter untuk regresi kuantil. Berdasarkan hasil yang diperoleh, diketahui nilai parameter *intercept* dan *slope* bertambah seiring dengan perubahan nilai kuantil. Hal ini menunjukkan bahwa regresi kuantil dapat memodelkan data sesuai dengan sebaran data, seperti yang terlihat pada Gambar 4.5 dan 4.6. Ketepatan dari hasil studi simulasi dapat diketahui apabila rata-rata nilai estimasi baik dari *slope* maupun *intercept* berada diantara batas bawah dan

batas atas. Model regresi kuantil untuk studi simulasi *error* heterogen dapat dituliskan sebagai berikut:

Model untuk kuantil 5% : $Q_{0,05}(y | \mathbf{x}) = 9,443 + 2,245X_i$.

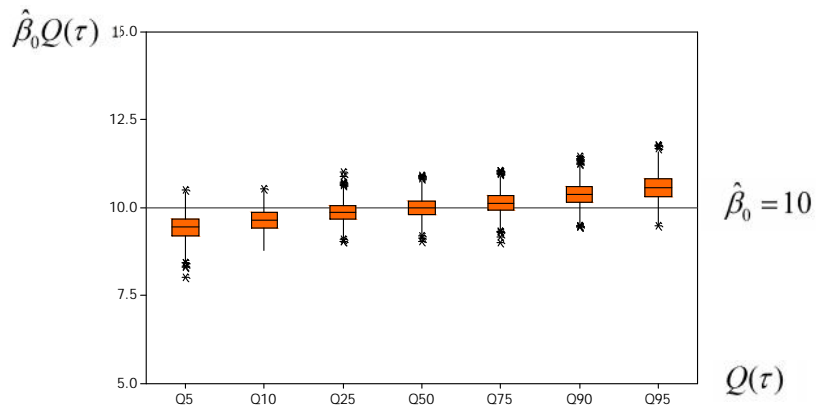
Model untuk kuantil 50% : $Q_{0,50}(y | \mathbf{x}) = 10,004 + 4,997X_i$.

Model untuk kuantil 95% : $Q_{0,95}(y | \mathbf{x}) = 10,576 + 7,754X_i$.

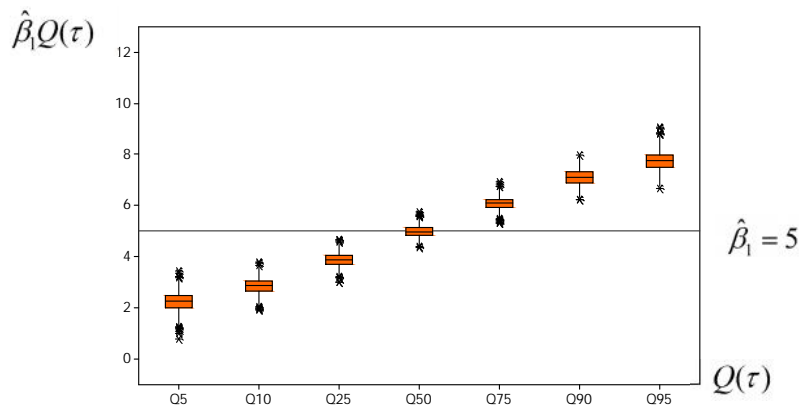
Tabel 4.4 Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Kuantil Studi Simulasi *Error* Heterogen

Sampel	‡	\widehat{S}_p	Nilai	Batas Bawah	Batas Atas	Akurasi
300	5%	\widehat{S}_0	9,443	9,195	9,695	100%
		\widehat{S}_1	2,247	2,008	2,485	100%
	10%	\widehat{S}_0	9,657	9,433	9,875	100%
		\widehat{S}_1	2,867	2,666	3,063	100%
	25%	\widehat{S}_0	9,872	9,677	10,059	100%
		\widehat{S}_1	3,881	3,723	4,04	100%
	50%	\widehat{S}_0	10,004	9,803	10,197	100%
		\widehat{S}_1	4,997	4,853	5,143	100%
	75%	\widehat{S}_0	10,142	9,942	10,337	100%
		\widehat{S}_1	6,105	5,95	6,257	100%
	90%	\widehat{S}_0	10,374	10,158	10,595	100%
		\widehat{S}_1	7,114	6,899	7,324	100%
	95%	\widehat{S}_0	10,576	10,305	10,834	100%
		\widehat{S}_1	7,754	7,489	7,994	100%

Berdasarkan hasil studi simulasi, dapat diketahui bahwa regresi kuantil dapat memodelkan sesuai dengan sebaran data, baik data yang memiliki pola seragam maupun data yang memiliki pola tidak seragam. Hal ini dibuktikan juga dengan plot pada kedua studi simulasi seperti pada Gambar 4.5 dan 4.6 berikut:



Gambar 4.5 *Boxplot* \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen



Gambar 4.6 *Boxplot* \hat{S}_1 Hasil Regresi Kuantil Studi Simulasi *Error* Heterogen

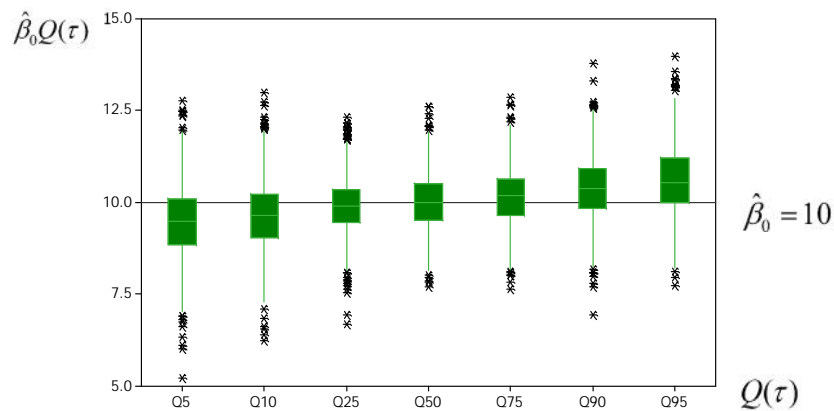
Boxplot pada Gambar 4.5 dan Gambar 4.6 merupakan estimasi parameter regresi kuantil. Berdasarkan Gambar tersebut terlihat bahwa metode regresi kuantil dapat memodelkan data sesuai dengan sebaran data. Secara teori, pada regresi kuantil nilai *intercept* dan *slope* semakin naik sesuai dengan bertambahnya nilai kuantil. Untuk kuantil 50% nilai parameter yang dihasilkan identik dengan parameter pada regresi linier.

Studi simulasi selanjutnya adalah membandingkan ukuran sampel untuk *error* heterogen sesuai dengan persamaan (4.31), dimana ukuran sampel berbeda-beda dengan varians tetap. Hasil estimasi dapat dilihat pada Tabel 4.5 berikut:

Tabel 4.5 Rata-rata Estimasi Parameter dan Taksiran Interval 95% Regresi Linier Berdasarkan Ukuran Sampel

Sampel	\widehat{S}_p	Nilai	Batas atas	Batas Bawah	Akurasi
50	\widehat{S}_0	10,036	9,316	10,75	100%
	\widehat{S}_1	4,977	4,546	5,452	100%
150	\widehat{S}_0	9,981	9,565	10,40	100%
	\widehat{S}_1	5,007	4,751	5,268	100%
300	\widehat{S}_0	10,01	9,703	10,30	100%
	\widehat{S}_1	4,994	4,824	5,180	100%

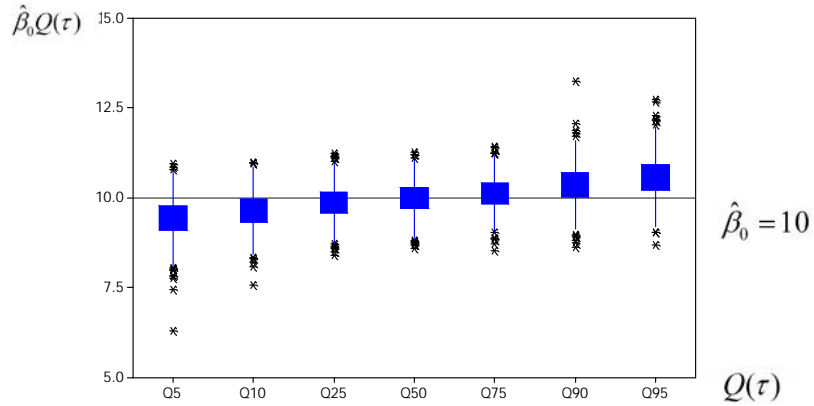
Berdasarkan Tabel 4.5 nilai parameter dari masing-masing sampel berada diantara batas atas dan batas bawah. Hal ini menunjukkan kebenaran dari hasil studi dimulasi. Dari hasil studi simulasi *error* heterogen seluruh nilai \widehat{S}_0 signifikan pada taraf signifikan $\alpha = 5\%$. Secara visual nilai \widehat{S}_0 dan \widehat{S}_1 dapat dilihat pada Gambar 4.7 dan 4.8 berikut.



Gambar 4.7 Boxplot \widehat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen dengan Ukuran Sampel 50.

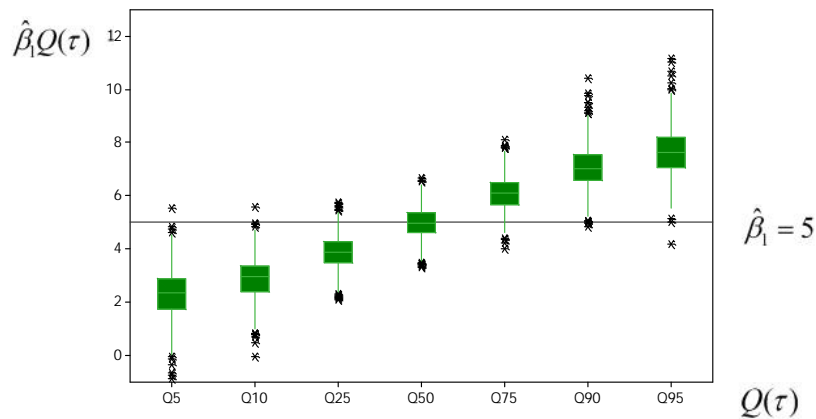
Gambar 4.7 menunjukkan hasil estimasi untuk S_0 pada ukuran sampel 50. Pada gambar tersebut dapat dilihat bahwa nilai \widehat{S}_0 sejajar dengan nilai \widehat{S}_0 pada regresi linier. Sedangkan untuk studi simulasi dimana ukuran sampel

sebanyak 150 dan 300, menunjukkan pola yang hampir sama yaitu meningkat seiring bertambahnya nilai kuantil.

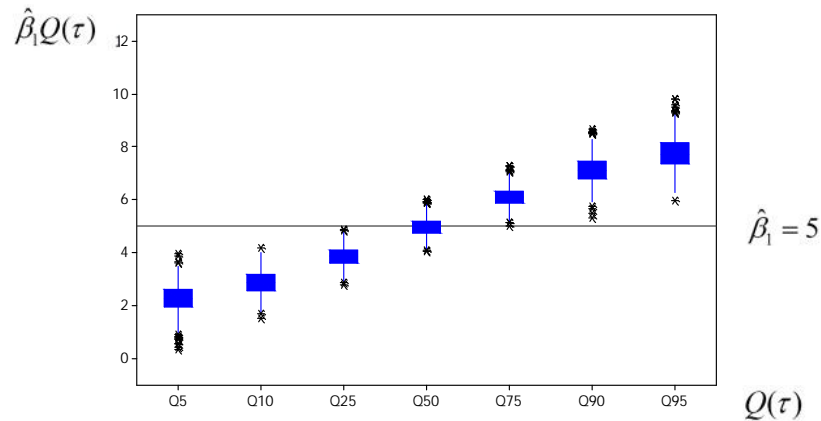


Gambar 4.8 *Boxplot* \hat{S}_0 Hasil Regresi Kuantil Studi Simulasi *Error* Heterogen dengan Ukuran Sampel 150

Hasil estimasi s_1 untuk jumlah sampel 50 dan 150 dapat dilihat pada Gambar 4.9 dan 4.10 berikut. Sedangkan untuk ukuran sampel 300 dapat dilihat pada Gambar 4.5



Gambar 4.9 *Boxplot* Regresi Kuantil \hat{S}_1 Studi Simulasi *Error* Heterogen dengan Ukuran Sampel 50.



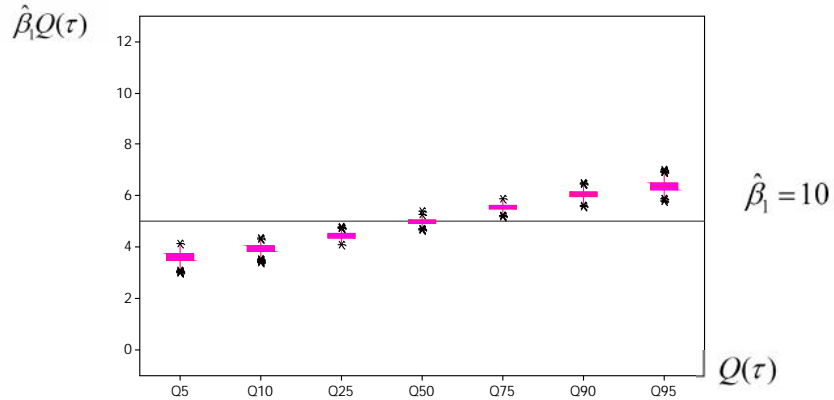
Gambar 4.10 *Boxplot* \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen dengan Ukuran Sampel 150

Secara visual dapat dilihat pola dari ketiga ukuran sampel yaitu 50, 150 dan 300 sama. *Boxplot* ketiga ukuran sampel menunjukkan pola semakin naik sesuai dengan teori pada regresi kuantil. Dari studi simulasi pada ukuran sampel berbeda dapat memberikan informasi bahwa ukuran sampel tidak mempengaruhi regresi kuantil, artinya regresi kuantil dapat digunakan berbagai ukuran sampel.

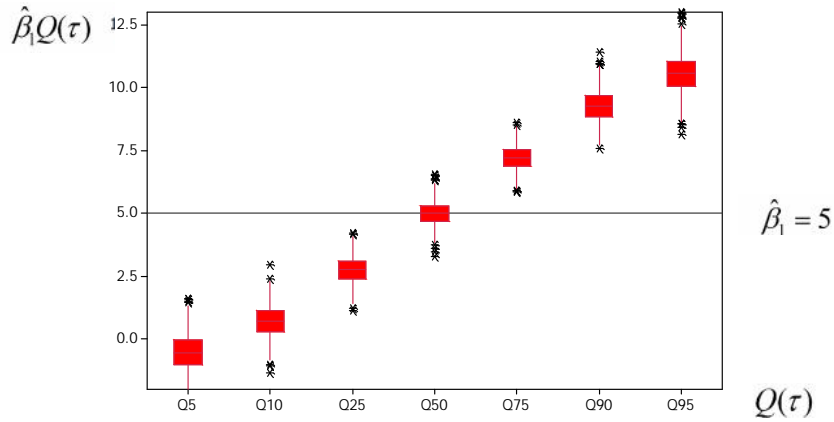
4.2.3 Studi Simulasi Berdasarkan Ukuran Varians

Studi simulasi selanjutnya adalah membandingkan ukuran varians dari *error*, dimana data yang dibangkitkan adalah *error* heterogen. Jumlah sampel yang digunakan untuk ketiga varians yang berbeda adalah 300. Varians pada *error* yang digunakan adalah 0.25, 1, dan 4. Berikut adalah *Boxplot* untuk varians *error* sebesar 0.25 dan 4, sedangkan untuk varians *error* sebesar 1 dapat dilihat pada Gambar 4.5 dan 4.6.

Berdasarkan Gambar 4.11 dan 4.12 dapat dilihat bahwa \hat{S}_1 menunjukkan pola yang sama. Hal ini menginformasikan bahwa perbedaan varians dari *error* tidak mempengaruhi hasil estimasi parameter pada regresi kuantil. Pola yang sama juga terlihat pada Gambar 4.6 dimana nilai varians *error* sebesar 1, artinya regresi kuantil dapat memodelkan data dengan varians berbeda.

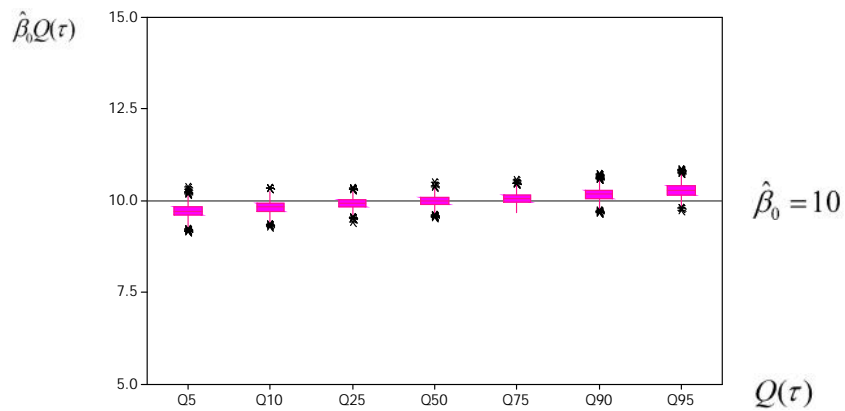


Gambar 4.11 *Boxplot* \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen untuk $\dagger_{v_i}^2 = 0.25$

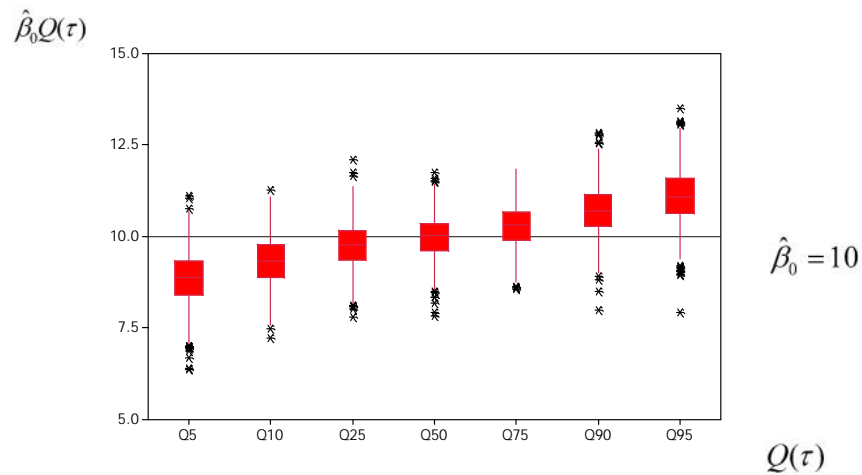


Gambar 4.12 *Boxplot* \hat{S}_1 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen dengan $\dagger_{v_i}^2 = 4$

Sedangkan secara visual untuk \hat{S}_0 dengan $\dagger_{v_i}^2 = 4$ dapat dilihat pada Gambar 4.13 dan Gambar 4.14 untuk $\dagger_{v_i}^2 = 0.25$. Secara visual dapat dilihat pola dari ketiga ukuran varians relatif sama. *Boxplot* ketiga ukuran varians menunjukkan pola semakin naik sesuai dengan teori pada regresi kuantil. Dari studi simulasi pada ukuran varians untuk *error* heterogen dapat memberikan informasi bahwa ukuran varians tidak mmepengaruhi regresi kuantil, artinya regresi kuantil dapat digunakan untuk berbagai ukuran varians.



Gambar 4.13 *Boxplot* \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen dengan $\dagger_{v_i}^2 = 0,25$



Gambar 4.14 *Boxplot* \hat{S}_0 Hasil Regresi Kuantil pada Studi Simulasi *Error* Heterogen dengan $\dagger_{v_i}^2 = 4$

Tabel 4.6 merupakan hasil estimasi regresi kuantil untuk masing-masing varians *error*. Dapat diketahui bahwa nilai estimasi parameter pada kuantil 5% untuk masing-masing varians berbeda. Rata-rata nilai *intercept* dan *slope* untuk varians *error* sebesar 0,25 lebih besar dibandingkan rata-rata pada varians 1 dan 4. Hal yang sama juga terjadi pada setiap kuantil.

Tabel 4.6 Rata-rata Estimasi Parameter Hasil Regresi Kuantil pada Studi Simulasi Berdasarkan Perbedaan Ukuran Varians

Sampel	‡	\widehat{S}_p	$\dagger^2 = 0.25$ $\dagger^2 = 1$ $\dagger^2 = 4$		
			5%	\widehat{S}_0	9,723
		\widehat{S}_1	3,621	2,247	-0,559
10%	\widehat{S}_0	9,833	9,657	9,323	
		\widehat{S}_1	3,936	2,867	0,704
25%	\widehat{S}_0	9,937	9,872	9,758	
		\widehat{S}_1	4,448	3,881	2,751
300	50%	\widehat{S}_0	10,002	10,004	9,995
		\widehat{S}_1	5,001	4,997	4,999
75%	\widehat{S}_0	10,068	10,142	10,297	
		\widehat{S}_1	5,557	6,105	7,210
90%	\widehat{S}_0	10,181	10,374	10,717	
		\widehat{S}_1	6,065	7,114	9,277
95%	\widehat{S}_0	10,282	10,576	11,103	
		\widehat{S}_1	6,384	7,754	10,561

4.3 Karakteristik Dosen ITS di Scopus.

Karakteristik dosen ITS akan digambarkan secara umum dosen ITS yang sudah menjabat sebagai PNS, kecuali dosen UPT dan MKU. Jumlah dosen yang diamati pada penelitian ini sebanyak 900 dosen. Hanya 457 dosen yang memiliki publikasi di Scopus. Berikut adalah deskripsi dari sitasi yang disajikan pada Tabel 4.7

Tabel 4.7 Statistika Deskriptif Sitasi Berdasarkan Jurusan

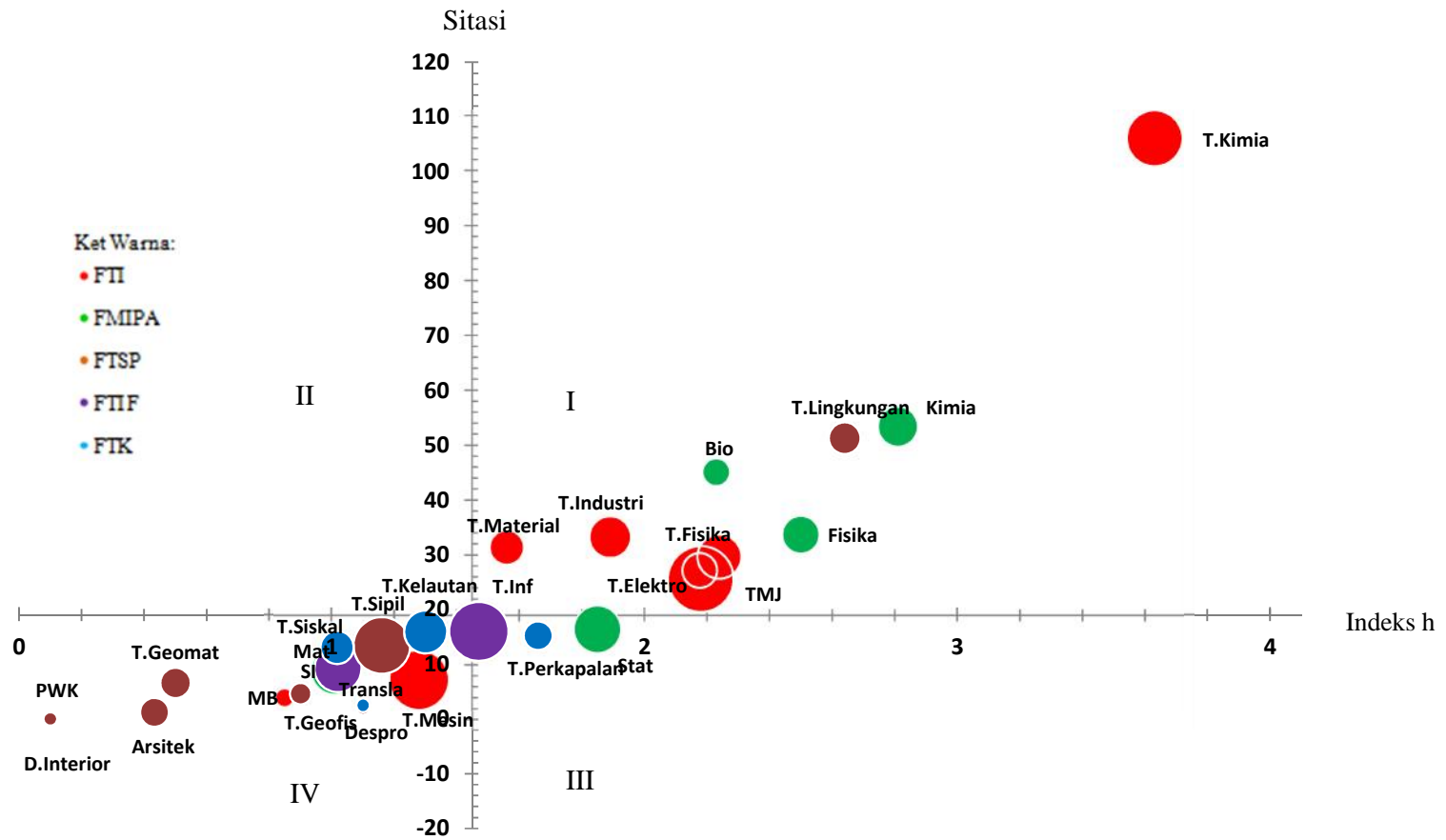
Fakultas	Jurusan	n	Mean	Median	StDev	Min	Max
FMIPA	Biologi	8	45,1	4,0	104,1	0	301
	Fisika	15	33,7	8,0	39,8	1	109
	Kimia	17	53,4	26,0	62,9	0	197
	Matematika	23	8,9	1,0	16,4	0	66
	Statistika	24	16,5	8,0	26,8	0	119
Total		87	31,5	8,0	34,9	0	301
FTI	MB	4	4,0	2,5	4,8	0	11
	T. Elektro	47	25,7	12,0	45,4	0	285
	T. Fisika	21	29,7	6,0	67,8	0	292
	T. Industri	19	33,3	5,0	75,5	0	327
	T. Kimia	34	105,9	29,5	154,7	0	682
	T. Mesin	39	7,3	4,0	10,1	0	43
	T. Material	13	31,4	4,0	61,3	0	188
	TMJ	13	27,2	13,0	41,0	0	147
Total		190	33,1	5,5	46,8	0	682
FTIF	SI	24	9,2	0,0	17,4	0	79
	T. Informatika	38	16,0	5,0	32,3	0	140
Total		62	12,6	2,5	10,6	0	140
FTK	T. Kelautan	20	15,9	1,0	50,6	0	224
	T. Perkapalan	9	15,2	6,0	24,2	0	77
	T. Siskal	12	13,1	0,0	32,5	0	114
	Transla	2	2,5	2,5	0,5	2	3
Total		43	11,7	28,4	1,0	0	224
FTSP	Arsitektur	9	1,2	0,0	2,7	0	8
	D. Interior	1	0,0	0,0	*	0	0
	D. Produk	2	2,0	2,0	1,4	1	3
	PWK	2	0,0	0,0	0,0	0	0
	T.Geofisika	5	4,6	4,0	6,2	0	15
	T.Geomatika	10	6,6	0,0	14,7	0	43
	T.Lingkungan	11	51,3	19,0	65,1	1	221
T.Sipil	35	13,5	4,0	27,0	0	130	
Total		75	9,9	1,0	23,4	0	221
TOTAL ITS		457	19,7	5,5	18,3	0	682

Jurusan yang mempunyai rata-rata sitasi tertinggi adalah Teknik Kimia dengan rata-rata sitasi sebesar 105,9. Jurusan yang ada di FTI mempunyai sitasi diatas rata-rata sitasi ITS, sedangkan jurusan yang ada di FTSP hanya Teknik Lingkungan yang mempunyai rata-rata sitasi diatas rata-rata sitasi ITS.

