

**APLIKASI REGRESI SURVIVAL BUCKLEY-JAMES UNTUK  
MENGATASI DATA OUTLIER YANG DISENSOR**

**(Studi Kasus : Data Rekam Medis Pasien *Tuberculosis* Rawat Inap Tahun  
2018-2019 di RSUD Soehadi Prijonegoro Sragen)**

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana Jurusan  
Statistika



**Erzylia Herlin Brilliant**

**16611077**

**PROGRAM STUDI STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS ISLAM INDONESIA  
YOGYAKARTA**

**2020**

**HALAMAN PERSETUJUAN PEMBIMBING**

**TUGAS AKHIR**

Judul : Aplikasi Regresi Survival Buckley-James Untuk Mengatasi Data Outlier yang Disensor

Nama Mahasiswa : Erzylia Herlin Brilliant


NIM : 16611077

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

Yogyakarta, 16 Januari 2020

**Pembimbing**

الجامعة الإسلامية  
الاندونيسية



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

**HALAMAN PENGESAHAN**

**TUGAS AKHIR**

**APLIKASI REGRESI SURVIVAL BUCKLEY-JAMES UNTUK  
MENGATASI DATA OUTLIER YANG DISENSOR**

**Nama Mahasiswa : Erzylia Herlin Brilliant**

**NIM : 16611077**

**TUGAS AKHIR INI TELAH DIUJIKAN  
PADA TANGGAL : 31 JANUARI 2020**

**Nama Penguji:**

1. Ir. Ali Parkhan, M.T.
2. Mujiati Dwi Kartikasari, S.Si., M.Sc.
3. Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc.

**Tanda Tangan**



البحث العلمى  
Mengetahui,

Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



(P. W. B.yanto, S.Pd., M.Si., Ph.D.)

## KATA PENGANTAR

*Assalamu'alaikum Wr.Wb*

Dengan mengucapkan syukur *Alhamdulillah*, penulis panjatkan puji syukur kehadirat Allah SWT karena atas rahmat dan karunia-Nya penulis dapat menyelesaikan Tugas Akhir yang berjudul “APLIKASI REGRESI SURVIVAL BUCKLEY-JAMES UNTUK MENGATASI DATA OUTLIER YANG DISENSOR” ini dengan baik. Shalawat serta salam penulis curahkan kepada junjungan Nabi Muhammad SAW beserta keluarga, sahabat dan umatnya.

Dalam penyelesaian Tugas Akhir ini, tidak lepas dari banyak pihak yang telah memberi bantuan yang berupa saran, kritik, bimbingan, maupun bantuan lainnya. Untuk itu, pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Keluarga yaitu Ayah, Ibu, dan Adik yang selalu mendukung dan memberikan doa yang tiada henti serta memberi motivasi kepada penulis untuk menyelesaikan tugas akhir.
2. Bapak Muhammad Hasan Sidiq Kurniawan, S.Si., M.Sc., selaku dosen pembimbing Tugas Akhir yang dengan sabar telah meluangkan waktu untuk membimbing dan mengarahkan penulis dalam penyusunan tugas akhir ini.
3. Bapak Dr. Edy Widodo, S.Si., M.Si., selaku Ketua Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Indonesia atas dukungan beliau mewujudkan pelaksanaan kerja praktik untuk mahasiswa/i Statistika FMIPA UII.
4. Seluruh dosen Program Studi Statistika yang telah memberikan arahan, bantuan, bimbingan serta ilmu di dalam maupun di luar lingkungan perkuliahan.
5. Bapak dr. Sunaryo, Sp.THT, selaku Direktur Utama RSUD Soehadi Prijonegoro Sragen yang telah memberikan izin untuk melaksanakan penelitian di RSUD Soehadi Prijonegoro Sragen.
6. Ibu dr. Indarsih selaku Kepala Bagian Diklat yang turut menerima dan memberikan bimbingan dengan baik selama melaksanakan penelitian.

7. Bapak dr. Imron Riyatno, Sp.P, M.Kes dan dr. Windu Prasetyo, Sp.P, selaku dokter spesialis paru di RSUD Soehadi Prijonegoro Sragen yang turut membantu dan memberikan banyak ilmu kepada peneliti.
8. Ibu Eri Kusumawati selaku staf rekam medis yang sangat membantu dalam proses penelitian dan pengambilan data sekunder.
9. Barlinda Titania, Syinta Nuri Mashita dan Iin Mutmainnah yang menjadi sahabat seperjuangan yang selalu mengingatkan kebaikan serta memberi dukungan satu sama lain.
10. Teman-teman yang selalu mendukung dan menemani selama kuliah yaitu Fauziah, Rizki Desi, Suharmi, Nur Anisa, Alfi Indah, Rinandha, Yosi, Ika Meiza dan Ria Amelia.
11. Semua pihak lain yang tidak dapat penulis sebutkan satu per satu. Terima kasih atas segala bantuan, dukungan, dan doa yang diberikan kepada penulis.

