

SKRIPSI

PERBANDINGAN KEMAMPUAN MODEL REGRESI LINEAR
INVERS GAUSSIAN DAN MODEL REGRESI LINEAR
NORMAL DALAM BEBERAPA JENIS DISTRIBUSI DATA



MATTHEW CLAUDIUS SUSANTIO

NPM: 6161901024

PROGRAM STUDI MATEMATIKA
FAKULTAS TEKNOLOGI INFORMASI DAN SAINS
UNIVERSITAS KATOLIK PARAHYANGAN
2023

FINAL PROJECT

CAPABILITIES COMPARISON OF THE INVERSE GAUSSIAN
LINEAR REGRESSION MODEL AND NORMAL LINEAR
REGRESSION MODEL IN SOME KIND OF DATA
DISTRIBUTION



MATTHEW CLAUDIUS SUSANTIO

NPM: 6161901024

DEPARTMENT OF MATHEMATICS
FACULTY OF INFORMATION TECHNOLOGY AND SCIENCES
PARAHYANGAN CATHOLIC UNIVERSITY
2023

LEMBAR PENGESAHAN

PERBANDINGAN KEMAMPUAN MODEL REGRESI LINEAR INVERS GAUSSIAN DAN MODEL REGRESI LINEAR NORMAL DALAM BEBERAPA JENIS DISTRIBUSI DATA

Matthew Claudius Susantio

NPM: 6161901024

Bandung, 19 Agustus 2023

Menyetujui,

Pembimbing 1

Pembimbing 2


Prof. M. Weno Setya Budhi


Rizky Reza Fauzi, D.Phil.Math.

Ketua Penguji

Anggota Penguji


Farah Kristiani, Ph.D.


Robyn Irawan, M.Sc.

Mengetahui,

Ketua Program Studi

Dr. Livia Owen

PERNYATAAN

Dengan ini saya yang bertandatangan di bawah ini menyatakan bahwa skripsi dengan judul:

PERBANDINGAN KEMAMPUAN MODEL REGRESI LINEAR INVERS GAUSSIAN DAN MODEL REGRESI LINEAR NORMAL DALAM BEBERAPA JENIS DISTRIBUSI DATA

adalah benar-benar karya saya sendiri, dan saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika keilmuan yang berlaku dalam masyarakat keilmuan.

Atas pernyataan ini, saya siap menanggung segala risiko dan sanksi yang dijatuhkan kepada saya, apabila di kemudian hari ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya, atau jika ada tuntutan formal atau non-formal dari pihak lain berkaitan dengan keaslian karya saya ini.

Dinyatakan di Bandung,
19 Agustus 2023

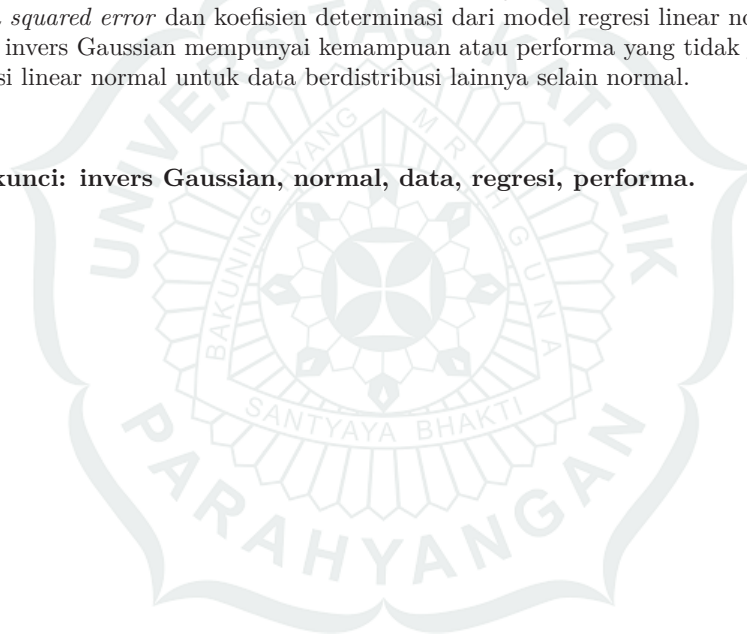


Matthew Claudius Susantio
NPM: 6161901024

ABSTRAK

Penelitian ini dilakukan untuk melihat kemampuan dari model regresi linear invers Gaussian sebagai model regresi untuk data yang berdistribusi normal dan lain-lain dengan menggunakan model regresi linear normal sebagai tolak ukurnya. Hal tersebut merupakan suatu langkah yang dapat digunakan dalam menyelesaikan masalah data secara umum, yaitu tidak berbentuk simetris, tetapi *skewed*. Kemampuan dari model regresi linear invers Gaussian dapat dilihat dari *mean squared error* dan koefisien determinasi dari model regresi. Kemudian akan dibandingkan dengan *mean squared error* dan koefisien determinasi dari model regresi linear normal. Model regresi linear invers Gaussian mempunyai kemampuan atau performa yang tidak jauh berbeda dengan regresi linear normal untuk data berdistribusi lainnya selain normal.