Tabel 4.8 Statistika Deskriptif Indeks h Setiap Jurusan

Fakultas	Jurusan	n	Mean	Median	StDev	Min	Max
FMIPA	Biologi	8	2,13	1	3,36	0	10
	Fisika	15	2,40	2	1,77	1	6
	Kimia	17	2,71	2	2,23	0	6
	Matematika	23	0,91	1	1,04	0	3
	Statistika	24	1,75	2	1,80	0	7
Total		87	1,98	2	0,85	0	10
FTI	MB	4	0,75	1	0,50	0	1
	T. Elektro	47	2,09	2	1,73	0	7
	T. Fisika	21	2,14	1	2,39	0	10
	T. Industri	19	1,79	1	2,02	0	8
	T. Kimia	34	3,53	2	3,44	0	14
	T. Mesin	39	1,18	1	1,21	0	4
	T. Material	13	1,46	1	1,66	0	5
	TMJ	13	2,08	2	1,75	0	6
Total		190	1,88	1	0,86	0	14
FTIF	SI	24	0,92	0	1,25	0	4
	T. Informatika	38	1,37	1	1,50	0	6
Total		62	1,15	0	0,18	0	6
FTK	T. Kelautan	20	1,20	1	2,09	0	9
	T. Perkapalan	9	1,56	1	1,01	0	3
	T. Siskal	12	0,92	0	2,02	0	7
	Transla	2	1,00	1	0,00	1	1
Total		43	1,17	1	0,99	0	9
FTSP	Arsitektur	9	0,33	0	0,71	0	2
	D. Interior	1	0,00	0	*	0	0
	D. Produk	2	1,00	1	0,00	1	1
	PWK	2	0,00	0	0,00	0	0
	T.Geofisika	5	0,80	1	0,84	0	2
	T.Geomatika	10	0,40	0	0,84	0	2
	T.Lingkungan	11	2,55	2	1,21	1	5
T.Sipil	35	1,06	1	1,08	0	4	
Total		75	0,77	0,50	0,49	0	5
Total ITS		457	1,39	1	0,33	0	14

Jurusan Tansportasi Laut dan Desain Produk mempunyai nilai indeks h maksimum dan minimum yang sama yaitu 1. Jadi semua dosen pada Jurusan Tansportasi Laut dan Desain Produk tidak ada yang memiliki indeks h kurang dari atau lebih dari sama dengan 1. Semua dosen pada jurusan Desain Interior dan Perencanaan Wilayah Kota tidak ada yang



Gambar 4.17 Bubble Chart antara Indeks h dan Sitasi

Berdasarkan Gambar 4.17 menunjukkan hubungan dari indeks h dan sitasi pada setiap jurusan, yang nilainya diperoleh dari rata-rata indeks h dan sitasi dosen pada setiap jurusan masing-masing. Berikut adalah penjelasan masing-masing kuadran.

Kuadran I : Jurusan yang memiliki sitasi dan indeks h yang sama-sama tinggi, yaitu dari **FTI** : Teknik Kimia, Teknik Fisika, Teknik Multimedia Jaringan, Teknik Elektro, Teknik Industri, Teknik Material. **FMIPA** : Kimia, Fisika, Biologi. **FTSP** : Teknik Lingkungan.

Kuadran II : Tidak ada satupun jurusan pada kuadran II, yaitu jurusan yang memiliki sitasi tinggi namun indeks h rendah.

Kuadran III : Jurusan yang memiliki sitasi rendah namun indeks h tinggi yaitu Teknik Perkapalan (**FTK**), Teknik Informatika (**FTIF**), dan Statistika (**FMIPA**).

Kuadran IV : Jurusan pada kuadran ini memiliki sitasi dan indeks h yang sama-sama rendah, yaitu dari **FTK** : Transportasi Laut, Teknik Sistem Perkapalan, dan Teknik Kelautan. **FTSP** : Perencanaan Wilayah Kota, Arsitektur, Teknik Sipil, Teknik Geofisika, Desain Produk, dan Geomatika. **FTI** : Manajemen Bisnis dan Teknik Mesin. **FMIPA** : Matematika. **FTIF** : Sistem Informasi.

Produktivitas dosen ITS di Scopus diketahui melalui metode *cluster analysis*. Metode ini, bertujuan untuk mengelompokkan publikasi dosen ITS berdasarkan variabel metrik. Sebelum dilakukan pengelompokkan, terlebih dahulu memfaktorkan. Analisis faktor dilakukan untuk mengetahui variabel mana yang lebih penting diantara yang lainnya. Hasil analisis faktor ditunjukkan pada Tabel 4.9, diketahui bahwa komponen pertama yang terbentuk terdiri dari variabel sitasi, indeks h, jumlah dokumen. Sedangkan komponen kedua yang terbentuk terdiri dari variabel usia dan lama bekerja.

Tabel 4.9 Analisis faktor

Variabel	Komponen	
	1	2
Indeks h	0,944*	0,051
Sitasi	0,904*	-0,017
Usia	0,027	0,990*
Lama bekerja	0,057	0,989*
Jumlah dokumen	0,834*	0,078

Tabel 4.10 dapat memberikan informasi jumlah masing-masing anggota cluster. Cluster yang dibentuk sebanyak tiga, dimana jumlah anggota cluster satu sebanyak 57 dosen. Jumlah dosen pada cluster kedua adalah 387 dosen, dan jumlah untuk cluster tiga adalah 13 dosen.

Dari Tabel 4.10 diketahui anggota pada cluster satu paling banyak berasal dari jurusan Teknik Elektro. Anggota cluster dua paling banyak berasal dari FTI dimana jurusan yang paling banyak menjadi anggota pada cluster dua adalah Teknik Elektro. Anggota cluster tiga terdiri dari satu dosen dari Biologi, dua dosen dari Teknik Elektro, satu dosen Teknik Fisika, satu dosen dari Teknik Industri, satu dosen dari Teknik Kelautan, satu dosen dari Teknik Multimedia Jaringan, dan enam dosen dari Teknik Kimia.

Karakteristik dosen ITS pada setiap cluster yang terbentuk ditampilkan pada Tabel 4.11 dapat diperoleh informasi bahwa karakteristik cluster satu yaitu memiliki sitasi dan indeks h tertinggi kedua jika dibandingkan dengan cluster yang lain. Rata-rata dosen pada cluster satu berusia 47 tahun dan paling lama bekerja yaitu selama 21 tahun. Anggota cluster satu rata-rata memiliki jumlah dokumen sebanyak 17, dimana rata-rata sitasi yang dimiliki sejumlah 93. Pada cluster dua anggota cluster rata-rata berusia 46 tahun dan telah bekerja selama 19 tahun. Anggota cluster dua memiliki jumlah dokumen lebih sedikit dibandingkan dengan cluster lainnya.

Tabel 4.10 Jumlah masing-masing Anggota Cluster

Fakultas	Jurusan	N	Cluster 1	Cluster 2	Cluster 3
FMIPA	Biologi	8	0	7	1
	Fisika	15	4	11	0
	Kimia	17	7	10	0
	Matematika	23	1	22	0
	Statistika	24	2	22	0
Total		87	14	72	1
FTI	MB	4	0	4	0
	T. Elektro	47	9	36	2
	T. Fisika	21	2	18	1
	T. Industri	19	2	16	1
	T. Kimia	34	7	21	6
	T. Mesin	39	1	38	0
	T. Material	13	2	11	0
	TMJ	13	3	9	1
Total		190	26	153	11
FTIF	SI	24	2	22	0
	T. Informatika	38	6	32	0
Total		62	8	54	0
FTSP	Arsitektur	9	0	9	0
	D. Interior	1	0	1	0
	D. Produk	2	0	2	0
	PWK	2	0	2	0
	T.Geofisika	5	0	5	0
	T.Geomatika	10	0	10	0
	T.Lingkungan	11	3	8	0
T.Sipil	35	2	33	0	
Total		75	5	70	0
FTK	T. Kelautan	20	1	18	1
	T. Perkapalan	9	2	7	0
	T. Siskal	12	1	11	0
	Transla	2	0	2	0
Total		43	4	38	1
Total ITS		457	57	387	13

Tabel 4.11 Karakteristik Dosen Masing-masing cluster.

Variabel	Cluster	n	Rata-rata	St.Dev	Median	Min	Max
Y1	1	57	93,3	61,1	77	19	337
	2	387	7,7	11,5	3	0	95
	3	13	292,5	142,5	285	100	682
Y2	1	57	4,4	1,3	4	2	8
	2	387	1	1	1	0	4
	3	13	8,7	2,5	9	5	14
X2	1	57	21,0	7,9	21	7	38
	2	387	19,3	10,3	19	1	43
	3	13	22,2	7	23	13	36
X3	1	57	16,7	11,7	13	3	58
	2	387	4,3	3,7	3	1	20
	3	13	44,1	29	37	19	125
X1	1	57	46,8	8	46	32	65
	2	387	46	10,3	46	25	68
	3	13	48,2	7,7	47	38	64

Cluster tiga adalah cluster yang memiliki pencapaian publikasi paling tinggi daripada cluster yang lain. Hal tersebut ditunjukkan dari rata-rata sitasi dan indeks h tertinggi. Selain itu, cluster tiga juga memiliki rata-rata dokumen paling banyak. Beberapa nama dosen yang termasuk pada cluster tiga ditampilkan pada Tabel 4.12. Anggota pada cluster tiga terdiri dari dosen FMIPA, FTK dan FTI. Jumlah dosen pada cluster tiga sebanyak 13 dosen, satu dosen dari Biologi, satu dosen dari Teknik Fisika, dua dosen dari Teknik Elektro, satu dosen dari Teknik Industri, satu dosen dari Teknik Kelautan, satu dosen dari Teknik Multimedia Jaringan, dan enam dosen dari Teknik Kimia.

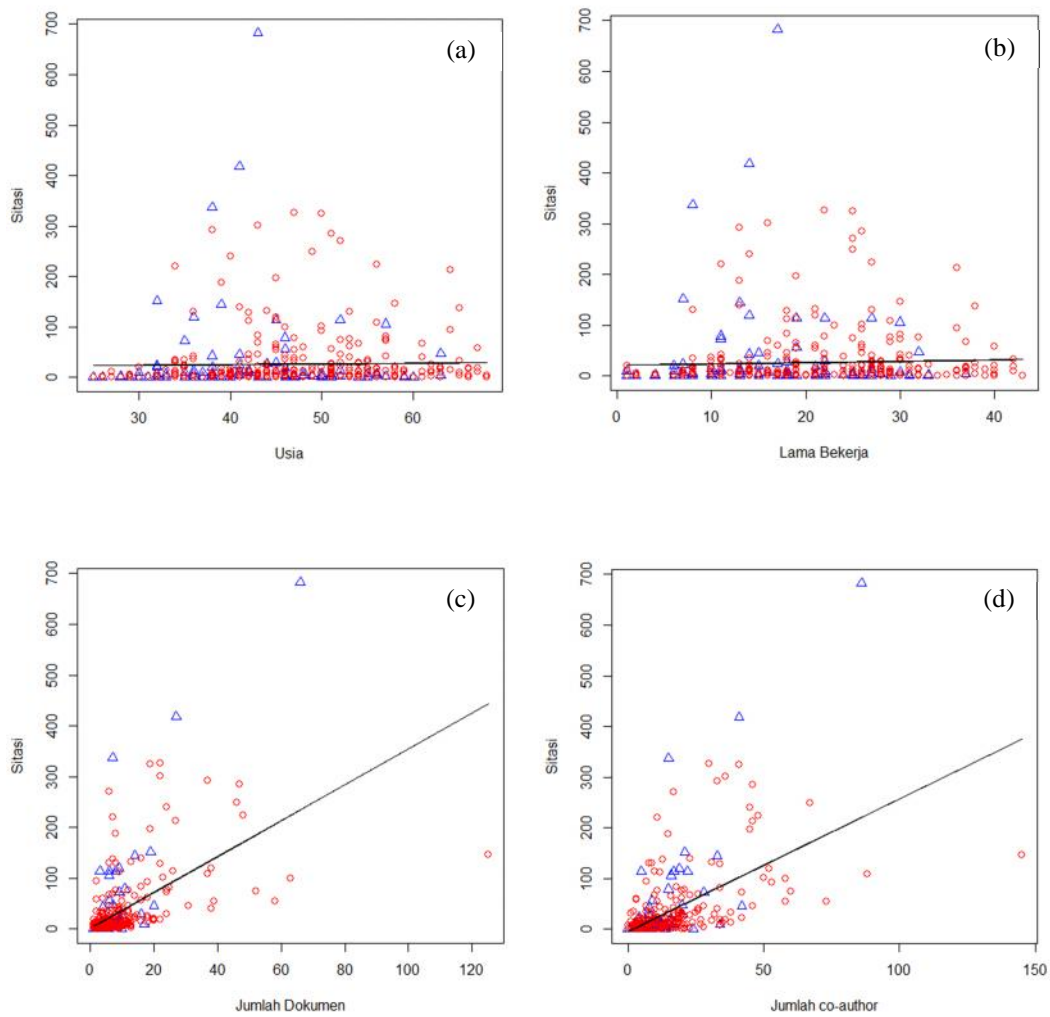
Tabel 4.12 Nama Anggota pada Cluster Tiga

Nama	Jurusan	Fakultas
Dr.techn Endry Nugroho Prasetyo, MT.	BIOLOGI	FMIPA
Prof. Dr. Ir Mochamad Ashari, M.Eng	T.ELEKTRO	FTI
Prof.Ir.Gamantyo Hendrantoro, M.Eng.Ph.D	T.ELEKTRO	FTI
Prof. Dr Sugeng Winardi, M.Eng	T.KIMIA	FTI
Prof.Dr.Ir. Arief Widjaja, M.Eng.	T.KIMIA	FTI
Setiyo Gunawan, S.T., Ph.D	T.KIMIA	FTI
Prof. Dr. Heru Setyawan	T.KIMIA	FTI
Dr. Widiyastuti, ST., MT.	T.KIMIA	FTI
Dr.Eng Siti Machmudah, , ST.,M.Eng.	T.KIMIA	FTI
Agus Muhamad Hatta, ST, MSi, PhD	T.FISIKA	FTI
Prof. Ir. I Nyoman Pujawan, M.Eng, Ph.D, CSCP	T.INDUSTRI	FTI
Prof.Dr.Ir. Mauridhi Hery Purnomo, M.Eng.	TMJ	FTI
Dr. Ir. Wahyudi, M.Sc.	T.KELAUTAN	FTK

4.4 Hubungan antara Variabel pada Analisis Regresi.

Hubungan antar variabel baik dengan variabel dependen maupun variabel independen yang akan digunakan untuk pemodelan regresi akan dijelaskan melalui Gambar 4.18. Tahap awal yang dilakukan adalah identifikasi data secara visual. Identifikasi dilakukan secara visual menggunakan *scatter plot* untuk variabel sitasi dengan seluruh variabel metrik, Gambar 4.19 untuk variabel sitasi dengan indeks h, dan Gambar 4.20 untuk variabel indeks h dengan seluruh variabel metrik.

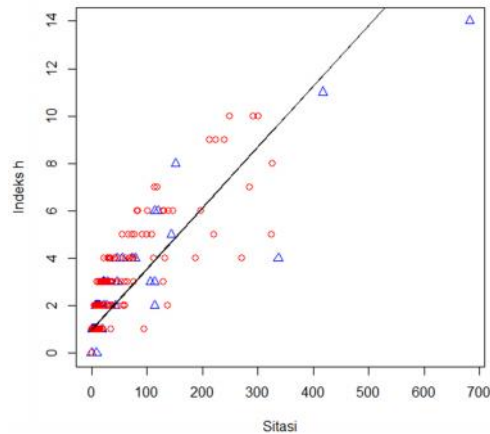
Keberadaan data *outlier* atau sebaran data yang tidak seragam akan mengganggu proses analisis data khususnya dalam hal estimasi parameter. Pada Gambar 4.18 diketahui bahwa sebaran data tidak seragam, ditandai dengan beberapa pengamatan jauh dari pusat rata-rata. Informasi lain menunjukkan bahwa semakin bertambah usia dan lama bekerja tidak mempengaruhi sitasi karena plot yang dihasilkan berupa garis lurus. Segitiga berwarna biru menunjukkan dosen perempuan, sedangkan plot berwarna merah menunjukkan dosen laki-laki.



Gambar 4.18 Scatter Plot (a) Variabel Usia dengan Sitasi, (b) Variabel Lama Bekerja dengan Sitasi, (c) Variabel Jumlah Dokumen dengan Sitasi, (d) Variabel Jumlah *co-authors* dengan Sitasi. Simbol \triangle Merupakan Dosen Laki-laki, dan Simbol \circ Merupakan Simbol untuk Dosen Perempuan

Sedangkan jumlah dokumen dan *co-authors* menunjukkan hal yang positif, karena semakin bertambah jumlah dokumen dan *co-authors* maka sitasi semakin naik. Hal yang sama juga ditunjukkan oleh Gambar 4.19 yaitu

hubungan antara sitasi dengan indeks h. Semakin bertambah sitasi maka indeks h semakin bertambah.



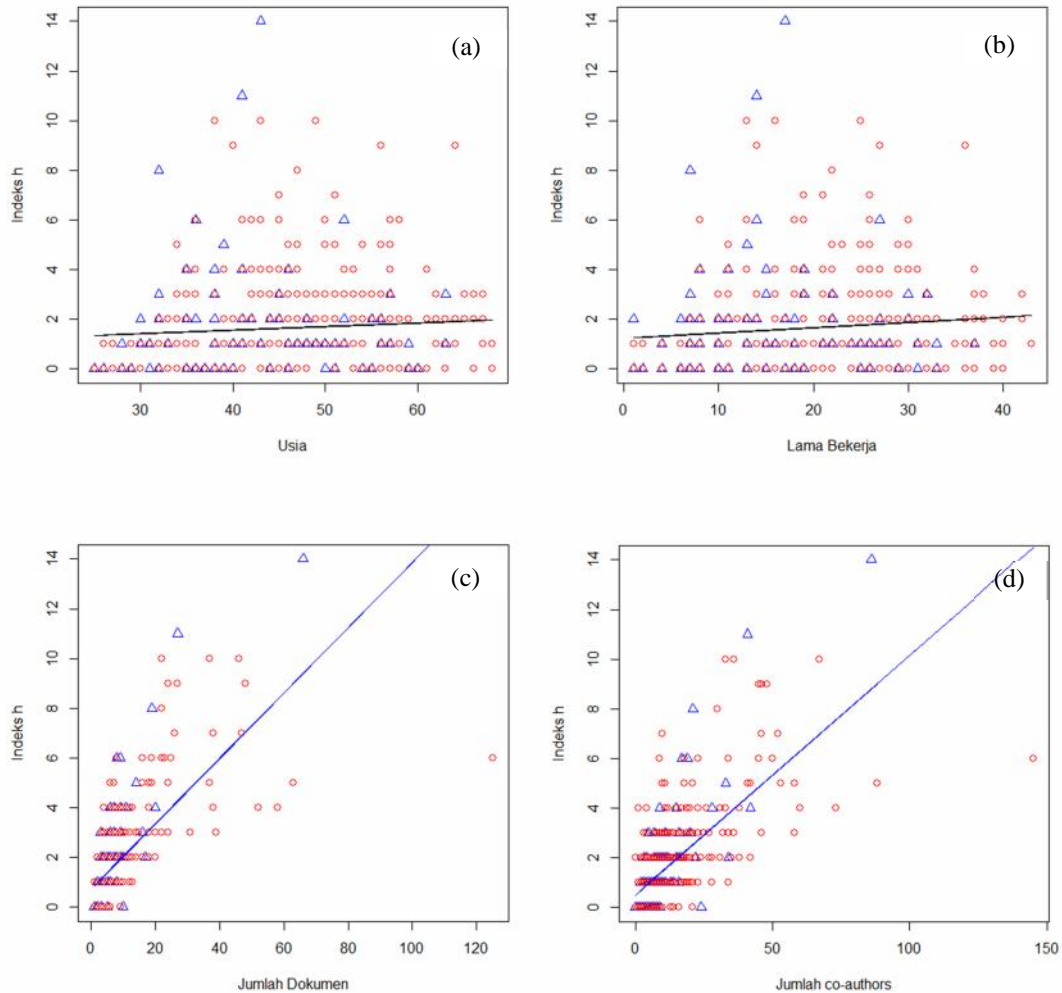
Gambar 4.19 Grafik Hubungan antara Sitasi dan Indeks h. Simbol \circ Merupakan Dosen Laki-laki, dan Simbol \triangle Merupakan Simbol untuk Dosen Perempuan

Tabel 4.13 adalah nilai korelasi antara sitasi dan indeks h terhadap variabel independen kontinu. Nilai korelasi tersebut menunjukkan hasil bahwa sitasi dan indeks h saling berhubungan, usia dan lama bekerja juga saling berhubungan. Hal ini dimungkinkan karena semakin bertambahnya usia maka lama bekerja juga akan bertambah. Jumlah dokumen berhubungan dengan sitasi dan indeks h, dan lama bekerja karena semakin banyak dokumen yang dimiliki seorang peneliti maka sitasi dan indeks h juga akan bertambah.

Tabel 4.13 Matriks Korelasi antar Variabel

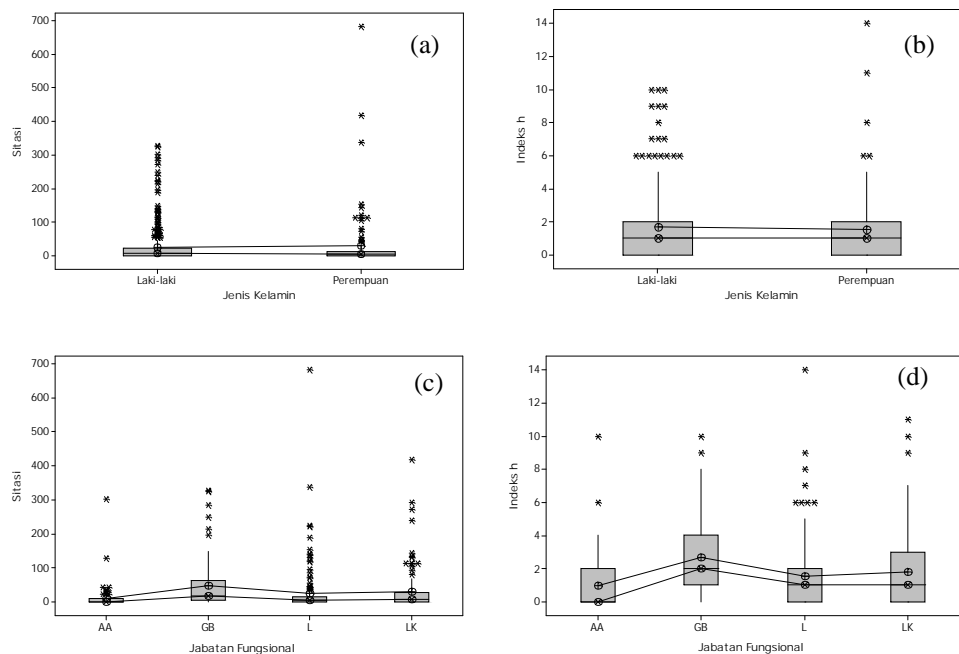
Variabel	Y ₂	Y ₁	X ₁	X ₂	X ₃	X ₄
Y ₂	1,000					
Y ₁	0,829	1,000				
X ₁	0,073	0,019	1,000			
X ₂	0,108	0,040	0,962	1,000		
X ₃	0,694	0,579	0,089	0,114	1,000	
X ₄	0,685	0,578	0,120	0,147	0,906	1,000

Selain itu jumlah *co-authors* juga berhubungan dengan sitasi, indeks h, usia, lama bekerja dan jumlah dokumen. Semakin bertambahnya usia maka lama bekerja dosen di ITS juga akan bertambah. Semakin bertambahnya jumlah dokumen maka jumlah *co-authors* juga akan semakin bertambah.