Akhir kata, semoga Allah SWT membalas segala kebaikan dari semua pihak yang telah mendampingi dan membantu hingga saat ini. Penulis berharap tugas akhir ini dapat bermanfaat .....

*Wassalamu'alaikum Wr.Wb*

Yogyakarta, Januari 2020

(Erzylia Herlin Brilliant)

## DAFTAR ISI

Halaman Judul.....	i
HALAMAN PERSETUJUAN PEMBIMBING .....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR .....	iv
DAFTAR ISI.....	vi
DAFTAR TABEL.....	viii
DAFTAR GAMBAR .....	ix
DAFTAR LAMPIRAN.....	x
PERNYATAAN.....	xi
ABSTRAK.....	xii
ABSTRACT.....	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah .....	4
1.4 Tujuan Penelitian.....	5
1.5 Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
BAB III LANDASAN TEORI.....	9
3.1 Sejarah Tuberkulosis .....	9
BAB IV METODE PENELITIAN .....	42
4.1 Populasi dan Sampel.....	42
BAB V PEMBAHASAN.....	49
5.1 Analisis Deskriptif.....	49
5.2 Kurva Survival Kaplan-Meier .....	55
5.3 Uji Outlier.....	58
5.4 Analisis Regresi Linier Berganda.....	60
5.5 Analisis Regresi Buckley-James .....	63

5.6	Perbandingan Metode Regresi Linier Berganda dengan Regresi <i>Buckley-James</i> .....	72
5.7	Koefisien Determinasi Regresi <i>Buckley-James</i> .....	73
BAB VI	PENUTUP .....	75
6.1	Kesimpulan.....	75
6.2	Saran .....	75
DAFTAR PUSTAKA	.....	77
Lampiran 1	Data Rekam Medis Pasien TB Rawat Inap Tahun 2018-2019 di RSUD Soehadi Prijonegoro Sragen .....	84
Lampiran 2	Syntax R .....	95
Lampiran 3	<i>Output R</i> Analisis Regresi Linier Berganda .....	97
Lampiran 4	<i>Output R</i> Analisis Regresi <i>Buckley-James</i> .....	102
Lampiran 5	Peramalan Nilai $y_i$ .....	104
Lampiran 6	Surat Ijin Pengambilan Data di RSUD Soehadi Prijonegoro Sragen .....	105

## DAFTAR TABEL

<b>Tabel 4.1.</b> Keterangan Variabel.....	43
<b>Tabel 5.1.</b> Fungsi Survival Estimasi Kaplan-Meier .....	55
<b>Tabel 5.2.</b> Statistik Uji, Keputusan dan Kesimpulan Uji Parsial Regresi Linier Berganda .....	61
<b>Tabel 5.3.</b> Statistik Uji, Keputusan dan Kesimpulan Uji Parsial Regresi Linier Berganda Setelah Proses Eliminasi.....	62
<b>Tabel 5.4.</b> Statistik uji, keputusan dan kesimpulan uji signifikansi parameter regresi.....	64
<b>Tabel 5.5.</b> Statistik uji, keputusan dan kesimpulan uji signifikansi parameter Regresi Buckley-James .....	66
<b>Tabel 5.6.</b> Perbandingan MAPE.....	72



## DAFTAR GAMBAR

<b>Gambar 3.1.</b> Pengaruh Jumlah Infeksi Bakteri dan Kekuatan Daya Tahan Tubuh .....	12
<b>Gambar 3.2.</b> <i>Boxplot</i> .....	20
<b>Gambar 3.3.</b> <i>Q-Q plot</i> .....	21
<b>Gambar 3.4.</b> Grafik Fungsi Survival .....	23
<b>Gambar 3.5.</b> Estimasi Fungsi $S(t)$ .....	24
<b>Gambar 3.6.</b> Ilustrasi Data Tersensor Kanan .....	25
<b>Gambar 4.1.</b> <i>Flowchart</i> Alur Penelitian .....	47
<b>Gambar 4.2.</b> <i>Flowchart</i> Alur Penelitian .....	48
<b>Gambar 5.1.</b> Histogram Usia Pasien <i>Tuberculosis</i> .....	49
<b>Gambar 5.2.</b> Persentase Jenis Kelamin Pasien <i>Tuberculosis</i> .....	50
<b>Gambar 5.3.</b> Jumlah Pasien Memiliki Komplikasi .....	51
<b>Gambar 5.4.</b> Jumlah Pasien yang Memiliki Penyakit Lain .....	51
<b>Gambar 5.5.</b> <i>Pie Chart</i> Pendidikan Pasien .....	52
<b>Gambar 5.6.</b> <i>Pie Chart</i> Pekerjaan Pasien .....	53
<b>Gambar 5.7.</b> Jumlah Kasus Lama / Baru Pasien <i>Tuberculosis</i> .....	54
<b>Gambar 5.8.</b> Diagnosa Penyakit Pasien <i>Tuberculosis</i> .....	54
<b>Gambar 5.9.</b> Kurva Estimasi Kaplan-Meier .....	57
<b>Gambar 5.10.</b> Kurva Estimasi Kaplan-Meier .....	57
<b>Gambar 5.11.</b> Uji <i>Outlier Boxplot</i> .....	58
<b>Gambar 5.12.</b> Uji <i>Outlier Q-Q Plot</i> .....	59