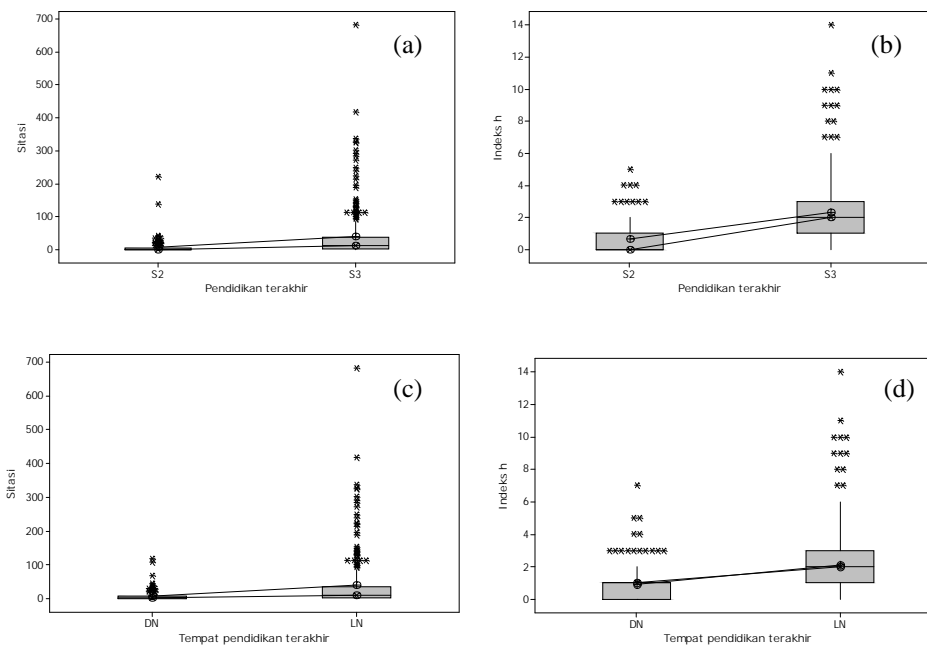
Gambar 4.20 Scatter Plot (a) Variabel Usia dengan Indeks h, (b) Variabel Lama Bekerja dengan Indeks h, (c) Variabel Jumlah Dokumen dengan Indeks h, (d) Variabel Jumlah *co-authors* dengan Indeks h. Simbol \triangle Merupakan Dosen Laki-laki, dan Simbol \circ Merupakan Simbol untuk Dosen Perempuan

Gambar 4.20 diketahui bahwa semakin bertambah usia dan lama bekerja tidak mempengaruhi indeks h karena plot yang dihasilkan berupa garis lurus. Sedangkan jumlah dokumen dan *co-authors* menunjukkan hal yang positif, karena semakin bertambah jumlah dokumen dan *co-authors* maka indeks h semakin naik. Informasi lain dari Gambar 4.20 yaitu terlihat sebaran data tidak seragam, ditandai dengan beberapa pengamatan yang jauh dari pusat rata-rata. Hubungan antara variabel dependen dengan variabel independen *dummy* dapat dilihat pada Gambar 4.21 dan Gambar 4.22. Gambar tersebut menunjukkan bahwa terdapat perbedaan rata-rata antar kategori dalam satu variabel yaitu jabatan fungsional jenis kelamin, pendidikan terakhir, dan tempat pendidikan terakhir. Selain itu, dapat diketahui adanya data *oulier* untuk masing-masing variabel independen.



Gambar 4.21 Boxplot (a) Variabel Jenis Kelamin dengan Sitasi, (b) Variabel Indeks h dengan Jenis Kelamin, (c) Variabel Jabatan Fungsional dengan Sitasi, (d) Variabel Indeks h dengan Jenis Kelamin

Data *outlier* pada Gambar 4.21 (a) untuk dosen dengan jenis kelamin perempuan yaitu Siti Machmudah, Siti Nurkhamidah, Diah Susanti, dan Widiyawati,. Sedangkan dosen laki-laki yang memiliki sitasi melebihi dosen lainnya adalah Arief Widjaja, I Nyoman Pujawan, Endry Nugroho dari jurusan Biologi dan Moh. Ashari. Data *outlier* pada Gambar 4.21 (a) sama dengan data *outlier* pada Gambar 4.21 (b). Adanya data outlier di sebabkan oleh sitasi dan indeks h masing-masing dosen yang berbeda. Dengan kata lain, terdapat beberapa dosen yang memiliki sitasi dan indeks h melebihi dosen ITS lainnya.



Gambar 4.22 Boxplot (a) Variabel Pendidikan Terakhir dengan Sitasi, (b) Variabel Indeks h dengan Pendidikan Terakhir, (c) Variabel Tempat Pendidikan Terakhir dengan Sitasi, (d) Variabel Indeks h dengan Tempat Pendidikan Terakhir

Gambar 4.21 (c) dan (d) menginformasikan bahwa terdapat *outlier* pada masing-masing kategori. Untuk jabatan fungsional dengan kategori Guru Besar (GB) sitasi tertinggi yaitu Arief Widjaja dan indeks h tertinggi adalah Heru Setyawan, kedua dosen berasal dari jurusan Teknik Kimia. Sedangkan untuk jabatan fungsional Lektor Kepala (LK) sitasi dan indeks h tertinggi adalah

Widiyastuti dari jurusan Teknik Kimia. Pada jabatan fungsional sebagai Lektor (L) adalah dosen yang memiliki sitasi dan indeks h melebihi dosen lain, yaitu Siti Machmudah.

Gambar 4.22 secara visual terlihat bahwa dosen dengan pendidikan terakhir S3 memiliki rata-rata yang lebih tinggi dibandingkan dengan dosen yang berpendidikan S2, baik dari sitasi maupun indeks h. Data *outlier* pada Boxplot menunjukkan dosen dengan sitasi dan indeks h melebihi rata-rata. Sitasi tertinggi dengan pendidikan terakhir S3 adalah Siti Machmudah dari jurusan Teknik Kimia. Gambar 4.22 (c) merupakan hubungan antara sitasi dengan tempat pendidikan terakhir dosen. Dosen yang menempuh pendidikan terakhir di luar negeri (LN) dan memiliki jumlah sitasi lebih tinggi dari dosen lainnya yaitu: Siti machmudah dan widyastuti dari jurusan Teknik kimia.

Perbedaan rata-rata masing-masing variabel independen pada Gambar 4.21 dan 4.22 lebih diperjelas pada Tabel 4.14 berikut.

Tabel 4.14 Rata-rata Sitasi dengan Variabel Independen *Dummy*

Variabel	Sitasi	Indeks h
Jenis Kelamin		
LK	25,38	1,68
PR	30,11	1,55
Jabatan Fungsional		
GB	46,85	2,68
LK	29,96	1,78
L	25,06	1,53
AA	9,644	0,95
Pendidikan Terakhir		
S3	39,73	2,30
S2	6,160	0,67
Tempat Pendidikan Terakhir		
LN	38,80	2,13
DN	6,680	0,90

Berdasarkan Tabel 4.14 diketahui bahwa rata-rata sitasi dosen perempuan lebih tinggi dibandingkan rata-rata sitasi dosen laki-laki yaitu sebesar 25,38 dan

30,11. Berbeda dengan indeks h, rata-rata indeks h dosen laki-laki lebih tinggi dibandingkan dengan dosen perempuan. Dosen yang memiliki jabatan fungsional sebagai GB mempunyai rata-rata sitasi yang lebih tinggi dibandingkan jabatan fungsional yang lain. Seorang dosen yang berpendidikan S3 mempunyai rata-rata sitasi dan indeks h yang lebih tinggi dibandingkan dosen yang pendidikannya S2. Seorang dosen yang melanjutkan studi terakhirnya diluar negeri (LN) mempunyai rata-rata sitasi yang lebih tinggi dibandingkan dosen yang melanjutkan studi terakhirnya di dalam negeri (DN). Hal tersebut juga berlaku untuk rata-rata indeks h. Dosen yang melanjutkan studi terakhirnya diluar negeri mempunyai rata-rata indeks h yang lebih tinggi dibandingkan dosen yang melanjutkan studinya didalam negeri.

4.4 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Linier Berbasis Model Rekursif.

Pada tahap pemodelan dengan menggunakan regresi linier berbasis model rekursif variabel dependen yang digunakan adalah sitasi untuk model pertama dan indeks h untuk model kedua, dimana pada model kedua sitasi sebagai variabel independen. Model yang dihasilkan dari regresi linier berbasis model rekursif adalah sebagai berikut:

$$\hat{Y}_1 = 12,244 - 0,255X_1 - 0,342X_2 + 1,673X_3 + 1,360X_4 + 13,042X_5 + 9,514X_{6(1)} + 14,842X_{6(2)} + 14,024X_{6(3)} - 6,409X_7 - 15,387X_8.$$

(4.32)

dan

$$\hat{Y}_2 = 1,622 + 0,019Y_1 - 0,022X_1 + 0,016X_2 + 0,037X_3 + 0,014X_4 + 0,0005X_5 + 0,156X_{6(1)} + 0,121X_{6(2)} + 0,036X_{6(3)} - 0,556X_7 - 0,104X_8.$$

Model rekursif pertama dilakukan pengujian serentak dengan menggunakan taraf signifikan Γ sebesar 5% dan menghasilkan nilai F sebesar 27,88. Nilai tersebut lebih besar dari nilai $F_{0,05;10;446}$ yaitu 1,85. Sehingga dapat

disimpulkan bahwa minimal ada satu variabel yang signifikan. Model tersebut memiliki nilai R^2 sebesar 37,1% artinya model tersebut dapat menjelaskan variabilitas data sebesar 37,1%. Sedangkan untuk model kedua, hasil pengujian serentak menunjukkan bahwa nilai F sebesar 145,2 lebih besar dari $F_{0,05;11;445}$ yaitu 1,81. Kesimpulan yang diperoleh adalah minimal ada satu variabel yang signifikan. Selanjutnya dilakukan uji secara parsial yang disajikan pada Tabel 4.15 untuk variabel dependen sitasi dan Tabel 4.17 untuk variabel dependen indeks h.

Hasil uji parsial pada Tabel 4.15 menunjukkan bahwa nilai t signifikan pada variabel jumlah dokumen, jumlah *co-authors*, jenis kelamin, dan tempat pendidikan terakhir. Namun beberapa variabel tidak berpengaruh terhadap sitasi, sedangkan pada pengujian serentak menunjukkan hasil signifikan. Hal ini mengindikasikan terjadi multikolinieritas. Oleh karena itu, perlu dilakukan pemilihan model terbaik seperti pada Tabel 4.16.

Tabel 4.15 Uji Parsial pada Variabel Dependen Sitasi

Variabel	Coef	SE Coef	t	p-value
Konstan	12,24	24,74	0,50	0,62
X ₁	-0,26	0,87	-0,29	0,77
X ₂	-0,34	0,90	-0,38	0,70
X ₃	1,67	0,54	3,09	0,00
X ₄	1,36	0,40	3,36	0,00
X ₅	13,04	5,86	2,23	0,03
X ₆₍₁₎	9,51	11,88	0,80	0,42
X ₆₍₂₎	14,84	8,82	1,68	0,09
X ₆₍₃₎	14,02	7,72	1,82	0,07
X ₇	-6,41	5,94	-1,08	0,28
X ₈	-15,39	5,44	-2,83	0,00

Pemilihan model terbaik dilakukan menggunakan metode *forward*. Tabel 4.16 menunjukkan bahwa seluruh variabel signifikan. Variabel tersebut yaitu jumlah dokumen, jumlah *co-authors*, jenis kelamin, dan tempat pendidikan terakhir.

Model rekursif dari hasil pemilihan model terbaik dapat dilihat pada persamaan (4.33).

Tabel 4.16 Pemilihan Model Terbaik pada Variabel Dependen Sitasi

Variabel	Coef	SE Coef	<i>t</i>	p-value
Konstan	1,43	4,05	0,35	0,72
X ₃	1,74	0,54	3,22	0,00
X ₄	1,36	0,40	3,40	0,00
X ₅	14,36	5,68	2,53	0,01
X ₈	-17,05	4,98	-3,43	0,00

Pengujian parsial untuk variabel dependen indeks h diperoleh hasil pada Tabel 4.17. Terdapat tiga variabel yang signifikan, yaitu variabel sitasi, jumlah dokumen, dan pendidikan terakhir. Jumlah variabel yang berpengaruh terhadap indeks h lebih sedikit dibandingkan dengan yang tidak berpengaruh.

Tabel 4.17 Uji Parsial pada Variabel Dependen Indeks h

Variabel	Coef	SE Coef	<i>t</i>	p-value
Konstan	1,621	0,46	3,55	0,00
Y ₁	0,019	0,001	21,65	0,00
X ₁	-0,022	0,02	-1,38	0,17
X ₂	0,016	0,02	1,00	0,32
X ₃	0,037	0,01	3,66	0,00
X ₄	0,013	0,01	1,81	0,07
X ₅	0,0005	0,11	-0,01	1,00
X ₆₍₁₎	0,156	0,22	0,71	0,48
X ₆₍₂₎	0,120	0,16	0,74	0,46
X ₆₍₃₎	0,036	0,14	0,26	0,80
X ₇	-0,556	0,11	-5,06	0,00
X ₈	-0,105	0,10	-1,03	0,30

Pada pengujian serentak diperoleh hasil signifikan, sedangkan pada pengujian parsial hanya terdapat tiga variabel yang signifikan. Hal ini menunjukkan adanya kasus multikolinieritas. Pemilihan model terbaik seperti pada Tabel 4.18 menjadi solusi dari permasalahan ini.

Seluruh variabel dari hasil pemilihan model terbaik dengan metode *forward* disajikan pada Tabel 4.17. Terdapat tiga variabel yang berpengaruh terhadap indeks h, yaitu sitasi, jumlah dokumen, dan pendidikan terakhir.

Tabel 4.18 Pemilihan Model Terbaik pada Variabel Dependen Indeks h

Variabel	Coef	SE Coef	t	p-value
Konstan	1,02	0,07	14,49	0,00
Y ₁	0,02	0,001	22,88	0,00
X ₃	0,05	0,01	10,27	0,00
X ₇	-0,62	0,09	-6,68	0,00

Model rekursif dari pemilihan model terbaik dapat dilihat pada persamaan (4.33) berikut:

$$\begin{aligned} \hat{Y}_1 &= 1,43 + 1,74X_3 + 1,36X_4 + 14,36X_5 - 17,05X_8. \\ \hat{Y}_2 &= 1,02 + 0,02Y_1 + 0,05X_3 - 0,62X_7. \end{aligned} \quad (4.33)$$

Berdasarkan hasil estimasi pada Tabel 4.16 dan 4.18 diketahui variabel yang berpengaruh terhadap sitasi adalah.

a. Jumlah dokumen

Apabila jumlah dokumen dosen bertambah sebanyak 1 dokumen, maka sitasi akan bertambah sebesar 1,74 dengan syarat variabel-variabel yang lain dianggap konstan.

b. Jumlah *co-authors*

Apabila jumlah *co-authors* dosen bertambah 10 orang, maka sitasi akan bertambah sebesar 13,6 dengan syarat variabel-variabel yang lain dianggap konstan.

c. Jenis kelamin

Rata-rata sitasi dosen perempuan lebih besar 14,36 dibandingkan dengan dosen laki-laki dengan syarat variabel-variabel yang lain dianggap konstan.

d. Tempat pendidikan terakhir

Rata-rata sitasi dosen yang menempuh pendidikan di dalam negeri lebih kecil 17,05 dibandingkan dengan dosen yang pernah menempuh pendidikan di luar negeri dengan syarat variabel-variabel yang lain dianggap konstan.

Sedangkan variabel yang berpengaruh terhadap indeks h adalah.

a. Sitasi

Apabila terdapat 50 orang yang mengutip karya dosen di Scopus, maka indeks h dosen tersebut akan bertambah sebesar 1 dengan syarat variabel-variabel yang lain dianggap konstan.

b. Jumlah dokumen

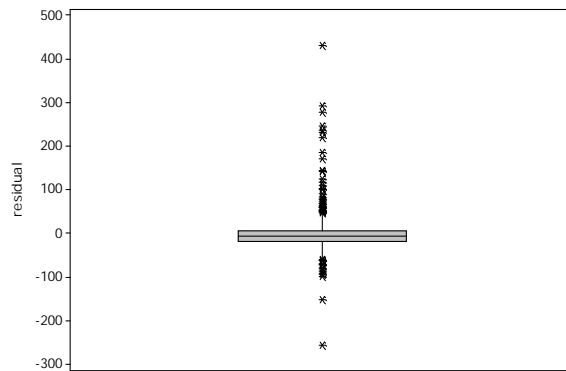
Apabila jumlah dokumen dosen bertambah sebanyak 1 dokumen, maka indeks h akan bertambah sebesar 0,05 dengan syarat variabel-variabel yang lain dianggap konstan.

c. Pendidikan terakhir

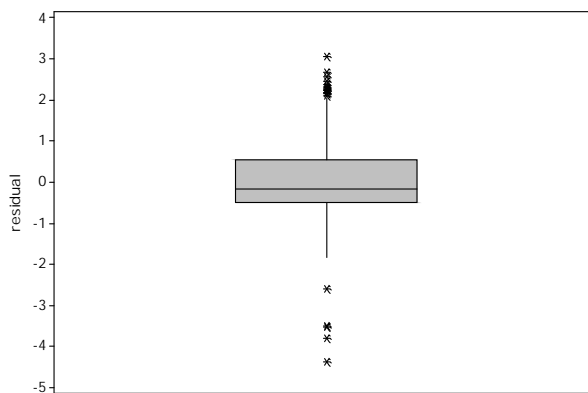
Rata-rata indeks h dosen lulusan S2 lebih kecil 0,62 dibandingkan dengan dosen lulusan S3 dengan syarat variabel-variabel yang lain dianggap konstan.

Model pada persamaan (4.32) merupakan model rekursif dari hasil pemilihan model terbaik menggunakan metode *forward*. Syarat dari model rekursif adalah tidak ada korelasi antara residual model pertama dan kedua. Hasil korelasi dari residual kedua model menunjukkan bahwa tidak ada korelasi. Nilai korelasi yang dihasilkan sebesar -0,021 dengan *p-value* sebesar 0,654. Pada taraf signifikan α 5%, diperoleh keputusan gagal menolak H_0 dan disimpulkan bahwa tidak ada hubungan antara residual dari model pertama dan model kedua. Dengan kata lain model rekursif dapat digunakan dalam persamaan tersebut.

Selanjutnya dilakukan pengujian asumsi distribusi normal untuk model pertama dimana sitasi menjadi variabel dependen. Hasil yang diperoleh menunjukkan bahwa data tidak berdistribusi normal, dibuktikan dengan nilai p-value sebesar $0,000 < \alpha (5\%)$. Untuk model kedua, hasil pengujian distribusi normal menunjukkan bahwa data tidak berdistribusi normal. Nilai p-value yang diperoleh sebesar 0,000, nilai tersebut kurang dari $\alpha (5\%)$. Asumsi yang tidak terpenuhi diduga karena adanya data *outlier* seperti yang terlihat pada Gambar 4.23 dan 4.24.



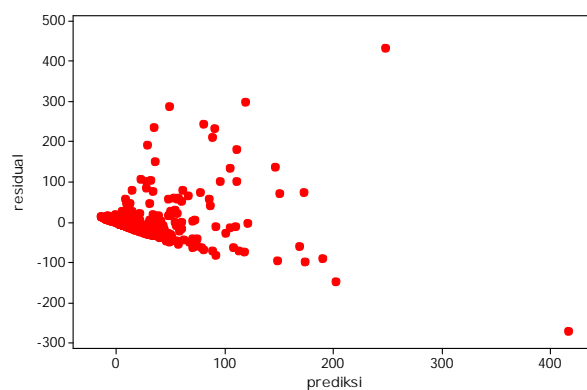
Gambar 4.23 Boxplot Residual dari Model Rekursif Pertama



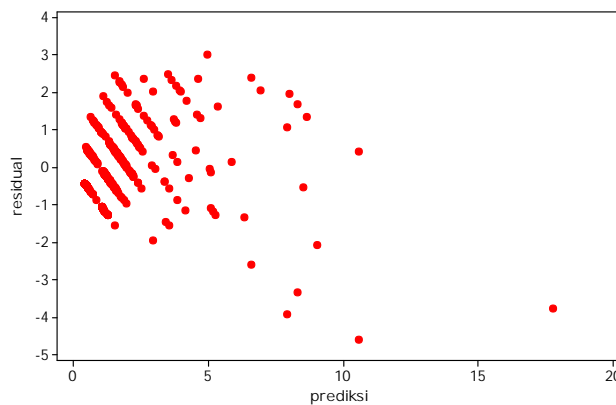
Gambar 4.24 Boxplot Residual dari Model Rekursif Kedua

Outlier pada residual Gambar 4.23 adalah Siti Nurkhamidah dari jurusan Teknik Kimia. Sedangkan data *outlier* kedua pada Gambar 4.24 adalah Siti machmudah dari Teknik Kimia. Asumsi normalitas tidak terpenuhi, diduga

karena adanya data outlier atau sebaran data yang tidak seragam. Terdapat beberapa cara menangani data *outlier* salah satunya dengan membuang pengamatan. Namun, cara tersebut bukan menyelesaikan masalah. Dengan membuang data *outlier*, maka akan mengakibatkan munculnya data outlier baru yaitu dosen yang memiliki sitasi dan indeks h tinggi. Sebaran data yang tidak seragam dapat diketahui melalui plot antara residual dengan hasil prediksi, seperti yang disajikan pada Gambar 4.25 dan 4.26.



Gambar 4.25 Scatter Plot Residual dan Prediksi dari Model Rekursif Pertama



Gambar 4.26 Scatter Plot Residual dan Prediksi dari Model Rekursif Kedua

Metode yang sesuai untuk memodelkan sebaran data tersebut adalah regresi kuantil. Pemodelan mengenai regresi kuantil akan dibahas pada sub bab 4.6.

4.6 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Kuantil Berbasis Model Rekursif.

Kuantil dapat dioperasikan dengan pengurutan sampel pengamatan dimana setiap kuantil mencirikan titik tertentu dari sebaran bersyarat, yaitu pada bagian ekor atau pusat. Kombinasi berbagai nilai kuantil akan menghasilkan deskripsi lengkap tentang sebaran bersyarat. Regresi kuantil sangat berguna untuk sebaran bersyarat yang tidak simetris, padat dibagian ekor, atau sebaran data yang terpotong. Terdapat dua pembahasan dalam sub bab ini, yaitu analisis regresi kuantil multivariabel dan analisis regresi kuantil univariat menggunakan variabel independen kontinu.

4.6.1 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Kuantil Berbasis Model Rekursif Multivariabel

Model regresi kuantil dalam penelitian ini digunakan untuk menggambarkan publikasi dosen ITS di Scopus. Tabel 4.19 adalah hasil estimasi untuk kuantil 5%, 50%, dan 95%, hasil estimasi secara keseluruhan dapat dilihat pada Lampiran V.

Tabel 4.19 Estimasi Parameter Regresi Kuantil untuk Variabel Dependen Sitasi

Variabel	Kuantil (‡)		
	5%	50%	95%
Konstan	-7.669	-0.4849	23.859
X ₁	0.144	-0.0443	0.316
X ₂	-0.119	0.0222	-0.316
X ₃	1.198	1.4880	9.626
X ₄	-0.232	0.4209	1.579
X ₅	0.561	0.3101	0.947
X ₆₍₁₎	-0.385	-0.3699	-15.942
X ₆₍₂₎	0.811	0.2215	-2.842
X ₆₍₃₎	-0.097	0.4431	-0.632
X ₇	0.085	-0.8924	-0.632
X ₈	0.490	-0.2215	-42.643

Berdasarkan Tabel 4.19 diperoleh hasil estimasi parameter dengan metode regresi kuantil. Variabel usia, jumlah dokumen, dan jumlah *co-authors* signifikan pada kuantil 5%. Pada kuantil 50% variabel yang berpengaruh terhadap sitasi adalah jumlah dokumen, dan pendidikan terakhir dosen. Untuk kuantil 95% variabel jumlah dokumen dan tempat pendidikan terakhir dosen berpengaruh terhadap sitasi. Dari hasil signifikansi parameter dapat diketahui bahwa variabel jumlah dokumen berpengaruh terhadap sitasi pada setiap kuantil kecuali kuantil 25%.

Hasil estimasi parameter dengan metode regresi kuantil untuk model rekursif kedua seperti yang ditampilkan pada Tabel 4.20 berikut:

Tabel 4.20 Estimasi Parameter Regresi Kuantil untuk Variabel Dependen Indeks h

Variabel	Kuantil (‡)		
	5%	50%	95%
Konstan	-0.1724	0.7526	1.420
Y ₁	0.0141	0.0202	0.039
X ₁	0.0024	-0.0041	0.009
X ₂	-0.0032	0.0045	0.001
X ₃	0.0425	0.0803	0.078
X ₄	-0.0093	0.0086	0.005
X ₅	0.0307	0.0280	-0.254
X ₆₍₁₎	0.0778	0.0469	-0.358
X ₆₍₂₎	0.0451	0.0098	-0.312
X ₆₍₃₎	0.0314	-0.0045	-0.416
X ₇	-0.0190	-0.7506	-0.160
X ₈	0.0353	0.0041	-0.031

Berdasarkan Tabel 4.20 menampilkan hasil estimasi parameter dimana variabel dependen yang digunakan adalah indeks h. Dari Tabel tersebut diketahui bahwa variabel sitasi berpengaruh terhadap indeks h di setiap kuantil. Variabel jumlah dokumen berpengaruh terhadap indeks h di setiap kuantil, kecuali pada kuantil 25%. Seluruh estimasi pada regresi kuantil dapat dilihat pada Lampiran V. Pada

kuantil 50% selain variabel sitasi dan jumlah dokumen, variabel yang berpengaruh terhadap sitasi adalah pendidikan terakhir yang ditempuh dosen.

Hasil pemodelan pada Tabel 4.19 dan 4.20 selanjutnya digunakan untuk menghitung nilai estimasi, yang kemudian dilakukan *backtesting* yaitu membandingkan y_i dengan \hat{y}_i . Hasil perhitungan *backtesting* disajikan pada Tabel 4.20. Metode *backtesting* digunakan untuk mengetahui sejauh mana suatu metode menghasilkan akurasi yang baik dalam pemodelan. Untuk mendapatkan nilai \dagger^* untuk pemodelan data publikasi sama dengan mencari nilai $(100 \times \dagger)\%$ kuantil dari y dengan syarat \mathbf{X} . Suatu model dikatakan akurat apabila $\dagger = \dagger^*$. Pada penelitian ini $P(R_t \leq V_t | F_{t-1}) = \dagger^*$, dimana $R_t = y_i$ dan $V_t | F_{t-1} = \hat{y}(\dagger) | \mathbf{X}$ (Smith *et al.*, 2009).

Tabel 4.21 Hasil *Backtesting* untuk Model Regresi Kuantil Berbasis Model Rekursif

Kuantil (\dagger)	Model 1 (\dagger^*)	Model 2 (\dagger^*)
5%	0,059	0,061
10%	0,089	0,116
25%	0,254	0,234
50%	0,492	0,490
75%	0,737	0,753
90%	0,897	0,912
95%	0,958	0,958

Hasil *backtesting* untuk publikasi dosen model rekursif persamaan pertama dan kedua pada Tabel 4.21 menunjukkan bahwa untuk model rekursif persamaan pertama pada kuantil 25% nilai $\dagger^* = \dagger$. Artinya terdapat 25 persen pengamatan yang nilai sitasinya lebih kecil dari nilai dugaan sitasi yang diperoleh dari model. Secara keseluruhan hanya tiga kuantil yang memperoleh nilai sama dengan nilai \dagger^* yaitu kuantil 0,25, 0,50, dan 0,90. Selanjutnya untuk kuantil 50% nilai $\dagger^* = \dagger$ artinya terdapat 50 persen pengamatan yang nilai observasi (sitasinya) lebih kecil dari nilai dugaan sitasi yang diperoleh dari model.

Sedangkan untuk model rekursif persamaan kedua, terdapat tiga $\dagger^*=\dagger$ yaitu pada kuantil 50%, 75%, dan 90%.

Hasil pemodelan pada Tabel 4.19 dan 4.20 untuk regresi kuantil dapat dituliskan seperti pada persamaan (4.33) berikut:

Model untuk variabel dependen sitasi

$$\begin{aligned}
 \hat{Q}_{0,05}(y | \mathbf{x}) &= -7,669 + 0,144X_1 - 0,119X_2 + 1,198X_3 - 0,232X_4 + 0,561X_5 \\
 &\quad - 0,385X_{6(1)} + 0,811X_{6(2)} - 0,097X_{6(3)} + 0,085X_7 + 0,490X_8. \\
 \hat{Q}_{0,50}(y | \mathbf{x}) &= -0,485 - 0,044X_1 + 0,022X_2 + 1,488X_3 + 0,421X_4 + 0,310X_5 \\
 &\quad - 0,3699X_{6(1)} + 0,222X_{6(2)} + 0,443X_{6(3)} - 0,892X_7 - 0,222X_8. \\
 \hat{Q}_{0,95}(y | \mathbf{x}) &= 23,859 + 0,316X_1 - 0,316X_2 + 9,626X_3 + 1,579X_4 + 0,947X_5 \\
 &\quad - 15,942X_{6(1)} - 2,842X_{6(2)} - 0,632X_{6(3)} - 0,632X_7 - 42,64X_8.
 \end{aligned} \tag{4.34}$$

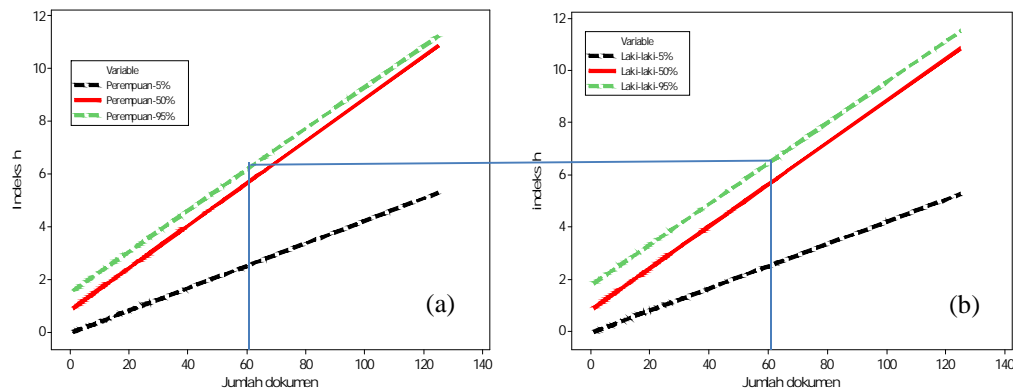
Berdasarkan hasil estimasi pada Tabel 4.19, diketahui bahwa koefisien variabel usia dan jumlah dokumen bertanda positif pada kuantil 5%. Artinya variabel usia dan jumlah dokumen berpengaruh positif terhadap sitasi. Dengan kata lain, rata-rata sitasi dosen ITS yang berusia lebih tua lebih tinggi 0,144 daripada dosen yang berusia muda. Untuk jumlah dokumen apabila jumlah dokumen bertambah sebanyak 1 dokumen, maka sitasi akan bertambah sebesar 1,2.

Model untuk variabel dependen indeks h

$$\begin{aligned}
 \hat{Q}_{0,05}(y | \mathbf{x}) &= -0,172 + 0,014Y_1 + 0,002X_1 - 0,003X_2 + 0,0425X_3 - 0,009X_4 \\
 &\quad + 0,031X_5 + 0,078X_{6(1)} + 0,045X_{6(2)} + 0,031X_{6(3)} - 0,019X_7 \\
 &\quad + 0,035X_8. \\
 \hat{Q}_{0,50}(y | \mathbf{x}) &= 0,752 + 0,020Y_1 - 0,004X_1 + 0,0045X_2 + 0,080X_3 + 0,008X_4 \\
 &\quad + 0,028X_5 + 0,047X_{6(1)} + 0,0098X_{6(2)} - 0,0045X_{6(3)} - 0,75X_7 \\
 &\quad + 0,0041X_8. \\
 \hat{Q}_{0,95}(y | \mathbf{x}) &= 1,420 + 0,039Y_1 + 0,009X_1 + 0,001X_2 + 0,078X_3 + 0,005X_4 \\
 &\quad - 0,254X_5 - 0,358X_{6(1)} - 0,312X_{6(2)} - 0,416X_{6(3)} - 0,160X_7 \\
 &\quad - 0,031X_8.
 \end{aligned} \tag{4.35}$$

Hasil estimasi pada Tabel 4.20, diketahui bahwa koefisien variabel sitasi dan jumlah dokumen berpengaruh positif terhadap indeks h pada kuantil 5%. Apabila jumlah dokumen bertambah sebanyak 10 dokumen, maka indeks h akan bertambah sebesar 0,425. Sedangkan untuk variabel sitasi, apabila terdapat 50 orang yang mengutip karya dosen, maka indeks h akan bertambah sebesar 1.

Gambar 4.27 berikut ini adalah regresi kuantil multivariabel yang menjelaskan indeks h dengan jumlah dokumen dengan syarat variabel independen lainnya tetap. Dari Gambar tersebut, dapat dijelaskan bahwa indeks h dengan variabel jumlah dokumen dengan syarat (a) dosen laki-laki sebagai Lektor Kepala (LK) yang memiliki sitasi sebanyak 5, berusia 45 tahun, lama bekerja 20 tahun, pendidikan terakhir yang ditempuh S3 di luar negeri (LN) dan memiliki jumlah *co-authors* sebanyak 5. Sedangkan syarat untuk (b) dosen perempuan sebagai Lektor Kepala (LK) yang memiliki sitasi sebanyak 5, berusia 45 tahun, lama bekerja 20 tahun, pendidikan terakhir yang ditempuh S3 di luar negeri (LN) dan memiliki jumlah *co-authors* sebanyak 5.



Gambar 4.27 Scatter Plot Regresi Kuantil untuk Indeks h dengan Jumlah Dokumen dengan Syarat X dimana dibedakan Berdasarkan Jenis Kelamin (a) Perempuan, (b) Laki-laki

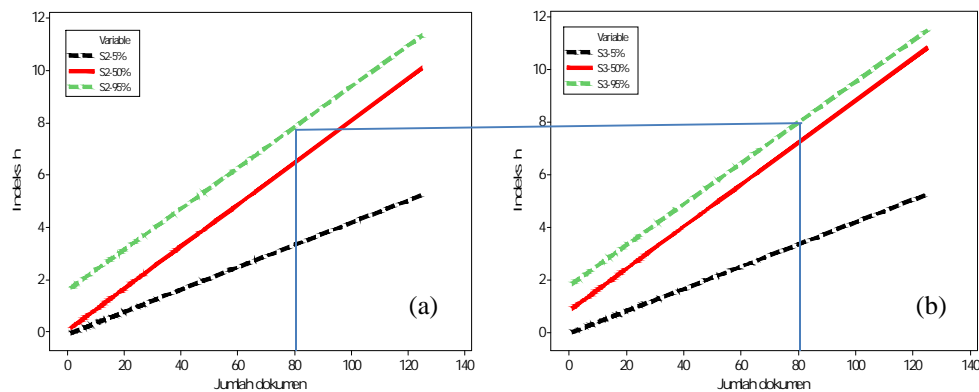
Nilai dugaan dari indeks h dengan jumlah dokumen dengan syarat X dapat dilihat pada Tabel 4.22. Terdapat perbedaan dugaan indeks h antara laki-laki

dengan perempuan dengan jumlah dokumen sebanyak 60 dan 80. Dari Tabel 4.22 diketahui bahwa dugaan indeks h dosen perempuan pada kuantil 5%. 50% dan 95% lebih tinggi dibandingkan dugaan indeks h dosen laki-laki.

Tabel 4.22 Hasil Dugaan Indeks h dengan Jumlah Dokumen Berdasarkan Jenis Kelamin

Jumlah dokumen	Laki-laki			Perempuan		
	5%	50%	95%	5%	50%	95%
60	2,493	5,628	6,435	2,524	5,656	6,181
80	3,344	7,233	7,993	3,375	7,261	7,739

Penjelasan mengenai indeks h dosen laki-laki dengan variabel jumlah dokumen dengan syarat X disajikan pada Gambar 4.28 berikut:



Gambar 4.28 Scatter Plot Regresi Kuantil untuk Indeks h dengan Jumlah Dokumen Dengan Syarat X dimana dibedakan Berdasarkan Pendidikan Terakhir Dosen (a) S3, (b) S2

Gambar 4.28 informasi mengenai indeks h dosen laki-laki dengan variabel jumlah dokumen berdasarkan perbedaan pendidikan terakhir dosen. *Scatter plot* yang disajikan merupakan plot regresi kuantil untuk hasil dugaan indeks h dengan syarat (a) Lektor Kepala (LK) yang memiliki sitasi sebanyak 5, berusia 45 tahun, lama bekerja 20 tahun, pendidikan terakhir yang ditempuh S3 di luar negeri (LN) dan memiliki jumlah *co-authors* sebanyak 5. Sedangkan syarat untuk (b) Lektor Kepala (LK) yang memiliki sitasi sebanyak 5, berusia 45

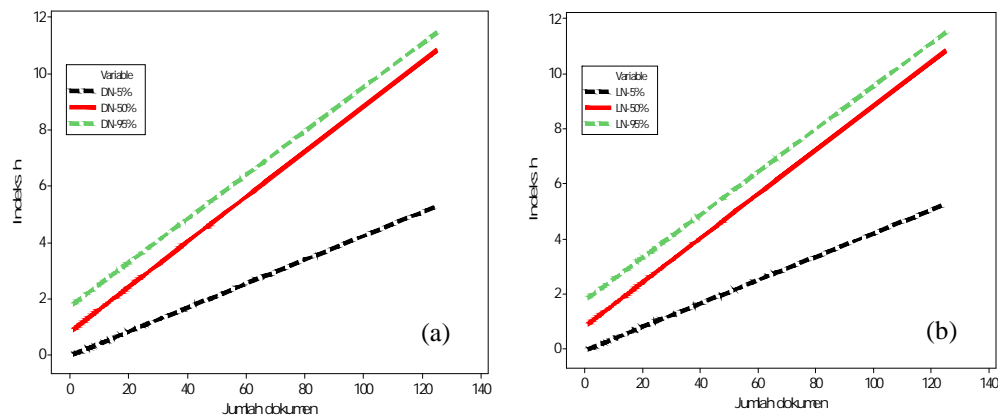
tahun, lama bekerja 20 tahun, pendidikan terakhir yang ditempuh S2 di luar negeri (LN). Secara visual terlihat adanya perbedaan pada kuantil 5%, 50% dan 95%. Perbedaan tersebut terlihat dari nilai dugaan untuk jumlah dokumen 60 dan 80 yang disajikan pada Tabel 4.23.

Tabel 4.23 Hasil Dugaan Indeks h dengan Jumlah Dokumen Berdasarkan Pendidikan Terakhir

Jumlah dokumen	S3			S2		
	5%	50%	95%	5%	50%	95%
60	2,493	5,628	6,435	2,474	4,877	6,275
80	3,344	7,233	7,993	3,325	6,482	7,834

Nilai dugaan dari indeks h dengan jumlah dokumen dengan syarat X diketahui terdapat perbedaan antara dosen laki-laki yang berpendidikan S3 dengan dosen laki-laki yang berpendidikan S2. Nilai dugaan dosen laki-laki dengan pendidikan S3 lebih tinggi daripada dosen dengan pendidikan S2.

Nilai dugaan indeks h berdasarkan tempat pendidikan terakhir dosen dapat dilihat pada Gambar 4.29.



Gambar 4.29 Scatter Plot Regresi Kuantil untuk Indeks h dengan Jumlah Dokumen Dengan Syarat X dimana dibedakan Berdasarkan Tempat Pendidikan Terakhir Dosen (a) DN, (b) LN

Hasil dugaan indeks h dosen laki-laki dengan pendidikan terakhir S3 di luar negeri untuk kuantil 5% dan 50% lebih rendah dibandingkan dengan dosen yang menempuh pendidikan terakhir di dalam negeri. Sedangkan untuk kuantil 95% dugaan indeks h dengan syarat X untuk dosen laki-laki dengan pendidikan terakhir S3 di luar negeri lebih tinggi dari pada dosen yang berpendidikan terakhir di dalam negeri. Hal tersebut dapat diketahui dari hasil Dugaan yang disajikan pada Tabel 4.24 berikut:

Tabel 4.24 Hasil Dugaan Indeks h dengan Jumlah Dokumen Berdasarkan Tempat Pendidikan Terakhir

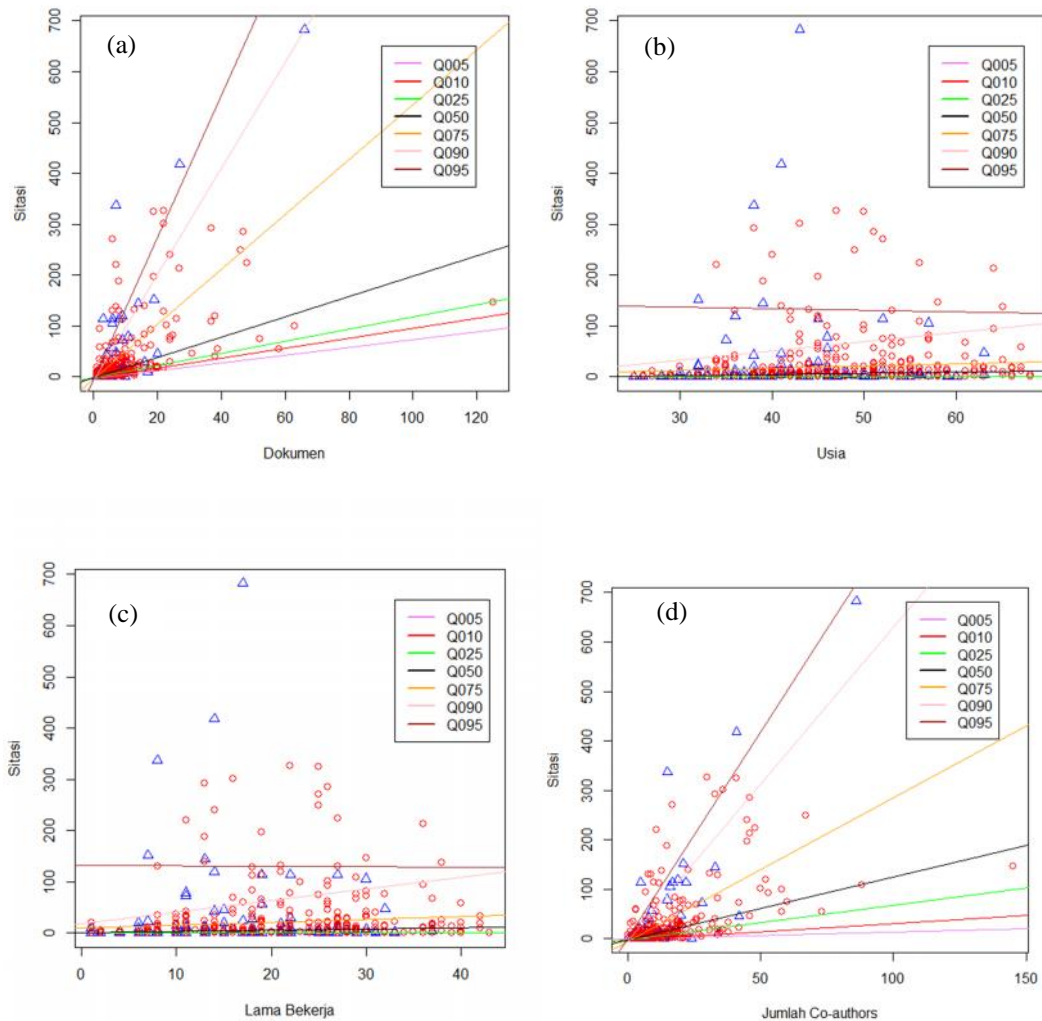
Jumlah dokumen	DN			LN		
	5%	50%	95%	5%	50%	95%
60	2,529	5,632	6,404	2,493	5,628	6,435
80	3,380	7,237	7,962	3,344	7,233	7,993

4.6.2 Pemodelan Publikasi Dosen ITS Menggunakan Regresi Kuantil Berbasis Model Rekursif Univariat untuk Variabel Independen Kontinu

Pemodelan regresi kuantil berbasis model rekursif univariat dilakukan hanya pada variabel independen kontinu, yaitu usia, lama bekerja, jumlah dokumen, dan jumlah *co-authors*. Hasil regresi kuantil secara univariat dapat dilihat pada Gambar 4.30. Garis regresi kuantil menggambarkan nilai dugaan variabel dependen, sementara varians mengukur penyimpangan observasi dari nilai dugaan. Pada regresi kuantil, *slope* yang sama antar kuantil merupakan indikator dari *homoscedasticity*. Hal ini dikarenakan kuantil merupakan sebagian dari variabel dependen. Pada *slope* yang sama antar kuantil, perbedaan nilai dugaan antar kuantil hanya diakibatkan oleh pergeseran rata-rata.

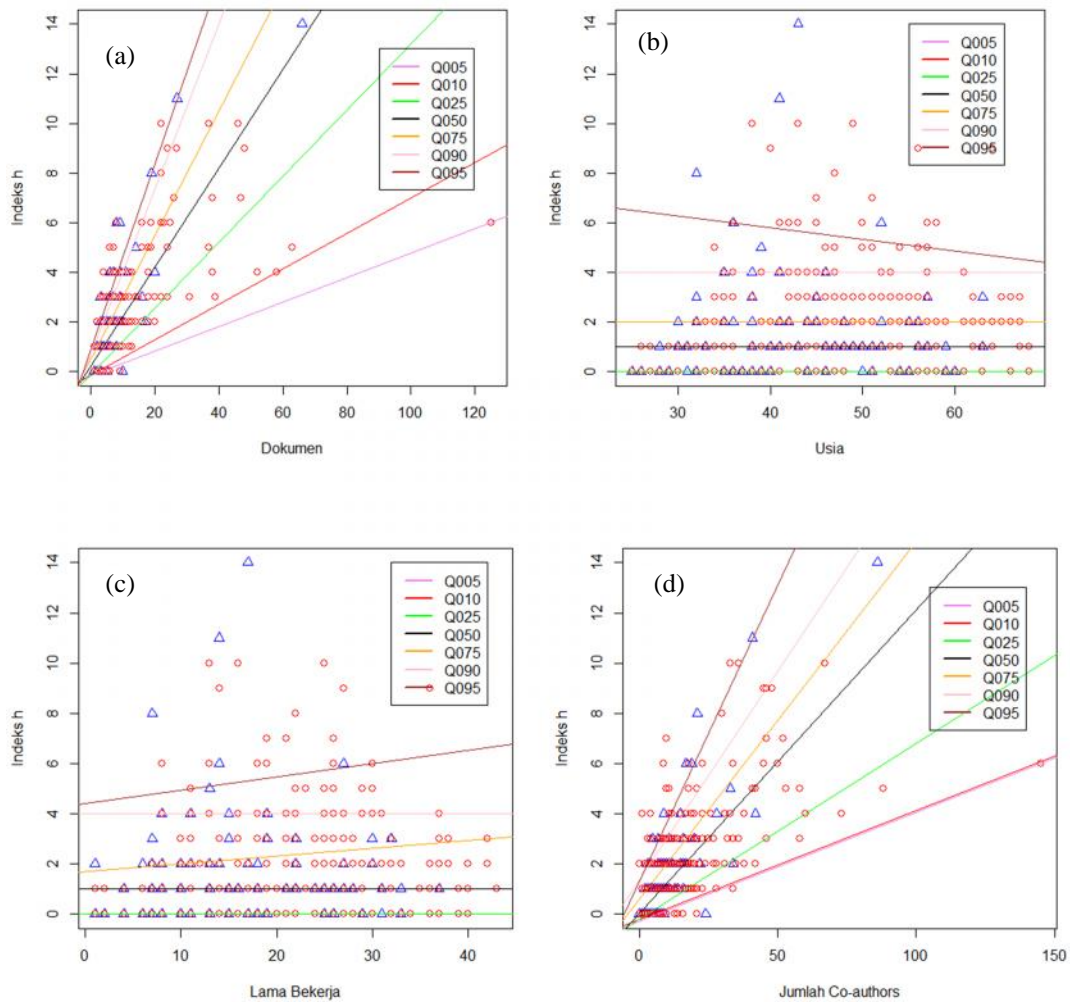
Hubungan sitasi dengan usia dan lama bekerja tidak berbanding lurus, terlihat dari plot data yang menyebar. Segitiga biru menunjukkan plot data untuk dosen perempuan, sedangkan plot merah untuk dosen laki-laki.

Perbandingan jumlah dosen perempuan yang memiliki akun di Scopus lebih sedikit daripada dosen laki-laki, namun sitasi dan indeks h tertinggi dimiliki oleh dosen perempuan.



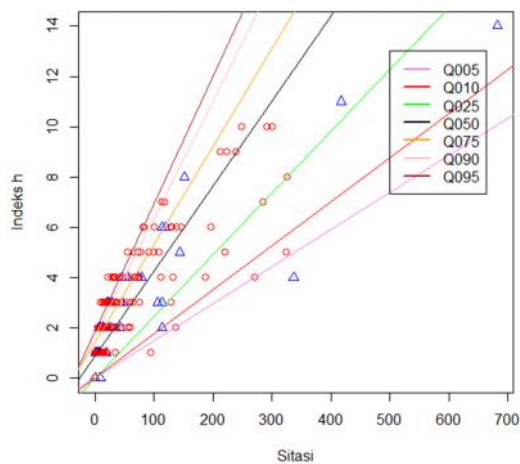
Gambar 4.30 Scatter Plot (a) Variabel Jumlah Dokumen dengan Sitasi, (b) Variabel Usia dengan Sitasi (c) Variabel Lama bekerja dengan Sitasi, (d) Variabel Jumlah *co-authors* dengan Sitasi. Simbol \triangle Merupakan Dosen Laki-laki, dan Simbol \circ Merupakan Simbol untuk Dosen Perempuan

Selain itu, semakin bertambahnya nilai sitasi maka jumlah dokumen dan jumlah *co-authors* selalu bertambah. Hal ini menunjukkan bahwa sitasi dan jumlah dokumen serta jumlah *co-authors* berbanding lurus.