## DAFTAR LAMPIRAN

Lampiran 1 Data Rekam Medis Pasien TB Rawat Inap Tahun 2018-2019 di RSUD Soehadi Prijonegoro Sragen .....	84
Lampiran 2 Syntax R .....	95
Lampiran 3 <i>Output R</i> Analisis Regresi Linier Berganda .....	97
Lampiran 4 <i>Output R</i> Analisis Regresi Buckley-James .....	102
Lampiran 5 Peramalan Nilai $y_i$ .....	104
Lampiran 6 Surat Ijin Pengambilan Data di RSUD Soehadi Prijonegoro Sragen .....	105



## PERNYATAAN

APLIKASI BERGEMER SORAWAL BEKREZY-JAMES UNTUK

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

Yogyakarta, 16 Januari 2020



(Erzylia Herlin Brilliant)

## ABSTRAK

### APLIKASI REGRESI SURVIVAL BUCKLEY-JAMES UNTUK MENGATASI DATA OUTLIER YANG DISENSOR

Erzylia Herlin Brilliant

Program Studi Statistika

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Islam Indonesia

Analisis survival atau analisis uji hidup adalah teknik statistik yang digunakan untuk menganalisis data yang bertujuan untuk mengetahui hasil dari variabel yang mempengaruhi suatu awal kejadian sampai akhir kejadian. Analisis regresi linear merupakan analisis yang bertujuan untuk mengetahui pengaruh dan hubungan linear variabel independen terhadap variabel dependen. Model regresi linear biasa tidak dapat digunakan untuk memodelkan data survival karena adanya data yang tersensor kanan. Jika dipaksakan untuk digunakan, regresi linear biasa akan memberikan hasil yang kurang akurat, karena data tersensor merupakan data yang diperoleh dari observasi yang tidak lengkap. Analisis Regresi Linear Berganda dapat digunakan apabila terdapat outlier namun kurang mampu memberikan hasil yang tepat. Data outlier tersebut yang akan dijadikan sebagai data tersensor. Untuk menganalisis data tersensor tersebut digunakan analisis Regresi Buckley-James. Pada penelitian ini, dilakukan studi kasus tentang faktor-faktor yang mempengaruhi lama waktu sembuh pasien Tuberkulosis. Faktor-faktor yang diduga mempengaruhi lama waktu sembuh pasien Tuberkulosis yaitu jenis kelamin, usia pasien, komplikasi, penyakit lain, pendidikan, pekerjaan, kasus, dan diagnosa. Dari hasil analisis memberikan kesimpulan bahwa metode Regresi Buckley-James memiliki nilai MAPE lebih kecil dibanding dengan metode Regresi Linier Berganda. Sehingga Regresi Buckley-James dapat lebih akurat digunakan pada data survival yang mengandung data tersensor.

**Kata Kunci:** Survival, Regresi Linear Berganda, Regresi Buckley-James, *Outlier*, Data Tersensor, Tuberkulosis

## ABSTRACT

### SURVIVAL BUCKLEY-JAMES REGRESSION APPLICATIONS TO RESOLVE CENSORED OUTLIER DATA

Erzylia Herlin Brilliant

Program Studi Statistika

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Islam Indonesia

*Survival analysis or life test analysis is a statistical technique used to analyze data that aims to find out the results of variables that affect an initial event to the end of an event. Linear regression analysis is an analysis that aims to determine the effect and linear relationship of independent variables on the dependent variable. Ordinary linear regression models cannot be used to model survival data because of the existence of right censored data. If forced to use, ordinary linear regression will give less accurate results, because censored data is data obtained from incomplete observations. Multiple Linear Regression Analysis can be used if there are outliers but are unable to provide the right results. Outlier data will be used as censored data. To analyze the censored data Buckley-James Regression analysis was used. In this research, a case study of factors that influence the duration of recovery of a tuberculosis patient is carried out. Factors that are thought to influence the length of time of a Tuberculosis patient's recovery are gender, patient age, complications, other diseases, education, occupation, cases, and diagnosis. From the results of the analysis concluded that the Buckley-James Regression method has a smaller MAPE value compared to the Multiple Linear Regression method. So that the Buckley-James Regression can be more accurately used in survival data containing censored data.*

**Keywords:** *Survival, Multiple Linear Regression, Buckley-James Regression, Outlier, Censored Data, Tuberculosis*