Gambar 4.31 Scatter Plot (a) Variabel Jumlah Dokumen dengan Indeks h, (b) Variabel Usia dengan Indeks h, (c) Variabel Lama bekerja dengan Indeks h, (d) Variabel Jumlah *co-authors* dengan Indeks h. Simbol \triangle Merupakan Dosen Laki-laki, dan Simbol \circ Merupakan Simbol untuk Dosen Perempuan

Garis regresi kuantil univariat ini menunjukkan bahwa regresi kuantil dapat memodelkan data mengikuti pola sebaran data. Hal yang sama juga terjadi pada variabel indeks h seperti pada Gambar 4.31 berikut. Gambar 4.31 menjelaskan mengenai hubungan indeks h dengan variabel usia, lama bekerja, jumlah dokumen, dan jumlah *co-authors*. Dapat dikatakan bahwa indeks h berbandng lurus dengan jumlah dokumen, dan jumlah *co-authors*, tetapi tidak berbanding lurus dengan usia dan lama bekerja. Hubungan antara indeks h dengan jumlah dokumen dan jumlah *co-authors* berpengaruh terhadap signfikansi hasil estimasi parameter pada regresi kuantil seperti pada Tabel 4.19 dan 4.20 dalam sub bab 4.6.1.



Gambar 4.32 *Scatter Plot* Variabel Indeks h dengan Sitasi. Simbol \circ Merupakan Dosen Laki-laki, dan Simbol \triangle Merupakan Simbol untuk Dosen Perempuan

Hubungan antara indeks h sebagai variabel dependen dengan sitasi dapat dilihat pada Gambar 4.32 diatas. Garis regresi kuantil mengikuti pola dari sebaran data dimana semakin bertambah sitasi seorang dosen, maka indeks h juga akan bertambah.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Setelah dilakukan analisis dan penjelasan dari pengolahan data pada bagian sebelumnya. Kesimpulan dari hasil analisis adalah sebagai berikut.

1. Taksiran titik untuk parameter pada regresi kuantil adalah sebagai berikut:

$$\begin{aligned}\hat{S}(\dagger) &= \min_{S \in \mathbb{R}^p} \left\{ \dagger \sum_{i=1; v \geq 0}^n |y - \mathbf{X}^T(\dagger)| + (1-\dagger) \sum_{i=1; v < 0}^n |y - \mathbf{X}^T(\dagger)| \right\}, \\ &= \min_{S \in \mathbb{R}^p} \sum_{i=1}^n \dots(\dagger)(v_i).\end{aligned}$$

dimana $\dots(\dagger)(v_i)$ adalah *loss function* dengan bentuk sebagai berikut:

$$\dots(\dagger)(v_i) = \begin{cases} \dagger v_i & , v_i \geq 0 \\ (\dagger - 1)v_i & , v_i < 0. \end{cases}$$

dengan, $\hat{S}(\dagger)$ adalah penduga parameter, \dagger adalah Indeks kuantil, dan v_i adalah residual dari estimasi parameter.

Studi simulasi untuk residual yang heterogen dan homogen dengan ukuran sampel sebesar 300 serta $S_0 = 10$ dan $S_1 = 5$. Selain itu, studi simulasi ini juga membandingkan ukuran sampel dan ukuran varians dimana *error* yang dibangkitkan adalah heterogen. Hasil studi simulasi yang dilakukan:

- a. Untuk *error* homogen nilai *intercept* pada regresi kuantil akan selalu naik, sedangkan nilai *slope* sejajar. Untuk *error* heterogen, pada regresi kuantil baik nilai *intercept* dan *slope* akan selalu naik mengikuti pola data.
- b. Untuk studi simulasi pada *error* heterogen dengan perbandingan ukuran sampel diketahui bahwa perbedaan ukuran sampel tidak mempengaruhi hasil analisis.

- c. Untuk studi simulasi pada *error* heterogen dengan membandingkan ukuran varians yaitu 0.25, 1, dan 4, diperoleh informasi bahwa perbedaan varians tidak mempengaruhi hasil dari regresi kuantil.
2. Taksiran interval untuk parameter $s_j(\dagger)$ dengan metode *direct* fungsi sparsity adalah sebagai berikut:

$$P\left(\hat{S}_j(\dagger) - t_{(r/2, df)} se(\hat{S}_j(\dagger)) \leq s_j(\dagger) \leq \hat{S}_j(\dagger) + t_{(r/2, df)} se(\hat{S}_j(\dagger))\right) = 1 - r,$$

dengan F' adalah fungsi distribusi kumulatif dan $f = F'$, $\hat{S}_j(\dagger)$ adalah taksiran titik dari $s_j(\dagger)$, $se(\hat{S}_j(\dagger))$ adalah standart *error* dari $\hat{S}_j(\dagger)$ yang diperoleh dari diagonal $\check{S}^2(\dagger)\mathbf{D}^{-1}$, dimana $\mathbf{D}^{-1} = (\mathbf{X}^T\mathbf{X})^{-1}$ dan $\check{S}^2(\dagger) = \dagger(1-\dagger) / f(F^{-1}(\dagger))^2$.

3. Model regresi kuantil untuk publikasi dosen ITS diperoleh hasil :
- a. Hasil analisis regresi kuantil secara multivariabel diketahui bahwa variabel yang berpengaruh pada setiap kuantil berbeda. Jumlah dokumen berpengaruh terhadap sitasi dan indeks h hampir di setiap kuantil. Untuk perbandingan jenis kelamin, dapat diinformasikan bahwa sitasi dan indeks h dosen perempuan lebih tinggi daripada dosen laki-laki pada kuantil rendah. Namun, pada kuantil 90% dan 95% dosen laki-laki memiliki indeks h dan sitasi lebih tinggi daripada dosen perempuan. Untuk perbandingan pendidikan dosen laki-laki, dosen dengan pendidikan S3 selalu memiliki indeks h dan sitasi lebih tinggi daripada dosen dengan pendidikan terakhir S2. Sedangkan pada perbandingan tempat pendidikan terakhir dosen laki-laki, sitasi dan indeks h dosen yang menyelesaikan studi di LN pada kuantil 90 dan 95% lebih tinggi daripada dosen yang menyelesaikan studi DN.
- b. Hasil analisis regresi kuantil secara univariabel diketahui bahwa sitasi selalu berbanding lurus dengan jumlah dokumen dan jumlah *co-authors*. Hal ini juga terjadi pada variabel indeks h. Hasil estimasi parameter regresi kuantil mengikuti pola data heterogen, sehingga metode ini dapat memodelkan sesuai sebaran data.

5.2 Saran

Saran yang bisa disampaikan berdasarkan hasil penelitian ini adalah, untuk pihak ITS diharapkan dapat memberikan fasilitas pelatihan supaya keilmuan dan ketrampilan dosen dalam hal menulis dokumen semakin meningkat. Selain itu ITS diharapkan memberi fasilitas berupa insentif sehingga dosen dapat melakukan penelitian dan menulis karya ilmiah. Dengan adanya fasilitas yang diberikan oleh ITS diharapkan dosen dapat meningkatkan jumlah dokumen khususnya dalam publikasi pada Scopus dan melakukan kolaborasi riset baik dengan pihak dalam negeri maupun dalam negeri, sehingga dapat meningkatkan sitasi dan indeks h.

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Archambault, E., Campbell, D., Gingras, Y., & Lariviere, V. (2009). "Comparing Bibliometric Statistics Obtained From the Web of Science and Scopus". *Journal of The American Society for Information Science and Technology*. Vol. 60, No.7, hal. 1320-1326.
- Buchinsky, M. (1994). "Change in the U.S. Wage Structure 1963-1987: Application of Quantile Regression". *Econometrica*. Vol. 62, hal. 405-458.
- Chen, C., dan Wei, Y. (2005). "Computational Issues for Quantile Regression". *The Indian Journal of Statistics*. Vol. 67, No. 2, hal. 399-417.
- Davino, C., Furno, M. (1998). "*Quantile Regression Theory and Application*". SPI Publishers Services, Pondicherry, India.
- Departemen Keuangan (2009). Peraturan yang Mengatur Tentang Dosen. <http://www.sjdih.depkeu.go.id/fulltext/2009/37PP.html>, Diakses pada 31 Agustus 2016.
- Eide, E., dan Showalter, M. H. (1998). "The Effect of School Quality on student Performance : A Quantile Regression Approach". *Economics*. Vol. 58, hal. 345-350.
- Goh, S. C., dan Knight, K. (2009). "Nonstandart Quantile Regression Inference". *Econometric Theory*. Vol. 25, hal. 1415-1432.
- Gujarati, D.N. (2004). *Basic Econometrics*, 4th edition. United States of America: The McGraw-Hill Companies.
- Hapsery, A. (2015). Pemodelan Faktor-faktor yang Mempengaruhi Kinerja Dosen ITS di *Google Scholar Citation* Menggunakan Model Rekursif. Tugas Akhir: Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Hemmings, B., dan Kay, R. (2010a). "University Lecturer Publication Output: Qualification, Time and Confidence Count". *Journal of Higher Education Policy and Management*. Vol. 32, No.2, hal. 185-197.

- Hemmings, B., dan Kay, R. (2010b). "Journal Ratings and the Publications of Australian Academics". *Issues in Educational Research*. Vol 20, No.3, hal. 234-243.
- Hirsch, J.E. (2007). "Does the h Index Have Predictive Power". *PNAS* 104, No.49, hal. 19193-19198.
- Hirsch, J.E. (2005). "An Index to Quantify An Individual's Scientific Research Output". *PNAS* 102, No. 46, hal. 16569-16572.
- John, O. O., dan Nduka, E.C. (2009). "Quantile Regression Analysis as a Robust Alternative to Ordinary Least Squares", *The Journal of Scientia Africana*, Vol. 8, No.2, hal. 61-65. Published by: University of Port Harcourt.
- Jenkins, N.R. (2015). "Variation in the h-Index and Use in the Assessment of Academic Output". <http://dx.doi.org/10.1016/j.wneu.2015.09041>. Diakses pada 31 Agustus 2016.
- Koenker, R. (2005). *Quantile Regression*. First published. Cambridge University Press.
- Koenker, R., dan Bassett, G. (1978). "Regression quantiles". *Econometrica*, hal. 33-50.
- Koenker, R., dan Hallock, K. F. (2001). "Quantile Regression". *Journal of Economic Perspective*. Vol. 15, No.4, hal. 143-156. Published by: American Economic Association.
- Koenker, R., dan D'Orey, V. (1993). "Computing Regression Quantiles". *J. Roy. Statist. Soc. Ser.C*. Vol. 43, hal. 410-414.
- Lei, Y., Tan, B. J., Zou, Z., Zhang, M. M., Song, R. P., Qu, S. H., et al. (2014). "Publication Patterns and Citation Analysis of APJTM during 2008 and June 2014". *Asian Pacific Journal of Tropical Medicine*, hal. 650-654.
- Navianti, D. R. (2014). Regresi Kuantil untuk Pemodelan Tingkat Pengangguran Terbuka di Indonesia. Tesis : Jurusan Statistika, Institut Teknologi Sepuluh Nopember.

- Noor, N. dan Muhammad, A., (2013). “Model of Robust Regression with Parametric and Nonparametric Methods”. *Mathematical Theory and Modeling*. Vol. 3, No. 5, hal. 27-39.
- Margaretha, M., dan Saragih, S. (2012). Faktor-faktor Penentu Produktivitas Penelitian Dosen Sebagai Implementasi Integritas Profesi. Tugas Akhir: Fakultas Ekonomi, Universitas Kristen Marantha 1, No,3, hal. 195-208.
- Pratt, M., Margaritis, D., dan Coy, D. (1999). “Developing a Research Culture in a University Faculty”. *Journal of Higher Education Policy and Management* . Vol. i21, No.1, hal. 43-55.
- Rahmawati, D. N. (2016). Pemodelan terhadap Faktor-Faktor yang Mempengaruhi Publikasi Dosen ITS di Scopus. Tugas Akhir : Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Smith, D., Gaglianone, W. P., Lima, L. R., Linton, O. (2009). “Evaluating Value-at Risk Models Via Quantile Regression”. *Working Paper Economic Series*. Vol. 25, hal 09-46.
- Suhartono, Rahmawati, D., Atok, R.M., Prastyo, D.D., dan Ahmad, I.S., (2016), “Determinant Factors of Publication in Scopus at ITS”, *Proceeding of 12-th International Conference on Mathematics, Statistics, and their Applications (ICMSA) 2016*, Banda Aceh.
- Sulistiyawati, V. K. (2016). Pemodelan Kepemilikan Publikasi, Jumlah Kutipan dan Indeks h Dosen ITS di Scopus. Tugas Akhir : Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Stack, S. (2004). “Gender, Children and Research Productivity”. *Research in Higher Education*. Vol. 45. No.8, hal. 891-920.
- Sax, L.J., Hagedorn, L.S., Arredondo, M., dan Dicrisi, F.A. (2002). “Faculty Research Productivity: Exploring the Role of Gender and Family-Related Factors”. *Research in Higher Education*. Vol. 43, No.4, hal. 423-446.
- Soekanto, S. (2014). *Pengertian Penelitian*. <http://PengertianPenelitianMenurut-ParaAhliSeputarPengetahuan.Com>. Diakses pada 31 Agustus 2016.

- Tien, F.F. (2000). "To What Degree Does the Desire for Promotion Motivate Faculty to Perform Research? Testing the Expectancy Theory". *Research in Higher Education*. Vol. 41, No,6, hal. 723-752.
- Uthami, I.A.P., Sukarsa, I. K. G., dan Kencana, I. P. K. (2013). "Regresi Kuantil Median untuk Mengatasi Heteroskedastisitas pada Analisis Regresi". *Jurnal Matematika*. Vol. 2. No. 1, hal. 6-13.
- Wahyudi, V. E. (2015). Analisis IPM di Pulau Jawa Menggunakan Analisis Regresi Kuantil. Tesis : Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Walpole, R.E. (1995). *Pengantar Metode Statistika* (3rd ed). (B. Sumantri, Penerj.) Jakarta:PT. Gramedia Putaka Utama.
- Wardani, D. K. (2014). Penerapan Regresi Kuantil pada Data dengan Pelanggaran Asumsi Kenormalaan Sisaan. Tugas Akhir : Jurusan Matematika, Program studi Statistika, Universitas Brawijaya.
- Zhou, K. Q., dan Portnoy, S. L. (1996). "Direct Use of Regression Quantiles to Construct Confidence Sets in Linier Models". *The Annals of Statistics. Value in Health*. Vol. 24, No. 1, hal. 287-306.

LAMPIRAN

Lampiran I. Data Publikasi Dosen ITS di Scopus.

Nama Dosen	Y1	Y2	X1	X2	X3	X4	X5.1	X6.1	X6.2	X6.3	X7.1	X8.1
Prof. Dr. Suasmoro	68	4	61	37	10	23	0	1	0	0	0	0
Prof. Drs. Suminar Pratapa, Ph.D.	101	6	50	26	23	50	0	1	0	0	0	0
Prof.Dr. Darminto, M.Sc	109	5	56	29	37	88	0	1	0	0	0	1
Dr.rer.nat. Triwikantoro, M.Sc.	92	5	50	26	18	53	0	0	1	0	0	0
Dr. Yono Hadi Pramono, M.Eng.	39	2	47	24	7	14	0	0	1	0	0	0
Dr. Mashuri, S.Si, M.Si	8	1	47	22	2	14	0	0	1	0	0	1
Prof.Dr. Bagus Jaya Santosa, S.U.	6	2	54	29	9	12	0	1	0	0	0	0
Drs. Zaenal Arifin, M.Si	1	1	51	25	2	6	0	0	0	1	1	1
Dr. M. Zainuri, M.Si	3	1	52	26	8	17	0	0	0	1	0	1
Drs. Yoyok Cahyono, M.Si.	6	1	53	26	1	5	0	0	0	1	1	1
Drs. M.Zainul Asrori, M.Si	7	1	51	25	1	7	0	0	0	1	1	1
Drs Agus Purwanto, M.Sc, D.Sc.	45	3	52	26	6	7	0	0	1	0	0	0
Dr.Dra Melania Suweni Muntini, MT	3	1	52	26	2	6	1	0	0	1	0	1
Malik Anjelh Baqiya, S.Si., M.Si.,Ph.D.	14	2	34	8	11	38	0	0	0	0	0	0
Sungkono, S.Si, M.Si	3	1	31	2	4	8	0	0	0	0	1	1
Prof. Dr. Basuki Widodo, MSc.	17	2	51	27	20	31	0	1	0	0	0	0
Dr. Drs. Chairul Imron, Mi. Komp	0	0	55	29	3	10	0	0	1	0	0	1
Dr. Subiono, M.Sc	15	2	59	32	4	3	0	0	1	0	0	0
Drs. Soehardjoepri, M.Si	0	0	54	29	1	3	0	0	1	0	1	1
Prof. Dr. Erna Apriliani, M.Si.	1	1	50	25	3	8	1	1	0	0	0	1
Dr. Drs. Hariyanto, M.Si.	0	0	63	34	1	8	0	0	1	0	0	1
Drs. Kamiran, M.Si	0	0	53	27	1	4	0	0	1	0	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Hatma Suryotrisongko, M.Eng	0	0	32	2	3	4	0	0	0	0	1	0
Rully Agus Hendrawan, M.Eng.	0	0	35	11	5	6	0	0	0	1	1	0
Amalia Utamima, S.Kom., M.BA.	10	2	30	1	3	3	1	0	0	0	1	0
Anisah Herdiyanti Prabowo, M.Sc.	0	0	29	1	1	2	1	0	0	0	1	0
Amna Shifia Nisafani, S.Kom, M.Sc	0	0	29	1	2	5	1	0	0	0	1	0

Lampiran II. Syntax Software R.

1. Syntax Studi Simulasi untuk Residual Homogen.

```
library(quantreg)
n = 300
r=1000
#regresi ols
b0 = matrix (ncol= 1, nrow=r)
b1 = matrix (ncol= 1, nrow=r)
pb0 = matrix (ncol= 1, nrow=r)
pb1 = matrix (ncol= 1, nrow=r)
p0 = matrix (ncol= 1, nrow=r)
p1 = matrix (ncol= 1, nrow=r)
#regresi kuantil 5%
b0q5 =matrix(ncol=1, nrow=r)
b1q5 =matrix(ncol=1, nrow=r)
pb0q5= matrix(ncol=1, nrow=r)
pb1q5=matrix(ncol=1, nrow=r)
p0q5=matrix(ncol=1, nrow=r)
p1q5=matrix(ncol=1, nrow=r)
#regresi kuantil 10%
b0q10 =matrix(ncol=1, nrow=r)
b1q10 =matrix(ncol=1, nrow=r)
pb0q10 = matrix(ncol=1, nrow=r)
pb1q10 = matrix(ncol=1, nrow=r)
p0q10 = matrix(ncol=1, nrow=r)
p1q10 = matrix(ncol=1, nrow=r)
#regresi kuantil 25%
b0q25 =matrix(ncol=1, nrow=r)
b1q25 =matrix(ncol=1, nrow=r)
pb0q25 = matrix(ncol=1, nrow=r)
pb1q25 = matrix(ncol=1, nrow=r)
p0q25 = matrix(ncol=1, nrow=r)
p1q25 = matrix(ncol=1, nrow=r)
#regresi kuantil 50%
b0q50 =matrix(ncol=1, nrow=r)
b1q50 =matrix(ncol=1, nrow=r)
pb0q50 = matrix(ncol=1, nrow=r)
pb1q50= matrix(ncol=1, nrow=r)
p0q50 = matrix(ncol=1, nrow=r)
p1q50 = matrix(ncol=1, nrow=r)
#regresi kuantil 75%
b0q75 =matrix(ncol=1, nrow=r)
b1q75 =matrix(ncol=1, nrow=r)
pb0q75 = matrix(ncol=1, nrow=r)
pb1q75 = matrix(ncol=1, nrow=r)
p0q75 = matrix(ncol=1, nrow=r)
p1q75 = matrix(ncol=1, nrow=r)
#regresi kuantil 90%
b0q90 =matrix(ncol=1, nrow=r)
b1q90 =matrix(ncol=1, nrow=r)
pb0q90 = matrix(ncol=1, nrow=r)
pb1q90 = matrix(ncol=1, nrow=r)
p0q90 = matrix(ncol=1, nrow=r)
p1q90 = matrix(ncol=1, nrow=r)
#regresi kuantil 95%
b0q95 =matrix(ncol=1, nrow=r)
b1q95 =matrix(ncol=1, nrow=r)
pb0q95 = matrix(ncol=1, nrow=r)
pb1q95 = matrix(ncol=1, nrow=r)
p0q95 = matrix(ncol=1, nrow=r)
p1q95 = matrix(ncol=1, nrow=r)
#pengulangan
for (i in 1:r)
{
x = rnorm(n,0,10)
e = rnorm(n,0,5)
b00 = 10
b10 = 5
y = b00 + (b10 * x) + e
xy = cbind(x,y)
#kedua regresikan
regression = lm(y~ x)
hasil = summary(regression)
b0[i]=hasil$coef[1,1]
b1[i]=hasil$coef[2,1]
pb0[i]=hasil$coef[1,4]
pb1[i]=hasil$coef[2,4]
if (pb0<0.05) p0[i]=1 else p0[i]=0
if (pb1<0.05) p1[i]=1 else p1[i]=0
#regresi kuantil
k1=rq(y~x, tau=c(0.01))
regq1=summary(k1, se="boot")
k2.5=rq(y~x, tau=c(0.025))
regq2.5=summary(k2.5, se="boot")
k5=rq(y~x, tau=c(0.05))
regq5=summary(k5, se="boot")
k10=rq(y~x, tau=c(0.10))
regq10=summary(k10, se="boot")
k25=rq(y~x, tau=c(0.25))
regq25=summary(k25, se="boot")
k50=rq(y~x, tau=c(0.50))
regq50=summary(k50, se="boot")
k75=rq(y~x, tau=c(0.75))
regq75=summary(k75, se="boot")
k90=rq(y~x, tau=c(0.90))
regq90=summary(k90, se="boot")
k95=rq(y~x, tau=c(0.95))
regq95=summary(k95, se="boot")
#regresi kuantil 5% pvalue
b0q5[i]=regq5$coef[1,1]
b1q5[i]=regq5$coef[2,1]
pb0q5[i]=regq5$coef[1,4]
pb1q5[i]=regq5$coef[2,4]
if (pb0q5<0.05) p0q5[i]=1 else p0q5[i]=0
if (pb1q5<0.05) p1q5[i]=1 else p1q5[i]=0
#regresi kuantil 10% pvalue
b0q10[i]=regq10$coef[1,1]
b1q10[i]=regq10$coef[2,1]
pb0q10[i]=regq10$coef[1,4]
pb1q10[i]=regq10$coef[2,4]
if (pb0q10<0.05) p0q10[i]=1 else p0q10[i]=0
if (pb1q10<0.05) p1q10[i]=1 else p1q10[i]=0
#regresi kuantil 25% pvalue
b0q25[i]=regq25$coef[1,1]
b1q25[i]=regq25$coef[2,1]
pb0q25[i]=regq25$coef[1,4]
pb1q25[i]=regq25$coef[2,4]
if (pb0q25<0.05) p0q25[i]=1 else p0q25[i]=0
if (pb1q25<0.05) p1q25[i]=1 else p1q25[i]=0
#regresi kuantil 50% pvalue
b0q50[i]=regq50$coef[1,1]
b1q50[i]=regq50$coef[2,1]
pb0q50[i]=regq50$coef[1,4]
pb1q50[i]=regq50$coef[2,4]
if (pb0q50<0.05) p0q50[i]=1 else p0q50[i]=0
if (pb1q50<0.05) p1q50[i]=1 else p1q50[i]=0
#regresi kuantil 75% pvalue
b0q75[i]=regq75$coef[1,1]
b1q75[i]=regq75$coef[2,1]
pb0q75[i]=regq75$coef[1,4]
pb1q75[i]=regq75$coef[2,4]
if (pb0q75<0.05) p0q75[i]=1 else p0q75[i]=0
if (pb1q75<0.05) p1q75[i]=1 else p1q75[i]=0
}
```

Lanjutan : Syntax Studi Simulasi untuk Residual Homogen.

```
#regresi kuantil 90% pvalue
b0q90[i]=regq90$coef[1,1]
b1q90[i]=regq90$coef[2,1]
pb0q90[i]=regq90$coef[1,4]
pb1q90[i]=regq90$coef[2,4]
if (pb0q90<0.05) p0q90[i]=1 else p0q90[i]=0
if (pb1q90<0.05) p1q90[i]=1 else p1q90[i]=0
#regresi kuantil 95% pvalue
b0q95[i]=regq95$coef[1,1]
b1q95[i]=regq95$coef[2,1]
pb0q95[i]=regq95$coef[1,4]
pb1q95[i]=regq95$coef[2,4]
if (pb0q95<0.05) p0q95[i]=1 else p0q95[i]=0
if (pb1q95<0.05) p1q95[i]=1 else p1q95[i]=0

hitung regresi
b0rata= mean(b0)
b1rata= mean(b1)
akurasip0=sum(p0/r)
akurasip1=sum(p1/r)
b0sort=sort(b0)
b1sort=sort(b1)

#hitung kuantil
#rata q5
b0q5rata= mean(b0q5)
b1q5rata= mean(b1q5)
akurasip0q5=sum(p0q5/r)
akurasip1q5=sum(p1q5/r)
b0q5sort=sort(b0q5)
b1q5sort=sort(b1q5)

#rata q10
b0q10rata= mean(b0q10)
b1q10rata= mean(b1q10)
akurasip0q10=sum(p0q10/r)
akurasip1q10=sum(p1q10/r)
b0q10sort=sort(b0q10)
b1q10sort=sort(b1q10)
#rata q25
b0q25rata= mean(b0q25)
b1q25rata= mean(b1q25)
akurasip0q25=sum(p0q25/r)
akurasip1q25=sum(p1q25/r)
b0q25sort=sort(b0q25)
b1q25sort=sort(b1q25)
#rata q50
b0q50rata= mean(b0q50)
b1q50rata= mean(b1q50)
akurasip0q50=sum(p0q50/r)
akurasip1q50=sum(p1q50/r)
b0q50sort=sort(b0q50)
b1q50sort=sort(b1q50)
#rata q75
b0q75rata= mean(b0q75)
b1q75rata= mean(b1q75)
akurasip0q75=sum(p0q75/r)
akurasip1q75=sum(p1q75/r)
b0q75sort=sort(b0q75)
b1q75sort=sort(b1q75)
#rata q90
b0q90rata= mean(b0q90)
b1q90rata= mean(b1q90)
akurasip0q90=sum(p0q90/r)
akurasip1q90=sum(p1q90/r)
b0q90sort=sort(b0q90)
b1q90sort=sort(b1q90)

#rata q95
b0q95rata= mean(b0q95)
b1q95rata= mean(b1q95)
akurasip0q95=sum(p0q95/r)
akurasip1q95=sum(p1q95/r)
b0q95sort=sort(b0q95)
b1q95sort=sort(b1q95)
```

2. Syntax Studi Simulasi untuk Residual Heterogen.

```

library(quantreg)
n = 300
r=1000
#regresi ols
b0 = matrix(ncol= 1, nrow=r)
b1 = matrix(ncol= 1, nrow=r)
pb0 = matrix(ncol= 1, nrow=r)
pb1 = matrix(ncol= 1, nrow=r)
p0 = matrix(ncol= 1, nrow=r)
p1 = matrix(ncol= 1, nrow=r)
#regresi kuantil 5%
b0q5 =matrix(ncol=1, nrow=r)
b1q5 =matrix(ncol=1, nrow=r)
pb0q5= matrix(ncol=1, nrow=r)
pb1q5=matrix(ncol=1, nrow=r)
p0q5=matrix(ncol=1, nrow=r)
p1q5=matrix(ncol=1, nrow=r)
#regresi kuantil 10%
b0q10 =matrix(ncol=1, nrow=r)
b1q10 =matrix(ncol=1, nrow=r)
pb0q10 = matrix(ncol=1, nrow=r)
pb1q10 = matrix(ncol=1, nrow=r)
p0q10 = matrix(ncol=1, nrow=r)
p1q10 = matrix(ncol=1, nrow=r)
#regresi kuantil 25%
b0q25 =matrix(ncol=1, nrow=r)
b1q25 =matrix(ncol=1, nrow=r)
pb0q25 = matrix(ncol=1, nrow=r)
pb1q25 = matrix(ncol=1, nrow=r)
p0q25 = matrix(ncol=1, nrow=r)
p1q25 = matrix(ncol=1, nrow=r)
#regresi kuantil 50%
b0q50 =matrix(ncol=1, nrow=r)
b1q50 =matrix(ncol=1, nrow=r)
pb0q50 = matrix(ncol=1, nrow=r)
pb1q50= matrix(ncol=1, nrow=r)
p0q50 = matrix(ncol=1, nrow=r)
p1q50 = matrix(ncol=1, nrow=r)
#regresi kuantil 75%
b0q75 =matrix(ncol=1, nrow=r)
b1q75 =matrix(ncol=1, nrow=r)
pb0q75 = matrix(ncol=1, nrow=r)
pb1q75 = matrix(ncol=1, nrow=r)
p0q75 = matrix(ncol=1, nrow=r)
p1q75 = matrix(ncol=1, nrow=r)
#regresi kuantil 90%
b0q90 =matrix(ncol=1, nrow=r)
b1q90 =matrix(ncol=1, nrow=r)
pb0q90 = matrix(ncol=1, nrow=r)
pb1q90 = matrix(ncol=1, nrow=r)
p0q90 = matrix(ncol=1, nrow=r)
p1q90 = matrix(ncol=1, nrow=r)
#regresi kuantil 95%
b0q95 =matrix(ncol=1, nrow=r)
b1q95 =matrix(ncol=1, nrow=r)
pb0q95 = matrix(ncol=1, nrow=r)
pb1q95 = matrix(ncol=1, nrow=r)
p0q95 = matrix(ncol=1, nrow=r)
p1q95 = matrix(ncol=1, nrow=r)

#pengulangan
for (i in 1:r)
{
x = runif(n,0,5)
e = rmorm(n,0,1)
s = exp(0.5*x)
ee = s*e
b00 = 10
b10 = 5
y = b00 + (b10 * x) + ee
xy = cbind(x,y)

#kedua regresikan
regression = lm(y~ x)
hasil = summary(regression)
b0[i]=hasil$coef[1,1]
b1[i]=hasil$coef[2,1]
pb0[i]=hasil$coef[1,4]
pb1[i]=hasil$coef[2,4]
if (pb0<0.05) p0[i]=1 else p0[i]=0
if (pb1<0.05) p1[i]=1 else p1[i]=0
#regresi kuantil
k1=rq(y~x, tau=c(0.01))
regq1=summary(k1, se="boot")
k2.5=rq(y~x, tau=c(0.025))
regq2.5=summary(k2.5, se="boot")
k5=rq(y~x, tau=c(0.05))
regq5=summary(k5, se="boot")
k10=rq(y~x, tau=c(0.10))
regq10=summary(k10, se="boot")
k25=rq(y~x, tau=c(0.25))
regq25=summary(k25, se="boot")
k50=rq(y~x, tau=c(0.50))
regq50=summary(k50, se="boot")
k75=rq(y~x, tau=c(0.75))
regq75=summary(k75, se="boot")
k90=rq(y~x, tau=c(0.90))
regq90=summary(k90, se="boot")
k95=rq(y~x, tau=c(0.95))
regq95=summary(k95, se="boot")

#regresi kuantil 5% pvalue
b0q5[i]=regq5$coef[1,1]
b1q5[i]=regq5$coef[2,1]
pb0q5[i]=regq5$coef[1,4]
pb1q5[i]=regq5$coef[2,4]
if (pb0q5<0.05) p0q5[i]=1 else p0q5[i]=0
if (pb1q5<0.05) p1q5[i]=1 else p1q5[i]=0
#regresi kuantil 10% pvalue
b0q10[i]=regq10$coef[1,1]
b1q10[i]=regq10$coef[2,1]
pb0q10[i]=regq10$coef[1,4]
pb1q10[i]=regq10$coef[2,4]
if (pb0q10<0.05) p0q10[i]=1 else p0q10[i]=0
if (pb1q10<0.05) p1q10[i]=1 else p1q10[i]=0
#regresi kuantil 25% pvalue
b0q25[i]=regq25$coef[1,1]
b1q25[i]=regq25$coef[2,1]
pb0q25[i]=regq25$coef[1,4]
pb1q25[i]=regq25$coef[2,4]
if (pb0q25<0.05) p0q25[i]=1 else p0q25[i]=0
if (pb1q25<0.05) p1q25[i]=1 else p1q25[i]=0

#regresi kuantil 50% pvalue
b0q50[i]=regq50$coef[1,1]
b1q50[i]=regq50$coef[2,1]
pb0q50[i]=regq50$coef[1,4]
pb1q50[i]=regq50$coef[2,4]
if (pb0q50<0.05) p0q50[i]=1 else p0q50[i]=0
if (pb1q50<0.05) p1q50[i]=1 else p1q50[i]=0
#regresi kuantil 75% pvalue
b0q75[i]=regq75$coef[1,1]
b1q75[i]=regq75$coef[2,1]
pb0q75[i]=regq75$coef[1,4]
pb1q75[i]=regq75$coef[2,4]
if (pb0q75<0.05) p0q75[i]=1 else p0q75[i]=0
if (pb1q75<0.05) p1q75[i]=1 else p1q75[i]=0
#regresi kuantil 90% pvalue
b0q90[i]=regq90$coef[1,1]
b1q90[i]=regq90$coef[2,1]
pb0q90[i]=regq90$coef[1,4]
pb1q90[i]=regq90$coef[2,4]
if (pb0q90<0.05) p0q90[i]=1 else p0q90[i]=0
if (pb1q90<0.05) p1q90[i]=1 else p1q90[i]=0

```


Lanjutan : Syntax Studi Simulasi Residual Heterogen.

```
#regresi kuantil 95% pvalue
b0q95[i]=regq95$coef[1,1]
b1q95[i]=regq95$coef[2,1]
pb0q95[i]=regq95$coef[1,4]
pb1q95[i]=regq95$coef[2,4]
if (pb0q95<0.05) p0q95[i]=1 else p0q95[i]=0
if (pb1q95<0.05) p1q95[i]=1 else p1q95[i]=0
}
#hitung regresi
b0rata= mean(b0)
b1rata= mean(b1)
akurasip0=sum(p0/r)
akurasip1=sum(p1/r)
b0sort=sort(b0)
b1sort=sort(b1)
#hitung kuantil
#rata q5
b0q5rata= mean(b0q5)
b1q5rata= mean(b1q5)
akurasip0q5=sum(p0q5/r)
akurasip1q5=sum(p1q5/r)
b0q5sort=sort(b0q5)
b1q5sort=sort(b1q5)
#rata q10
b0q10rata= mean(b0q10)
b1q10rata= mean(b1q10)
akurasip0q10=sum(p0q10/r)
akurasip1q10=sum(p1q10/r)
b0q10sort=sort(b0q10)
b1q10sort=sort(b1q10)
#rata q25
b0q25rata= mean(b0q25)
b1q25rata= mean(b1q25)
akurasip0q25=sum(p0q25/r)
akurasip1q25=sum(p1q25/r)
b0q25sort=sort(b0q25)
b1q25sort=sort(b1q25)
#rata q50
b0q50rata= mean(b0q50)
b1q50rata= mean(b1q50)
akurasip0q50=sum(p0q50/r)
akurasip1q50=sum(p1q50/r)
b0q50sort=sort(b0q50)
b1q50sort=sort(b1q50)
#rata q75
b0q75rata= mean(b0q75)
b1q75rata= mean(b1q75)
akurasip0q75=sum(p0q75/r)
akurasip1q75=sum(p1q75/r)
b0q75sort=sort(b0q75)
b1q75sort=sort(b1q75)
#rata q90
b0q90rata= mean(b0q90)
b1q90rata= mean(b1q90)
akurasip0q90=sum(p0q90/r)
akurasip1q90=sum(p1q90/r)
b0q90sort=sort(b0q90)
b1q90sort=sort(b1q90)
#rata q95
b0q95rata= mean(b0q95)
b1q95rata= mean(b1q95)
akurasip0q95=sum(p0q95/r)
akurasip1q95=sum(p1q95/r)
b0q95sort=sort(b0q95)
b1q95sort=sort(b1q95)
#simpan file
write.csv(hasil,"E:hasil.csv")
write.csv(hasilqb0,"E:hasilqb0.csv")
write.csv(hasilqb1,"E:hasilqb1.csv")
write.csv(hasilqsort,"E:hasilqsort.csv")
write.csv(hasilqakurasi,"E:hasilqakurasi.csv")
write.table(xy,"E:xy1.xls")
```

3. Syntax Regresi Linier Berbasis Model Rekursif.

```

#regresi rekursif variabel sitasi
data=read.csv("E:/ALFI/Thesis me/8. Olah
Data/Olah.csv",sep=",")
#regresi ols sitasi
regresi = lm(Y1~
X1+X2+X3+X4+X5.1+X6.1+X6.2+X6.3+X7.1+X8.1,
data=data)
summary (regresi)
anova(regresi)
summary.aov(regresi)
confint(regresi)
fitted(regresi)

#regresi ols indeks h
regresi_h = lm(Y2~
Y1+X1+X2+X3+X4+X5.1+X6.1+X6.2+X6.3+X7.1+X8.1,
data=data)
summary (regresi_h)
anova(regresi_h)

#plot masing-masing
Sitasi=data$Y1
Usia=data$X1
Lama_bekerja=data$X2
Jumlah_Dokumen=data$X3
Jumlah_coauthor=data$X4
Indeks_h=data$Y2
Jenis_kelamin = data$X5.1
Pendidikan_terakhir=data$X7.1
Tempat_Pendidikan_Terakhir=data$X8.1
Sitasi2=Sitasi*Sitasi

#plot data
regresi1 = lm(Sitasi~ Usia, data=data)
y1het = fitted(regresi1)
plot(Usia,Sitasi, col="red")
lines(Usia, y1het, col="blue")

regresi2 = lm(Sitasi~ Lama_bekerja, data=data)
y2het = fitted(regresi2)
plot(Lama_bekerja,Sitasi, col="red")
lines(Lama_bekerja, y2het, col="blue")

regresi3 = lm(Sitasi~ Jumlah_Dokumen, data=data)
y3het = fitted(regresi3)
plot(Jumlah_Dokumen,Sitasi, col="red")
lines(Jumlah_Dokumen, y3het, col="blue")

regresi4 = lm(Sitasi~ Jumlah_coauthor, data=data)
y4het = fitted(regresi4)
plot(Jumlah_coauthor,Sitasi, col="red")
lines(Jumlah_coauthor, y4het, col="blue")

#indeks h
regresi5 = lm(Indeks_h~ Sitasi+Sitasi2, data=data)
regresi5.1 = lm(Indeks_h~ Sitasi, data=data)
y5het = fitted(regresi5)
y5het.1 = fitted(regresi5.1)
plot(Sitasi,Indeks_h, col="red")
lines(Sitasi, y5het, col="blue")
lines(Sitasi, y5het.1, col="blue")

regresi6 = lm(Indeks_h~ Usia, data=data)
y6het = fitted(regresi6)
plot(Usia,Indeks_h, col="red")
lines(Usia, y6het, col="blue")

regresi7 = lm(Indeks_h~ Lama_bekerja, data=data)
y7het = fitted(regresi7)
plot(Lama_bekerja,Indeks_h, col="red")
lines(Lama_bekerja, y7het, col="blue")

regresi8 = lm(Indeks_h~ Jumlah_Dokumen, data=data)
y8het = fitted(regresi8)
plot(Jumlah_Dokumen,Indeks_h, col="red")
lines(Jumlah_Dokumen, y8het, col="blue")

regresi9 = lm(Indeks_h~ Jumlah_coauthor, data=data)
y9het = fitted(regresi9)
plot(Jumlah_coauthor,Indeks_h, col="red")
lines(Jumlah_coauthor, y9het, col="blue")

#korelasi
data=read.csv("E:/ALFI/Thesis me/8. Olah
Data/Olah_korelasi.csv",sep=",")
library(psych)
X=as.matrix(data[,1:6])
Y=as.matrix(data[,1:6])
head(X)
head(Y)
#nilai korelasi
cor(X, Y)

#nilai korelasi dan pvalue
cor.test(Sitasi, Usia)
cor.test(Sitasi, Lama_bekerja)
cor.test(Sitasi, Jumlah_Dokumen)
cor.test(Sitasi, Jumlah_coauthor)
cor.test(Indeks_h, Usia)
cor.test(Indeks_h, Lama_bekerja)
cor.test(Indeks_h, Jumlah_Dokumen)
cor.test(Indeks_h, Jumlah_coauthor)
cor.test(Usia, Lama_bekerja)
cor.test(Usia, Jumlah_Dokumen)
cor.test(Usia, Jumlah_coauthor)
cor.test(Lama_bekerja, Jumlah_Dokumen)
cor.test(Lama_bekerja, Jumlah_coauthor)
cor.test(Jumlah_Dokumen, Jumlah_coauthor)

#boxplot
boxplot(Sitasi~Jenis_kelamin, data=data,
xlab="Jenis_kelamin",ylab="Sitasi")
boxplot(Sitasi~Pendidikan_terakhir, data=data,
xlab="Pendidikan_terakhir",ylab="Sitasi")
boxplot(Sitasi~Tempat_Pendidikan_Terakhir,
data=data,
xlab="Tempat_Pendidikan_Terakhir",ylab="Sitasi")
boxplot(Indeks_h~Jenis_kelamin, data=data,
xlab="Jenis_kelamin",ylab="Indeks_h")
boxplot(Indeks_h~Pendidikan_terakhir, data=data,
xlab="Pendidikan_terakhir",ylab="Indeks_h")
boxplot(Indeks_h~Tempat_Pendidikan_Terakhir,
data=data,
xlab="Tempat_Pendidikan_Terakhir",ylab="Indeks_h")

#prediksi manual dengan x tertentu
X3=data$X3
X4=data$X4
X5.1=data$X5.1
X6.2=data$X6.2
X6.3=data$X6.3
X8.1=data$X8.1
predict(regresi, list(X3=20, X4=1, X5.1=1, X6.2=1
,X6.3=0, X8.1=1 ))
predict(regresi)
#stepwise
stepAIC(regresi)
step(regresi, direction="backward", k=4)
step(regresi, direction="forward")
stepAIC(regresi, direction="both")
library(MASS) #stepwise
library(leaps) #subset

#regresi parameter yang dignifikan sitasi
regresi_sig_sitasi = lm(Y1~ X3+X4+X5.1+X8.1,
data=data)
summary (regresi_sig_sitasi)
anova(regresi_sig_sitasi)
resi=residuals(regresi_sig_sitasi)
yhet1= predict(regresi_sig_sitasi)

#regresi parameter yang dignifikan indeks h
regresi_sig_h = lm(Y2~ Y1+X3+X7.1, data=data)
summary (regresi_sig_h)
anova(regresi_sig_h)
resi2 =residuals(regresi_sig_h)

```

4. Syntax Regresi Kuantil Berbasis Model Rekursif. Model Pertama.

```

#memanggil data
library(quantreg)
data=read.csv("E:/ALFI/Thesis me/8. Olah
Data/Olah.csv",sep=",")

#pakage secara langsung
k1=rq(Y1~X1+X2+X3+X4+X5.1+X6.1+X6.2+X6.3+X7.1+X8.1,
data=data, tau=c(,0.05,0.1,0.25,0.5,0.75,0.9,0.95))
summary(k1)
summary(k1,se="nid") #ada metode "ker", "boot",
"iid", "nid"
yhet=fitted(k1)
resi=residuals(k1)
#membuat garis regresi
attach(data)
#variabel dokumen
plot(X3,Y1, type="n", xlab="Dokumen", ylab="Sitasi")
#cex is ukuran dari bulatan data
points(X3, Y1, cex=1.1,col="black")
abline(rq(Y1~X3,tau=0.05), col="dark violet")
abline(rq(Y1~X3,tau=0.1), col="red")
abline(rq(Y1~X3, tau=0.25), col="green")
abline(rq(Y1~X3, tau=0.5), col="violet")
abline(rq(Y1~X3, tau=0.75), col="orange")
abline(rq(Y1~X3, tau=0.90), col="pink")
abline(rq(Y1~X3, tau=0.95), col="brown")

#variabel Usia
plot(X1,Y1, type="n", xlab="Usia", ylab="Sitasi")
#cex is ukuran dari bulatan data
points(X1, Y1, cex=1.1,col="black")
abline(rq(Y1~X1,tau=0.05), col="dark violet")
abline(rq(Y1~X1,tau=0.1), col="red")
abline(rq(Y1~X1, tau=0.25), col="green")
abline(rq(Y1~X1, tau=0.5), col="violet")
abline(rq(Y1~X1, tau=0.75), col="orange")
abline(rq(Y1~X1, tau=0.90), col="pink")
abline(rq(Y1~X1, tau=0.95), col="brown")

#variabel Lama Bekerja
plot(X2,Y1, type="n", xlab="Lama Bekerja", ylab="Sitasi")
#cex is ukuran dari bulatan data
points(X2, Y1, cex=1.1,col="black")
abline(rq(Y1~X2,tau=0.01), col="blue")
abline(rq(Y1~X2,tau=0.05), col="dark violet")
abline(rq(Y1~X2,tau=0.1), col="red")
abline(rq(Y1~X2, tau=0.25), col="green")
abline(rq(Y1~X2, tau=0.5), col="violet")
abline(rq(Y1~X2, tau=0.75), col="orange")
abline(rq(Y1~X2, tau=0.90), col="pink")
abline(rq(Y1~X2, tau=0.95), col="brown")
abline(rq(Y1~X2, tau=0.99), col="gold")
#variabel Jumlah co-author
plot(X4,Y1, type="n", xlab="Jumlah co-author",
ylab="Sitasi")
#cex is ukuran dari bulatan data
points(X4, Y1, cex=1.1,col="black")
abline(rq(Y1~X4,tau=0.05), col="dark violet")
abline(rq(Y1~X4,tau=0.1), col="red")
abline(rq(Y1~X4, tau=0.25), col="green")
abline(rq(Y1~X4, tau=0.5), col="violet")
abline(rq(Y1~X4, tau=0.75), col="orange")
abline(rq(Y1~X4, tau=0.90), col="pink")
abline(rq(Y1~X4, tau=0.95), col="brown")

#menghitung tau bintang
n=nrow(data)
Y1=as.matrix(data$Y1)
X=as.matrix(data[,4:13])
koef001=as.matrix(k1$coef[,1])
koef05=as.matrix(k1$coef[,2])
koef10=as.matrix(k1$coef[,3])
koef25=as.matrix(k1$coef[,4])
koef50=as.matrix(k1$coef[,5])
koef75=as.matrix(k1$coef[,6])
koef90=as.matrix(k1$coef[,7])
koef95=as.matrix(k1$coef[,8])
koef99=as.matrix(k1$coef[,9])

yhat05=cbind(1,X)%%%koef05
tauhat05=sum(ifelse(Y1<=yhat05,yes=1,no=0))/n
yhat10=cbind(1,X)%%%koef10
tauhat10=sum(ifelse(Y1<=yhat10,yes=1,no=0))/n
yhat25=cbind(1,X)%%%koef25
tauhat25=sum(ifelse(Y1<=yhat25,yes=1,no=0))/n
yhat50=cbind(1,X)%%%koef50
tauhat50=sum(ifelse(Y1<=yhat50,yes=1,no=0))/n
yhat75=cbind(1,X)%%%koef75
tauhat75=sum(ifelse(Y1<=yhat75,yes=1,no=0))/n
yhat90=cbind(1,X)%%%koef90
tauhat90=sum(ifelse(Y1<=yhat90,yes=1,no=0))/n
yhat95=cbind(1,X)%%%koef95
tauhat95=sum(ifelse(Y1<=yhat95,yes=1,no=0))/n

tauhat=rbind(tauhat001,tauhat05,tauhat10,tauhat2
5,tauhat50,tauhat75,tauhat90,tauhat95,tauhat99)
rownames(tauhat)=c("kuantil001","kuantil05","kuan
til10","kuantil25","kuantil50","kuantil75","kuantil90
","kuantil95","kuantil99")
colnames(tauhat)="tauhat"

#mengeksport yhat
QRyhat=data.frame(yhat001,yhat05,yhat10,yhat25,y
hat50,yhat75,yhat90,yhat95,yhat99)
write.csv(QRyhat,"D:/QRyhat.csv")
tauhat

```

5. Syntax Regresi Kuantil Berbasis Model Rekursif. Model Kedua.

```

#memanggil data
library(quantreg)
data=read.csv("E:/ALFI/Thesis me/8. Olah
Data/Olah.csv",sep=",")
head(data)

#pakage secara langsung
k1=rq(Y2~Y1+X1+X2+X3+X4+X5.1+X6.1+X6.2+X6.3+X7.
1+X8.1, data=data,
tau=c(,0.05,0.1,0.25,0.5,0.75,0.9,0.95))
summary(k1)
summary(k1,se="nid") #ada metode "ker", "boot",
"iid", "nid"
yhet=fitted(k1)
resi=residuals(k1)
anova(k1)
plot(Y1,X1)

#membuat garis regresi
attach(data)
#variabel dokumen
plot(X3,Y2, type="n",xlab="Dokumen",ylab="Indeks h")
#cex is ukuran dari bulatan data
points(X3, Y2, cex=1.1,col="black")
abline(rq(Y2~X3,tau=0.05), col="dark violet")
abline(rq(Y2~X3,tau=0.1), col="red")
abline(rq(Y2~X3, tau=0.25), col="green")
abline(rq(Y2~X3, tau=0.5), col="violet")
abline(rq(Y2~X3, tau=0.75), col="orange")
abline(rq(Y2~X3, tau=0.90), col="pink")
abline(rq(Y2~X3, tau=0.95), col="brown")

#variabel Usia
plot(X1,Y2, type="n", xlab="Usia", ylab="Indeks h")
#cex is ukuran dari bulatan data
points(X1, Y2, cex=1.1,col="black")
abline(rq(Y2~X1,tau=0.05), col="dark violet")
abline(rq(Y2~X1,tau=0.1), col="red")
abline(rq(Y2~X1, tau=0.25), col="green")
abline(rq(Y2~X1, tau=0.5), col="violet")
abline(rq(Y2~X1, tau=0.75), col="orange")
abline(rq(Y2~X1, tau=0.90), col="pink")
abline(rq(Y2~X1, tau=0.95), col="brown")
#variabel Lama Bekerja
plot(X2,Y2, type="n", xlab="Lama Bekerja",
ylab="Indeks h")
#cex is ukuran dari bulatan data
points(X2, Y2, cex=1.1,col="black")
abline(rq(Y2~X2,tau=0.05), col="dark violet")
abline(rq(Y2~X2,tau=0.1), col="red")
abline(rq(Y2~X2, tau=0.25), col="green")
abline(rq(Y2~X2, tau=0.5), col="violet")
abline(rq(Y2~X2, tau=0.75), col="orange")
abline(rq(Y2~X2, tau=0.90), col="pink")
abline(rq(Y2~X2, tau=0.95), col="brown")
#variabel Jumlah co-author
plot(X4,Y2, type="n", xlab="Jumlah co-author",
ylab="Indeks h")
#cex is ukuran dari bulatan data
points(X4, Y2, cex=1.1,col="black")
abline(rq(Y2~X4,tau=0.05), col="dark violet")
abline(rq(Y2~X4,tau=0.1), col="red")
abline(rq(Y2~X4, tau=0.25), col="green")
abline(rq(Y2~X4, tau=0.5), col="violet")
abline(rq(Y2~X4, tau=0.75), col="orange")
abline(rq(Y2~X4, tau=0.90), col="pink")
abline(rq(Y2~X4, tau=0.95), col="brown")

#variabel indeks h dan sitasi
plot(Y2,Y1, type="n", xlab="Indeks h", ylab="Sitasi")
#cex is ukuran dari bulatan data
points(Y2, Y1, cex=1.1,col="black")
abline(rq(Y1~Y2,tau=0.05), col="dark violet")
abline(rq(Y1~Y2,tau=0.1), col="red")
abline(rq(Y1~Y2, tau=0.25), col="green")
abline(rq(Y1~Y2, tau=0.5), col="violet")
abline(rq(Y1~Y2, tau=0.75), col="orange")
abline(rq(Y1~Y2, tau=0.90), col="pink")
abline(rq(Y1~Y2, tau=0.95), col="brown")

#jika semua
tau = c(0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95)
for (i in 1:length(tau)) {abline(rq(Y1~X3, tau=tau[i]),
col="black")}

#menghitung tau bintang
n=nrow(data)
Y2=as.matrix(data$Y2)
X=as.matrix(data[,4:14])
koef001=as.matrix(k1$coef[,1])
koef05=as.matrix(k1$coef[,2])
koef10=as.matrix(k1$coef[,3])
koef25=as.matrix(k1$coef[,4])
koef50=as.matrix(k1$coef[,5])
koef75=as.matrix(k1$coef[,6])
koef90=as.matrix(k1$coef[,7])
koef95=as.matrix(k1$coef[,8])
koef99=as.matrix(k1$coef[,9])
yhat05=cbind(1,X)%*%koef05
tauhat05=sum(ifelse(Y1<=yhat05,yes=1,no=0))/n
yhat10=cbind(1,X)%*%koef10
tauhat10=sum(ifelse(Y1<=yhat10,yes=1,no=0))/n
yhat25=cbind(1,X)%*%koef25
tauhat25=sum(ifelse(Y1<=yhat25,yes=1,no=0))/n
yhat50=cbind(1,X)%*%koef50
tauhat50=sum(ifelse(Y1<=yhat50,yes=1,no=0))/n
yhat75=cbind(1,X)%*%koef75
tauhat75=sum(ifelse(Y1<=yhat75,yes=1,no=0))/n
yhat90=cbind(1,X)%*%koef90
tauhat90=sum(ifelse(Y1<=yhat90,yes=1,no=0))/n
yhat95=cbind(1,X)%*%koef95
tauhat95=sum(ifelse(Y1<=yhat95,yes=1,no=0))/n

tauhat=rbind(tauhat001,tauhat05,tauhat10,tauhat25,t
auhat50,tauhat75,tauhat90,tauhat95,tauhat99)
rownames(tauhat)=c("kuantil001","kuantil05","kuantil
10","kuantil25","kuantil50","kuantil75","kuantil90","ku
antil95","kuantil99")
colnames(tauhat)="tauhat"

tauhat

```

Lampiran III. Output Software R.

1. Output Studi Simulasi Data Homogen.

```
> b0rata
[1] 0.9879784
> b1rata
[1] 2.998926
> akurasi0
[1] 1
> akurasi1
[1] 1

> b0q5rata
[1] -7.199994
> b1q5rata
[1] 2.998221
> akurasi0q5
[1] 1
> akurasi1q5
[1] 1

> b0q10rata
[1] -5.420329
> b1q10rata
[1] 2.997911
> akurasi0q10
[1] 1
> akurasi1q10
[1] 1

> b0q25rata
[1] -2.383562
> b1q25rata
[1] 2.998874
> akurasi0q25
[1] 1
> akurasi1q25
[1] 1

> b0q50rata
[1] 0.9928693
> b1q50rata
[1] 2.999583
> akurasi0q50
[1] 1
> akurasi1q50
[1] 1

> b0q75rata
[1] 4.357591
> b1q75rata
[1] 2.999987
> akurasi0q75
[1] 1
> akurasi1q75
[1] 1

> b0q90rata
[1] 7.365733
> b1q90rata
[1] 2.999
> akurasi0q90
[1] 1
> akurasi1q90
[1] 1

> b0q95rata
[1] 9.179529
> b1q95rata
[1] 2.998646
> akurasi0q95
[1] 1
> akurasi1q95
[1] 1
```

2. Output Studi Simulasi Data Heterogen

```
> b0rata
[1] 0.4834363
> b1rata
[1] 5.011052
> akurasi0
[1] 0
> akurasi1
[1] 1

#regresi kuantil
> b0q5rata
[1] -0.07032985
> b1q5rata
[1] 2.254454
> akurasi0q5
[1] 0
> akurasi1q5
[1] 1

> b0q10rata
[1] 0.1402083
> b1q10rata
[1] 2.878287
> akurasi0q10
[1] 0
> akurasi1q10
[1] 1

> b0q25rata
[1] 0.3701076
> b1q25rata
[1] 3.887866
> akurasi0q25
[1] 1
> akurasi1q25
[1] 1

> b0q50rata
[1] 0.4955261
> b1q50rata
[1] 5.007408
> akurasi0q50
[1] 1
> akurasi1q50
[1] 1

> b0q75rata
[1] 0.6320907
> b1q75rata
[1] 6.118276
> akurasi0q75
[1] 0
> akurasi1q75
[1] 1

> b0q90rata
[1] 0.8500199
> b1q90rata
[1] 7.135987
> akurasi0q90
[1] 0
> akurasi1q90
[1] 1

> b0q95rata
[1] 1.066606
> b1q95rata
[1] 7.766282
> akurasi0q95
[1] 0
> akurasi1q95
[1] 1
```

2. Output Model Regresi Linier Berbasis Model Rekursif.

```
> summary (regresi)
Call:
lm(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1)
Residuals:
    Min     1Q   Median     3Q      Max
-256.07 -18.66  -5.04   7.06  432.07
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 12.2444   24.7395   0.495 0.620890
X1          -0.2550    0.8670  -0.294 0.768810
X2          -0.3419    0.9009  -0.379 0.704519
X3           1.6733    0.5418   3.089 0.002137 **
X4           1.3600    0.4047   3.360 0.000846 ***
X5.1        13.0426    5.8624   2.225 0.026595 *
X6.1         9.5143   11.8828   0.801 0.423743
X6.2        14.8425    8.8176   1.683 0.093019 .
X6.3        14.0239    7.7169   1.817 0.069843 .
X7.1        -6.4095    5.9400  -1.079 0.281155
X8.1       -15.3873    5.4350  -2.831 0.004848 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 50.05 on 446 degrees of freedom
Multiple R-squared:  0.3847, Adjusted R-squared:  0.3709
F-statistic: 27.88 on 10 and 446 DF, p-value: < 2.2e-16
```

Stepwise Regression: Y1 versus X1, X2,

```
Alpha-to-Enter: 0.05 Alpha-to-Remove:
0.05
Response is Y1 on 10 predictors, with N =
457
Step           1           2           3           4
Constant      1.745    -2.845    4.402    1.428
X3             3.53     1.88     1.69     1.74
T-Value       15.16     3.45     3.12     3.22
P-Value       0.000     0.001    0.002    0.001
X4             1.35     1.36     1.36
T-Value       3.33     3.40     3.40
P-Value       0.001     0.001    0.001
X8.1          -15.8    -17.1
T-Value       -3.17    -3.43
P-Value       0.002     0.001
X5.1          14.4
T-Value       2.53
P-Value       0.012
S             51.5     50.9     50.4     50.1
R-Sq          33.55    35.14    36.55    37.44
R-Sq(adj)     33.41    34.85    36.13    36.88
Mallows Cp    28.6     19.1     10.9     6.5
```

```
> summary (regresi_h)
Call:
lm(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 +
X6.1 + X6.2 + X6.3 + X7.1 + X8.1)
Residuals:
    Min     1Q   Median     3Q      Max
-4.3949 -0.4921 -0.1797  0.5276  3.0629
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.6219790  0.4575046  3.545 0.000434 ***
Y1           0.0189534  0.0008754  21.651 < 2e-16 ***
X1          -0.0221825  0.0160306  -1.384 0.167125
X2           0.0165839  0.0166582   0.996 0.320015
X3           0.0370924  0.0101231   3.664 0.000278 ***
X4           0.0136793  0.0075764   1.806 0.071670 .
X5.1        -0.0005197  0.1089828  -0.005 0.996198
X6.1         0.1557485  0.2198447   0.708 0.479038
X6.2         0.1208387  0.1635346   0.739 0.460346
X6.3         0.0364772  0.1431962   0.255 0.799045
X7.1        -0.5559565  0.1099610  -5.056 6.27e-07 ***
X8.1        -0.1048510  0.1013802  -1.034 0.301588
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9253 on 445 degrees of
freedom
Multiple R-squared:  0.7821, Adjusted R-squared:
0.7768
F-statistic: 145.2 on 11 and 445 DF, p-value: < 2.2e-16
```

Stepwise Regression: Y2 versus Y1, X1, ...

```
Alpha-to-Enter: 0.05 Alpha-to-Remove:
0.05
Response is Y2 on 11 predictors, with N =
457
Step           1           2           3
Constant      0.9759    0.7032    1.0156
Y1            0.02574  0.01996  0.01937
T-Value       31.67     22.65     22.88
P-Value       0.000     0.000     0.000
X3            0.0608    0.0538
T-Value       11.32     10.27
P-Value       0.000     0.000
X7.1          -0.625
T-Value       -6.68
P-Value       0.000
S             1.10     0.968     0.925
R-Sq          68.80    75.67    77.85
R-Sq(adj)     68.73    75.56    77.71
Mallows Cp    184.3     46.0     3.4
```

3. Output Model Regresi linier Berbasis Model Rekursif Pemilihan Model Terbaik.

```
> summary(regresi_signifikan)
Call:
lm(formula = Y2 ~ Y1 + X3 + X7.1)

Residuals:
    Min       1Q   Median       3Q      Max
-4.5823 -0.4445 -0.1963  0.4630  3.0193
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.0156084  0.0701043  14.487 < 2e-16 ***
Y1           0.0193665  0.0008465   22.877 < 2e-16 ***
X3           0.0537588  0.0052334   10.272 < 2e-16 ***
X7.1        -0.6248267  0.0935142   -6.682 6.95e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9247 on 453 degrees of freedom
Multiple R-squared:  0.7785, Adjusted R-squared:  0.7771
F-statistic: 530.8 on 3 and 453 DF, p-value: < 2.2e-1
```

```
> summary(regresi_signifikan)
Call:
lm(formula = Y1 ~ X3 + X4 + X5.1 + X8.1)

Residuals:
    Min       1Q   Median       3Q      Max
-268.86 -18.39  -5.92    7.61  434.56
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.4281    4.0485   0.353 0.724447
X3           1.7392    0.5398   3.222 0.001366 **
X4           1.3589    0.3993   3.403 0.000726 ***
X5.1        14.3618    5.6818   2.528 0.011822 *
X8.1        -17.0511   4.9771  -3.426 0.000669 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 50.13 on 452 degrees of freedom
Multiple R-squared:  0.3744, Adjusted R-squared:  0.3688
F-statistic: 67.61 on 4 and 452 DF, p-value: < 2.2e-16
```

```
> cor.test(a,b)
Pearson's product-moment correlation
data: a and b
t = -0.44853, df = 455, p-value = 0.654
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.11253309 0.07084071
sample estimates:
cor
-0.02102299
> cor(a,b)
[1] -0.02102299
> ks.test(a, "pnorm")
One-sample Kolmogorov-Smirnov test
data: a
D = 0.60369, p-value < 2.2e-16
alternative hypothesis: two-sided
> ks.test(b, "pnorm")
One-sample Kolmogorov-Smirnov test
data: b
D = 0.12362, p-value = 1.719e-06
alternative hypothesis: two-sided
```

4. Output Model Regresi Kuantil Berbasis Model Rekursif Model Pertama.

```
> summary(k1,se="nid") #ada metode "ker", "boot",
"iid","nid"
Call: rq(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1,
0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)
```

```
.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75,
0.9, 0.95, 0.975, 0.99), data = data)
```

```
tau: [1] 0.05
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-7.66899	1.69225	-4.53184	0.00001
X1	0.14355	0.06384	2.24844	0.02504
X2	-0.11891	0.06257	-1.90033	0.05803
X3	1.19818	0.16267	7.36558	0.00000
X4	-0.23192	0.05275	-4.39694	0.00001
X5.1	0.56133	0.29634	1.89420	0.05885
X6.1	-0.38511	0.81632	-0.47177	0.63732
X6.2	0.81093	0.65460	1.23882	0.21606
X6.3	-0.09695	0.44108	-0.21980	0.82613
X7.1	0.08516	0.37344	0.22805	0.81971
X8.1	0.49009	0.37489	1.30731	0.19178

```
Call: rq(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1,
0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)
```

```
tau: [1] 0.1
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-5.71759	1.79398	-3.18710	0.00154
X1	0.03504	0.06358	0.55109	0.58185
X2	0.03194	0.07642	0.41798	0.67616
X3	1.20179	0.07194	16.70524	0.00000
X4	-0.16897	0.05346	-3.16095	0.00168
X5.1	0.62870	0.35092	1.79160	0.07388
X6.1	-0.14084	0.85761	-0.16422	0.86963
X6.2	0.09461	0.70244	0.13468	0.89292
X6.3	0.19329	0.54366	0.35553	0.72236
X7.1	1.01319	0.41034	2.46912	0.01392
X8.1	0.34458	0.27353	1.25975	0.20842

```
Call: rq(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1,
0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)
```

```
tau: [1] 0.5
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-0.48485	1.82518	-0.26564	0.79064
X1	-0.04431	0.05479	-0.80861	0.41917
X2	0.02215	0.06369	0.34782	0.72814
X3	1.48798	0.59328	2.50808	0.01249
X4	0.42090	0.22438	1.87579	0.06134
X5.1	0.31014	0.16754	1.85113	0.06481
X6.1	-0.36991	0.93789	-0.39440	0.69347
X6.2	0.22153	0.84036	0.26361	0.79220
X6.3	0.44305	0.65519	0.67622	0.49925
X7.1	-0.89237	0.44503	-2.00518	0.04555
X8.1	-0.22153	0.44720	-0.49536	0.62059

```
Call: rq(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1,
0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)
```

```
tau: [1] 0.75
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-1.13337	8.43605	-0.13435	0.89319
X1	-0.10350	0.24743	-0.41828	0.67594
X2	0.07010	0.22218	0.31549	0.75254
X3	4.34885	1.49583	2.90731	0.00383
X4	0.51749	0.60005	0.86241	0.38892
X5.1	0.12043	1.53795	0.07831	0.93762
X6.1	0.79751	5.50730	0.14481	0.88493
X6.2	0.21405	2.99566	0.07145	0.94307
X6.3	0.49067	2.76008	0.17777	0.85898
X7.1	-0.49773	1.49113	-0.33379	0.73869
X8.1	-1.09849	2.19117	-0.50132	0.61639

```
Call: rq(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1,
0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)
```

```
tau: [1] 0.9
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	12.11037	22.84266	0.53016	0.59626
X1	-0.44482	0.76501	-0.58145	0.56123
X2	0.32776	0.84903	0.38604	0.69965
X3	7.71906	1.91227	4.03659	0.00006
X4	1.19398	0.94707	1.26070	0.20807
X5.1	2.57525	2.64844	0.97236	0.33140
X6.1	-1.71237	11.87137	-0.14424	0.88537
X6.2	2.10702	8.67899	0.24277	0.80829
X6.3	3.13712	6.65618	0.47131	0.63765
X7.1	0.21070	3.44270	0.06120	0.95123
X8.1	-10.04348	10.28471	-0.97654	0.32932

```
Call: rq(formula = Y1 ~ X1 + X2 + X3 + X4 + X5.1 + X6.1 +
X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.025, 0.05, 0.1, 0.25,
0.5, 0.75, 0.9, 0.95, 0.975), data = data)
```

```
tau: [1] 0.95
```

```
Coefficients:
```

	Value	Std. Error	t value	Pr(> t)
(Intercept)	23.85873	53.99609	0.44186	0.65880
X1	0.31579	1.96452	0.16075	0.87237
X2	-0.31579	2.27125	-0.13904	0.88948
X3	9.62604	2.96369	3.24799	0.00125
X4	1.57895	1.88582	0.83727	0.40289
X5.1	0.94737	9.53660	0.09934	0.92091
X6.1	-15.94183	41.24031	-0.38656	0.69927
X6.2	-2.84211	20.35450	-0.13963	0.88902
X6.3	-0.63158	18.49432	-0.03415	0.97277
X7.1	-0.63158	11.74350	-0.05378	0.95713
X8.1	-42.64266	19.33138	-2.20588	0.02790

```
tauhat
kuantil05 0.05908096
kuantil10 0.08971554
kuantil25 0.25382932
kuantil50 0.49234136
kuantil75 0.73741794
kuantil90 0.89715536
kuantil95 0.95842451
```


5. Output model Regresi Kuantil Berbasis Model Rekursif Model kedua.

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.05

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	-0.17244	0.12985	-1.32806	0.18484
Y1	0.01411	0.00164	8.60477	0.00000
X1	0.00238	0.00464	0.51300	0.60821
X2	-0.00315	0.00532	-0.59132	0.55460
X3	0.04254	0.00853	4.98951	0.00000
X4	-0.00925	0.00546	-1.69465	0.09084
X5.1	0.03066	0.05296	0.57889	0.56296
X6.1	0.07777	0.04949	1.57140	0.11680
X6.2	0.04507	0.03147	1.43211	0.15282
X6.3	0.03144	0.03327	0.94507	0.34514
X7.1	-0.01899	0.02566	-0.74001	0.45968
X8.1	0.03533	0.02353	1.50174	0.13387

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.1

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.16766	0.21746	0.77102	0.44111
Y1	0.01378	0.00218	6.33574	0.00000
X1	-0.00915	0.00697	-1.31141	0.19040
X2	0.00448	0.00657	0.68224	0.49544
X3	0.05479	0.01932	2.83639	0.00477
X4	0.00220	0.00921	0.23840	0.81168
X5.1	0.02964	0.02789	1.06267	0.28851
X6.1	0.16967	0.11371	1.49220	0.13636
X6.2	0.09476	0.05938	1.59590	0.11122
X6.3	0.08287	0.05171	1.60273	0.10970
X7.1	-0.05708	0.03432	-1.66330	0.09696
X8.1	-0.00375	0.03122	-0.12012	0.90444

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.25

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.54292	0.28015	1.93792	0.05326
Y1	0.01641	0.00342	4.80332	0.00000
X1	-0.00239	0.00866	-0.27572	0.78289
X2	0.00217	0.00919	0.23615	0.81343
X3	0.03409	0.02160	1.57853	0.11515
X4	0.00964	0.00957	1.00706	0.31445
X5.1	0.01515	0.03178	0.47662	0.63387
X6.1	0.16918	0.22399	0.75528	0.45048
X6.2	-0.00182	0.06809	-0.02675	0.97867
X6.3	0.01897	0.04595	0.41285	0.67991
X7.1	-0.55941	0.13976	-4.00264	0.00007
X8.1	0.00551	0.03066	0.17976	0.85742

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.5

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	0.75257	0.34372	2.18949	0.02908
Y1	0.02016	0.00266	7.58317	0.00000
X1	-0.00410	0.01153	-0.35536	0.72249
X2	0.00453	0.01161	0.38985	0.69684
X3	0.08026	0.01913	4.19652	0.00003
X4	0.00859	0.01102	0.77950	0.43610
X5.1	0.02796	0.09660	0.28940	0.77241
X6.1	0.04692	0.20156	0.23277	0.81605
X6.2	0.00980	0.13920	0.07039	0.94391
X6.3	-0.00452	0.10634	-0.04254	0.96609
X7.1	-0.75060	0.08023	-9.35560	0.00000
X8.1	0.00406	0.07459	0.05447	0.95659

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.75

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	1.06790	0.40673	2.62554	0.00895
Y1	0.02582	0.00396	6.52286	0.00000
X1	0.00173	0.01334	0.12990	0.89670
X2	0.00227	0.01411	0.16068	0.87242
X3	0.05848	0.01725	3.39010	0.00076
X4	0.01678	0.00795	2.11095	0.03533
X5.1	-0.02758	0.09964	-0.27681	0.78205
X6.1	-0.10990	0.20438	-0.53770	0.59105
X6.2	-0.04472	0.16949	-0.26388	0.79200
X6.3	-0.00657	0.14050	-0.04673	0.96275
X7.1	-0.44331	0.10027	-4.42122	0.00001
X8.1	-0.08087	0.08939	-0.90472	0.36610

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

Lanjutan : Output model Regresi Kuantil Berbasis Model Rekursif Model Kedua.

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.9

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	1.06525	0.54361	1.95958	0.05067
Y1	0.03470	0.00451	7.70025	0.00000
X1	0.01003	0.01772	0.56566	0.57191
X2	-0.00414	0.01835	-0.22564	0.82159
X3	0.06134	0.01139	5.38407	0.00000
X4	0.01551	0.00950	1.63275	0.10323
X5.1	0.02140	0.11421	0.18739	0.85144
X6.1	-0.30206	0.30277	-0.99763	0.31900
X6.2	0.07982	0.23482	0.33990	0.73409
X6.3	-0.14459	0.21072	-0.68618	0.49296
X7.1	-0.40592	0.15148	-2.67973	0.00764
X8.1	-0.11605	0.11227	-1.03368	0.30185

Call: rq(formula = Y2 ~ Y1 + X1 + X2 + X3 + X4 + X5.1 + X6.1 + X6.2 + X6.3 + X7.1 + X8.1, tau = c(0.01, 0.025, 0.05, 0.1, 0.25, 0.5, 0.75, 0.9, 0.95, 0.975, 0.99), data = data)

tau: [1] 0.95

Coefficients:

	Value	Std. Error	t value	Pr(> t)
(Intercept)	1.42046	0.67620	2.10065	0.03623
Y1	0.03868	0.00199	19.43036	0.00000
X1	0.00910	0.02171	0.41886	0.67552
X2	0.00129	0.02116	0.06098	0.95140
X3	0.07791	0.01050	7.41923	0.00000
X4	0.00473	0.00624	0.75716	0.44935
X5.1	-0.25430	0.12967	-1.96113	0.05049
X6.1	-0.35821	0.34330	-1.04343	0.29732
X6.2	-0.31245	0.18495	-1.68934	0.09185
X6.3	-0.41569	0.12967	-3.20574	0.00144
X7.1	-0.15965	0.15737	-1.01452	0.31089
X8.1	-0.03115	0.13369	-0.23300	0.81587

Lanjutan : Output Regresi Kuantil Berbasis Model Rekursif Model Kedua

korelasi

```
> cor(residual1, residual2)
      tau..0.01 tau..0.05 tau..0.10 tau..0.25 tau..0.50 tau..0.75 tau..0.90 tau..0.95 tau..0.99
tau..0.01 0.5323992 0.4581569 0.376319284 0.23067431 -0.25228053 -0.4819558 -0.7404116 -0.78005201 -0.8204483
tau..0.05 0.5284466 0.4535946 0.375750058 0.22811811 -0.24660422 -0.4779804 -0.7354777 -0.77567637 -0.8216222
tau..0.10 0.5252518 0.4502927 0.375297577 0.22706867 -0.24277251 -0.4748404 -0.7331230 -0.77271041 -0.8216301
tau..0.25 0.5106202 0.4357854 0.369138587 0.22224851 -0.23181970 -0.4672991 -0.7256102 -0.76618729 -0.8243676
tau..0.50 0.4587900 0.3831914 0.348330769 0.20335833 -0.17900522 -0.4263360 -0.6822955 -0.72252104 -0.8224913
tau..0.75 0.1844026 0.1180277 0.210818681 0.08989612 0.06922109 -0.1751505 -0.3677746 -0.39017934 -0.6390467
tau..0.90 -0.1578135 -0.1916000 -0.001762689 -0.05633823 0.28726782 0.1294802 0.0645533 0.06147923 -0.2515698
tau..0.95 -0.2758904 -0.2953841 -0.089109874 -0.10629201 0.33558736 0.2219809 0.2076639 0.20952395 -0.1017544
tau..0.99 -0.3162422 -0.3072932 -0.140026672 -0.10315482 0.20133175 0.1348228 0.1301994 0.07845476 -0.1222556
```

```
> cor.test(d, f)
```

Pearson's product-moment correlation

data: d and f

t = 13.416, df = 455, p-value < 2.2e-16

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.4632974 0.5950660

sample estimates:

cor

0.5323992

```
> cor.test(d, f)
```

Pearson's product-moment correlation

data: d and f

t = 10.857, df = 455, p-value < 2.2e-16

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.3775773 0.5235388

sample estimates:

cor

0.4535946

```
> cor.test(d, f)
```

Pearson's product-moment correlation

data: d and f

t = 8.6367, df = 455, p-value < 2.2e-16

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

0.2936804 0.4514824

sample estimates:

cor

0.3752976

Lanjutan : Output Regresi Kuantil Berbasis Model Rekursif Model Kedua

```
> cor.test(d, f)
```

```
Pearson's product-moment correlation
```

```
data: d and f
t = 4.8623, df = 455, p-value = 1.601e-06
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1332376 0.3077027
sample estimates:
   cor
0.2222485
```

```
> cor.test(d, f)
```

```
Pearson's product-moment correlation
```

```
data: d and f
t = -3.881, df = 455, p-value = 0.0001194
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.26635880 -0.08873511
sample estimates:
   cor
-0.1790052
```

```
> cor.test(d, f)
```

```
Pearson's product-moment correlation
```

```
data: d and f
t = -3.7947, df = 455, p-value = 0.0001678
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.26265770 -0.08478553
sample estimates:
   cor
-0.1751505
```

```
> cor.test(d, f)
```

```
Pearson's product-moment correlation
```

```
data: d and f
t = 1.3798, df = 455, p-value = 0.1683
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.02733566 0.15536047
sample estimates:
   cor
0.0645533
```

Lanjutan : Output Regresi Kuantil Berbasis Model Rekursif Model Kedua

```
> cor.test(d, f)
```

```
    Pearson's product-moment correlation
```

```
data: d and f
```

```
t = 4.5708, df = 455, p-value = 6.27e-06
```

```
alternative hypothesis: true correlation is not equal to 0
```

```
95 percent confidence interval:
```

```
0.1201052 0.2955705
```

```
sample estimates:
```

```
cor  
0.2095239
```

```
> cor.test(d, f)
```

```
    Pearson's product-moment correlation
```

```
data: d and f
```

```
t = -2.6275, df = 455, p-value = 0.008892
```

```
alternative hypothesis: true correlation is not equal to 0
```

```
95 percent confidence interval:
```

```
-0.21160964 -0.03087469
```

```
sample estimates:
```

```
cor  
-0.1222556
```

Lampiran IV. Statistika Deskriptif.

1. Output Statistika Deskriptif Indeks h Berdasarkan Jurusan.

Variable	Jurusan	Total Count	Mean	StDev	Minimum	Median
Indeks h	ARSITEKTUR	9	0,333	0,707	0,000	0,000
	BIOLOGI	8	2,13	3,36	0,00	1,00
	D. INTERIOR	1	0,000000	*	0,000000	0,000000
	DESPRO	2	1,0000	0,000000	1,0000	1,0000
	FISIKA	15	2,400	1,765	1,000	2,000
	KIMIA	17	2,706	2,229	0,000	2,000
	MAN. BISNIS	4	0,750	0,500	0,000	1,000
	MATEMATIKA	23	0,913	1,041	0,000	1,000
	PWK	2	0,000000	0,000000	0,000000	0,000000
	SI	24	0,917	1,248	0,000	0,000
	STATISTIKA	24	1,750	1,800	0,000	2,000
	T. ELEKTRO	47	2,085	1,730	0,000	2,000
	T. FISIKA	21	2,143	2,393	0,000	1,000
	T. GEOFISIKA	5	0,800	0,837	0,000	1,000
	T. GEOMATIKA	10	0,400	0,843	0,000	0,000
	T. INDUSTRI	19	1,789	2,016	0,000	1,000
	T. INFORMATIKA	38	1,368	1,496	0,000	1,000
	T. KELAUTAN	20	1,200	2,093	0,000	1,000
	T. KIMIA	34	3,529	3,440	0,000	2,000
	T. LINGKUNGAN	11	2,545	1,214	1,000	2,000
	T. MESIN	39	1,179	1,211	0,000	1,000
	T. MMT	13	1,462	1,664	0,000	1,000
	T. PERKAPALAN	9	1,556	1,014	0,000	1,000
	T. SIPIL	35	1,057	1,083	0,000	1,000
	T. SISKAL	12	0,917	2,021	0,000	0,000
	TMJ	13	2,077	1,754	0,000	2,000
	TRANSLA	2	1,0000	0,000000	1,0000	1,0000

2. Output Statistika Deskriptif Sitasi Berdasarkan Jurusan.

Variable	Jurusan	Total Count	Mean	StDev	Minimum	Median
Sitasi	ARSITEKTUR	9	1.222	2.728	0.000	0.000
	BIOLOGI	8	45.1	104.1	0.0	4.0
	D. INTERIOR	1	0.000000	*	0.000000	0.000000
	DESPRO	2	2.00	1.41	1.00	2.00
	FISIKA	15	33.7	39.8	1.0	8.0
	KIMIA	17	53.4	62.9	0.0	26.0
	MAN. BISNIS	4	4.00	4.83	0.00	2.50
	MATEMATIKA	23	8.87	16.35	0.00	1.00
	PWK	2	0.000000	0.000000	0.000000	0.000000
	SI	24	9.21	17.37	0.00	0.00
	STATISTIKA	24	16.46	26.75	0.00	8.00
	T. ELEKTRO	47	25.66	45.36	0.00	12.00
	T. FISIKA	21	29.7	67.8	0.0	6.0
	T. GEOFISIKA	5	4.60	6.15	0.00	4.00
	T. GEOMATIKA	10	6.60	14.69	0.00	0.00
	T. INDUSTRI	19	33.3	75.5	0.0	5.0
	T. INFORMATIKA	38	16.00	32.30	0.00	5.00
	T. KELAUTAN	20	15.9	50.5	0.0	1.0
	T. KIMIA	34	105.9	154.7	0.0	29.5
	T. LINGKUNGAN	11	51.3	65.1	1.0	19.0
	T. MESIN	39	7.28	10.07	0.00	4.00
	T. MMT	13	31.4	61.3	0.0	4.0
	T. PERKAPALAN	9	15.22	24.20	0.00	6.00
	T. SIPIL	35	13.46	26.98	0.00	4.00
	T. SISKAL	12	13.08	32.53	0.00	0.00
	TMJ	13	27.2	41.0	0.0	13.0
	TRANSLA	2	2.500	0.707	2.000	2.500

3. Output Analisis Faktor

Rotated Component Matrix^a

	Component	
	1	2
indeks_h	.944	.051
Sitasi	.904	-.017
Usia	.027	.990
Lama_Bekerja	.057	.989
Jumlah_dokumen	.834	.078

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser

Normalization.

a. Rotation converged in 3 iterations.

4. Output Jumlah Masing-masing Anggota Cluster.

Cluster Number of Case

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1	57	12.5	12.5	12.5
2	387	84.7	84.7	97.2
3	13	2.8	2.8	100.0
Total	457	100.0	100.0	

Jurusan * Cluster Number of Case Crosstabulation

Count		Cluster Number of Case			Total
		1	2	3	
Jurusan	ARSITEKTUR	0	9	0	9
	BIOLOGI	0	7	1	8
	D.INTERIOR	0	1	0	1
	DESPRO	0	2	0	2
	FISIKA	4	11	0	15
	KIMIA	7	10	0	17
	MAN. BISNIS	0	4	0	4
	MATEMATIKA	1	22	0	23
	PWK	0	2	0	2
	SI	2	22	0	24
	STATISTIKA	2	22	0	24
	T.ELEKTRO	9	36	2	47
	T.FISIKA	2	18	1	21
	T.GEOFISIKA	0	5	0	5
	T.GEOMATIKA	0	10	0	10
	T.INDUSTRI	2	16	1	19
	T.INFORMATIKA	6	32	0	38
	T.KELAUTAN	1	18	1	20
	T.KIMIA	7	21	6	34
	T.LINGKUNGAN	3	8	0	11
	T.MESIN	1	38	0	39
	T.MMT	2	11	0	13
	T.PERKAPALAN	2	7	0	9
	T.SIPIL	2	33	0	35
	T.SISKAL	1	11	0	12
	TMJ	3	9	1	13
TRANSLA	0	2	0	2	
Total		57	387	13	457

5. Karakteristik Dosen Per Cluster.

Descriptive Statistics: Y2, Y1, X1, X2, X3

Variable	cluster	Mean	StDev	Minimum	Median	Maximum
Y2	1	4.351	1.275	2.000	4.000	8.000
	2	1.0233	0.9775	0.0000	1.0000	4.0000
	3	8.692	2.529	5.000	9.000	14.000
Y1	1	93.30	61.11	19.00	77.00	337.00
	2	7.651	11.497	0.000	3.000	95.000
	3	292.5	142.5	100.0	285.0	682.0
X1	1	46.77	8.02	32.00	46.00	65.00
	2	46.028	10.326	25.000	46.000	68.000
	3	48.15	7.65	38.00	47.00	64.00
X2	1	21.02	7.93	7.00	21.00	38.00
	2	19.341	10.320	1.000	19.000	43.000
	3	22.15	7.01	13.00	23.00	36.00
X3	1	16.67	11.65	3.00	13.00	58.00
	2	4.328	3.728	1.000	3.000	20.000
	3	44.08	28.95	19.00	37.00	125.00

6. Jumlah Dokumen Dosen ITS di Scopus.

Fakultas	Jurusan	n	Mean	StDev	Median	Min	Max
FMIPA	Biologi	8	4,75	7,11	2,50	1	22
	Fisika	15	9,40	9,92	7,00	1	37
	Kimia	17	6,24	4,52	5,00	1	19
	Matematika	23	4,30	4,84	3,00	1	20
	Statistika	24	6,88	7,99	5,00	1	38
Total		87	6,31	2,25	4,00		
FTI	MB	4	2,75	2,22	2,00	1	6
	T. Elektro	47	12,32	13,57	9,00	1	63
	T. Fisika	21	8,57	8,42	6,00	1	37
	T. Industri	19	5,11	6,53	4,00	1	24
	T. Kimia	34	10,68	13,87	6,00	1	66
	T. Mesin	39	4,74	3,59	4,00	1	12
	T. Material	13	6,00	4,67	5,00	1	14
	TMJ	13	21,31	35,50	1,00	1	125
Total		190	8,94	10,78	5,00		
FTIf	SI	24	3,79	2,80	3,00	1	11
	T. Informatika	38	6,68	6,15	5,00	1	24
Total		62	5,24	2,37	4,50		
FTK	T. Kelautan	20	6,40	10,41	3,00	1	48
	T. Perkapalan	9	6,22	5,78	6,00	1	20
	T. Siskal	12	5,17	7,60	1,00	1	26
	Transla	2	2,00	0,00	2,00	2	2
Total		43	4,95	4,40	3,00		
FTSP	Arsitektur	9	2,22	2,28	1,00	1	8
	D. Interior	1	1,00	*	1,00	1	1
	D. Produk	2	2,50	2,12	2,50	1	4
	PWK	2	1,00	0,00	1,00	1	1
	T.Geofisika	5	1,60	0,89	1,00	1	3
	T.Geomatika	10	2,00	1,05	2,00	1	4
	T.Lingkungan	11	5,36	2,98	6,00	2	12
T.Sipil	35	4,83	5,14	2,00	1	18	
Total		75	2,56	1,68	2,00		
Total ITS		457	5,60	3,77	4,00		

7. Persentase Kepemilikan Publikasi Berdasarkan Jenis Kelamin.

Fakultas	Jurusan	Pemilik Scopus		
		n	L	P
FMIPA	Biologi	8	0,50	0,50
	Fisika	15	0,93	0,07
	Kimia	17	0,65	0,35
	Matematika	23	0,78	0,22
	Statistika	24	0,54	0,46
Total		87	0,69	0,31
FTI	MB	4	0,75	0,25
	T. Elektro	47	0,89	0,11
	T. Fisika	21	0,81	0,19
	T. Industri	19	0,74	0,26
	T. Kimia	34	0,68	0,32
	T. Mesin	39	0,90	0,10
	T. Material	13	0,62	0,38
TMJ	13	0,85	0,15	
Total		190	0,81	0,19
FTIf	SI	24	0,71	0,29
	T. Informatika	38	0,68	0,32
Total		62	0,69	0,31
FTK	T. Kelautan	20	0,95	0,05
	T. Perkapalan	9	1,00	0,00
	T. Siskal	12	1,00	0,00
	Transla	2	1,00	0,00
Total		43	0,98	0,02
FTSP	Arsitektur	9	0,56	0,44
	D. Interior	1	1,00	0,00
	D. Produk	2	1,00	0,00
	PWK	2	0,50	0,50
	T.Geofisika	5	0,80	0,20
	T.Geomatika	10	0,70	0,30
T.Lingkungan	11	0,64	0,36	
T.Sipil	35	0,86	0,14	
Total		75	0,76	0,24
Total ITS		457	0,78	0,22

8. Deskripsi Jabatan Dosen ITS per Jurusan.

Fakultas	Jurusan	n	AA	L	LK	GB
FMIPA	Biologi	8	0,63	0,13	0,25	0,00
	Fisika	15	0,13	0,33	0,27	0,27
	Kimia	17	0,06	0,35	0,35	0,24
	Matematika	23	0,17	0,39	0,30	0,13
	Statistika	24	0,25	0,33	0,33	0,08
Total		87	0,21	0,33	0,31	0,15
FTI	MB	4	0,25	0,75	0,00	0,00
	T. Elektro	47	0,26	0,34	0,26	0,15
	T. Fisika	21	0,14	0,33	0,48	0,05
	T. Industri	19	0,16	0,32	0,16	0,37
	T. Kimia	34	0,21	0,26	0,21	0,32
	T. Mesin	39	0,21	0,33	0,23	0,23
	T. Material	13	0,08	0,62	0,23	0,08
TMJ	13	0,23	0,46	0,23	0,08	
Total		190	0,20	0,36	0,25	0,19
FTIF	SI	24	0,54	0,17	0,25	0,04
	T. Informatika	38	0,32	0,24	0,34	0,11
Total		62	0,40	0,21	0,31	0,08
FTK	T. Kelautan	20	0,15	0,45	0,20	0,20
	T. Perkapalan	9	0,11	0,33	0,22	0,33
	T. Siskal	12	0,08	0,42	0,42	0,08
	Transla	2	0,00	1,00	0,00	0,00
Total		43	0,12	0,44	0,26	0,19
FTSP	Arsitektur	9	0,11	0,22	0,56	0,11
	D. Interior	1	0,00	1,00	0,00	0,00
	D. Produk	2	0,00	0,50	0,50	0,00
	PWK	2	0,50	0,50	0,00	0,00
	T.Geofisika	5	0,20	0,60	0,20	0,00
	T.Geomatika	10	0,40	0,50	0,00	0,10
T.Lingkungan	11	0,09	0,36	0,27	0,27	
T.Sipil	35	0,29	0,29	0,23	0,20	
Total		75	0,24	0,36	0,24	0,16
Total ITS		457	0,23	0,34	0,27	0,16

9. Kepemilikan Publikasi Berdasarkan Pendidikan Terakhir untuk Masing-masing Jurusan.

Fakultas	Jurusan	Pemilik <i>Scopus</i>		
		n	S2	S3
FMIPA	Biologi	8	0,50	0,50
	Fisika	15	0,27	0,73
	Kimia	17	0,12	0,88
	Matematika	23	0,43	0,57
	Statistika	24	0,29	0,71
Total		87	0,31	0,69
FTI	MB	4	0,50	0,50
	T. Elektro	47	0,34	0,66
	T. Fisika	21	0,52	0,48
	T. Industri	19	0,37	0,63
	T. Kimia	34	0,24	0,76
	T. Mesin	39	0,36	0,64
	T. Material	13	0,46	0,54
	TMJ	13	0,46	0,54
Total		190	0,37	0,63
FTIF	SI	24	0,75	0,25
	T. Informatika	38	0,66	0,34
Total		62	0,69	0,31
FTK	T. Kelautan	20	0,25	0,75
	T. Perkapalan	9	0,33	0,67
	T. Siskal	12	0,33	0,67
	Transla	2	0,00	1,00
Total		43	0,28	0,72
FTSP	Arsitektur	9	0,22	0,78
	D. Interior	1	0,00	1,00
	D. Produk	2	1,00	0,00
	PWK	2	1,00	0,00
	T. Geofisika	5	0,20	0,80
	T. Geomatika	10	0,60	0,40
	T. Lingkungan	11	0,18	0,82
T. Sipil	35	0,40	0,60	
Total		75	0,39	0,61
Total ITS		457	0,40	0,60

10. Persentase Kepemilikan Scopus Berdasarkan Tempat Pendidikan Terakhir untuk Masing-masing Jurusan.

Fakultas	Jurusan	Pemilik Scopus		
		n	DN	LN
FMIPA	Biologi	8	0,50	0,50
	Fisika	15	0,53	0,47
	Kimia	17	0,29	0,71
	Matematika	23	0,78	0,22
	Statistika	24	0,75	0,25
	Total	87	0,61	0,39
FTI	MB	4	0,50	0,50
	T. Elektro	47	0,43	0,57
	T. Fisika	21	0,67	0,33
	T. Industri	19	0,37	0,63
	T. Kimia	34	0,15	0,85
	T. Mesin	39	0,26	0,74
	T. Material	13	0,23	0,77
	TMJ	13	0,46	0,54
	Total	190	0,35	0,65
FTIf	SI	24	0,42	0,58
	T. Informatika	38	0,34	0,66
	Total	62	0,37	0,63
FTK	T. Kelautan	20	0,05	0,95
	T. Perkapalan	9	0,11	0,89
	T. Siskal	12	0,17	0,83
	Transla	2	0,00	1,00
	Total	43	0,09	0,91
FTSP	Arsitektur	9	0,44	0,56
	D. Interior	1	1,00	0,00
	D. Produk	2	1,00	0,00
	PWK	2	0,00	1,00
	T. Geofisika	5	0,80	0,20
	T. Geomatika	10	0,50	0,50
	T. Lingkungan	11	0,18	0,82
T. Sipil	35	0,31	0,69	
	Total	75	0,39	0,61
	Total ITS	457	0,39	0,61

11. Deskripsi Jumlah *Co-author* Dosen ITS di Scopus untuk Masing-masing Jurusan.

Variabel	Jurusan	N	Mean	StDev	Median	Min	Max
FMIPA	Biologi	8	11,88	12,22	7,00	0	36
	Fisika	15	23,20	23,96	14,00	5	88
	Kimia	17	16,53	11,32	17,00	3	45
	Matematika	23	7,22	6,59	6,00	0	31
	Statistika	24	10,79	11,41	8,50	1	52
Total		87	13,92	6,46	8,00		
FTI	MB	4	5,00	2,71	4,00	3	9
	T. Elektro	47	17,85	13,91	15,00	1	60
	T. Fisika	21	14,71	10,49	13,00	2	42
	T. Industri	19	7,63	7,19	6,00	0	30
	T. Kimia	34	20,35	19,00	14,50	3	86
	T. Mesin	39	7,36	5,40	5,00	1	22
	T. Material	13	11,92	10,63	9,00	1	33
TMJ	13	26,90	41,90	8,00	0	145	
Total		190	13,96	12,39	9,50		
FTIf	SI	24	6,25	4,96	5,00	1	19
	T. Informatika	38	12,35	10,27	10,00	2	42
Total		62	9,30	3,75	7,00		
FTSP	Arsitektur	9	4,78	4,32	1,00	1	15
	D. Interior	1	2,00	*	2,00	2	2
	D. Produk	2	4,00	1,41	3,00	3	5
	PWK	2	2,00	0,00	2,00	2	2
	T. Geofisika	5	5,40	2,30	5,00	3	9
	T. Geomatika	10	4,56	2,40	4,00	2	8
	T. Lingkungan	11	10,64	6,89	9,00	2	26
T. Sipil	35	9,26	9,56	6,00	0	35	
Total		75	5,33	3,35	5,00		
FTK	T. Kelautan	20	8,60	10,58	5,00	1	48
	T. Perkapalan	9	8,63	5,24	9,00	3	19
	T. Siskal	12	5,67	3,77	4,50	2	14
	Transla	2	5,00	0,00	5,00	5	5
Total		43	6,98	4,38	5,00		
Total ITS		457	9,90	3,73	7,00		

12. Nilai Korelasi Antar Variabel.

```
> cor(X, Y,)
      Y1   Y2   X1   X2   X3   X4
Y1 1.00000000 0.82944897 0.01942708 0.03999361 0.57926326 0.5782404
Y2 0.82944897 1.00000000 0.07344532 0.10756414 0.69416754 0.6851080
X1 0.01942708 0.07344532 1.00000000 0.96191672 0.08891851 0.1203620
X2 0.03999361 0.10756414 0.96191672 1.00000000 0.11356948 0.1467870
X3 0.57926326 0.69416754 0.08891851 0.11356948 1.00000000 0.9064029
X4 0.57824037 0.68510803 0.12036198 0.14678702 0.90640288 1.0000000
```

13. Hasil Uji Korelasi

```
> cor.test(Sitasi, Usia)
Pearson's product-moment correlation
data: Sitasi and Usia
t = 0.41447, df = 455, p-value = 0.6787
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.07242909 0.11095646
sample estimates:
cor
0.01942708
```

```
> cor.test(Sitasi, Lama_bekerja)
Pearson's product-moment correlation
data: Sitasi and Lama_bekerja
t = 0.85378, df = 455, p-value = 0.3937
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.05192397 0.13123926
sample estimates:
cor
0.03999361
```

Lanjutan : Hasil Uji Korelasi.

```
> cor.test(Sitasi, Jumlah_Dokumen)
```

Pearson's product-moment correlation

```
data: Sitasi and Jumlah_Dokumen
t = 15.158, df = 455, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5148946 0.6371367
sample estimates:
   cor
0.5792633
```

```
> cor.test(Sitasi, Jumlah_coauthor)
```

Pearson's product-moment correlation

```
data: Sitasi and Jumlah_coauthor
t = 15.118, df = 455, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5137634 0.6362221
sample estimates:
   cor
0.5782404
```

```
> cor.test(Indeks_h, Usia)
```

Pearson's product-moment correlation

```
data: Indeks_h and Usia
t = 1.5709, df = 455, p-value = 0.1169
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.01840578 0.16406711
sample estimates:
   cor
0.07344532
```

```
> cor.test(Indeks_h, Lama_bekerja)
```

Pearson's product-moment correlation

```
data: Indeks_h and Lama_bekerja
t = 2.3078, df = 455, p-value = 0.02146
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.01599485 0.19734413
sample estimates:
   cor
0.1075641
```

```
> cor.test(Indeks_h, Jumlah_Dokumen)
```

Pearson's product-moment correlation

```
data: Indeks_h and Jumlah_Dokumen
t = 20.571, df = 455, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6434088 0.7388492
sample estimates:
   cor
0.6941675
```

```
> cor.test(Indeks_h, Jumlah_coauthor)
```

Pearson's product-moment correlation

```
data: Indeks_h and Jumlah_coauthor
t = 20.062, df = 455, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.6331713 0.7309030
sample estimates:
   cor
0.685108
```

```
> cor.test(Usia, Lama_bekerja)
```

Pearson's product-moment correlation

```
data: Usia and Lama_bekerja
t = 75.065, df = 455, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.9544000 0.9682145
sample estimates:
   cor
0.9619167
```

```
> cor.test(Usia, Jumlah_Dokumen)
```

Pearson's product-moment correlation

```
data: Usia and Jumlah_Dokumen
t = 1.9042, df = 455, p-value = 0.05751
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.002831683 0.179184147
sample estimates:
   cor
0.08891851
```

```
> cor.test(Usia, Jumlah_coauthor)
```

Pearson's product-moment correlation

```
data: Usia and Jumlah_coauthor
t = 2.5862, df = 455, p-value = 0.01001
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.02895455 0.20977309
sample estimates:
   cor
0.120362
```

```
> cor.test(Lama_bekerja, Jumlah_Dokumen)
```

Pearson's product-moment correlation

```
data: Lama_bekerja and Jumlah_Dokumen
t = 2.4383, df = 455, p-value = 0.01514
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.02207231 0.20317997
sample estimates:
   cor
```

Lampiran V. Regresi Kuantil.

1. Estimasi Parameter Regresi Kuantil pada Setiap kuantil untuk Variabel Dependen Sitasi.

	5%	10%	25%	50%	75%	90%	95%
Konstan	-7.669	-5.718	-1.179	-0.4849	-1.1334	12.110	23.859
X ₁	0.144	0.035	-0.059	-0.0443	-0.1035	-0.445	0.316
X ₂	-0.119	0.032	0.082	0.0222	0.0701	0.328	-0.316
X ₃	1.198	1.202	1.089	1.4880	4.3489	7.719	9.626
X ₄	-0.232	-0.169	0.093	0.4209	0.5175	1.194	1.579
X ₅	0.561	0.629	0.774	0.3101	0.1204	2.575	0.947
X ₆₍₁₎	-0.385	-0.141	-0.473	-0.3699	0.7975	-1.712	-15.942
X ₆₍₂₎	0.811	0.095	-0.098	0.2215	0.2141	2.107	-2.842
X ₆₍₃₎	-0.097	0.193	-0.223	0.4431	0.4907	3.137	-0.632
X ₇	0.085	1.013	0.239	-0.8924	-0.4977	0.211	-0.632
X ₈	0.490	0.345	-0.087	-0.2215	-1.0985	-10.043	-42.643

2. Estimasi Parameter Regresi Kuantil pada Setiap kuantil untuk Variabel Dependen Indeks h.

	5%	10%	25%	50%	75%	90%	95%
Konstan	-0.1724	0.1677	0.5429	0.7526	1.068	1.065	1.420
Y ₁	0.0141	0.0138	0.0164	0.0202	0.026	0.035	0.039
X ₁	0.0024	-0.0092	-0.0024	-0.0041	0.002	0.010	0.009
X ₂	-0.0032	0.0045	0.0022	0.0045	0.002	-0.004	0.001
X ₃	0.0425	0.0548	0.0341	0.0803	0.058	0.061	0.078
X ₄	-0.0093	0.0022	0.0096	0.0086	0.017	0.016	0.005
X ₅	0.0307	0.0296	0.0152	0.0280	-0.028	0.021	-0.254
X ₆₍₁₎	0.0778	0.1697	0.1692	0.0469	-0.110	-0.302	-0.358
X ₆₍₂₎	0.0451	0.0948	-0.0018	0.0098	-0.045	0.080	-0.312
X ₆₍₃₎	0.0314	0.0829	0.0190	-0.0045	-0.007	-0.145	-0.416
X ₇	-0.0190	-0.0571	-0.5594	-0.7506	-0.443	-0.406	-0.160
X ₈	0.0353	-0.0038	0.0055	0.0041	-0.081	-0.116	-0.031

Lampiran VI. Langkah Analisis Menggunakan Stata 12.