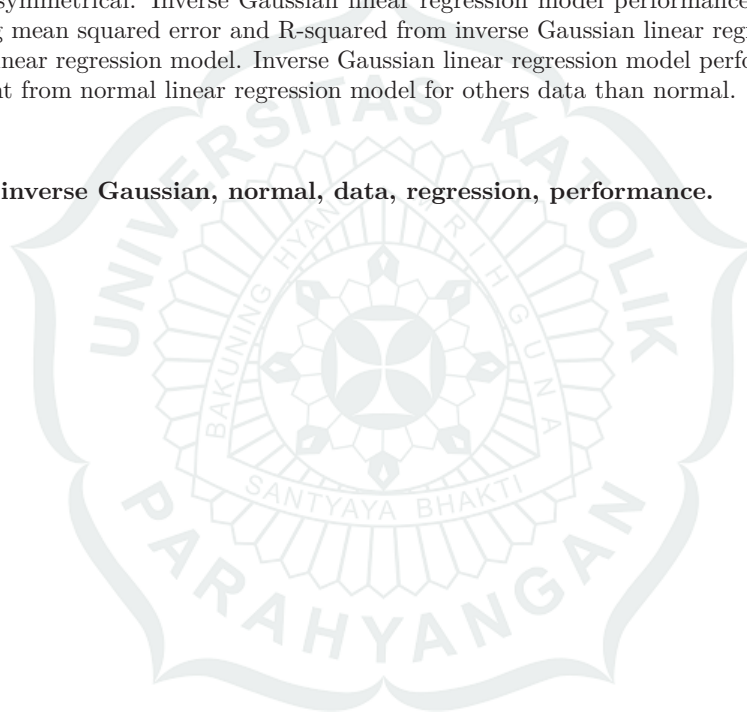
Kata-kata kunci: invers Gaussian, normal, data, regresi, performa.



ABSTRACT

This research goal is to see the performance of inverse Gaussian linear regression model as a regression model for normal distribution data and others by using normal linear regression model as a benchmark. This model is a step that can be used to solve data problems in general, which skewed, not symmetrical. Inverse Gaussian linear regression model performance's can be see by comparing mean squared error and R-squared from inverse Gaussian linear regression model and normal linear regression model. Inverse Gaussian linear regression model performance's not much different from normal linear regression model for others data than normal.

Keywords: inverse Gaussian, normal, data, regression, performance.



KATA PENGANTAR

Segala puji dan syukur penulis ucapkan kepada Tuhan Yang Maha Esa karena atas anugerah, berkat, kasih, dan curahan Roh Kudus-Nya, penulis diberi kesempatan untuk belajar dan menyelesaikan skripsi yang berjudul **“Perbandingan Kemampuan Model Regresi Linear Invers Gaussian dan Model Regresi Linear Normal dalam Beberapa Jenis Distribusi Data”**. Selama masa perkuliahan di Universitas Katolik Parahyangan dan penulisan skripsi ini, penulis juga selalu memperoleh ilmu baru, dukungan, dan bantuan tulus dari berbagai pihak yang luar biasa. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang tulus kepada:

1. Keluarga, yaitu papi, mami, dan dede atas doa dan dukungan kepada penulis. Terima kasih karena telah berjuang dan selalu mengusahakan memberi bantuan yang terbaik kepada penulis. Terima kasih karena selalu percaya kepada penulis bahwa bisa melakukan segala hal positif walaupun kadang sifat pesimis yang penulis miliki muncul. Terima kasih atas kasih sayang yang selalu keluarga berikan kepada penulis. Semoga kita selalu menjadi keluarga yang saling mendukung satu sama lain apapun kondisi yang sedang di hadapi. Penulis bersyukur mempunyai keluarga yang sangat peduli.
2. Bapak Prof. M. Wono Setya Budhi selaku dosen pembimbing yang sangat berwawasan dan penulis hormati. Terima kasih atas ilmu, bimbingan, didikan, dan perhatiannya yang diberikan selama ini. Penulis banyak belajar dari bapak dan termotivasi untuk menjadi versi terbaik dari diri penulis. Setiap saran dari bapak akan penulis terima dan terapkan ke depannya. Mohon maaf atas kesalahan penulis dan semoga Bapak selalu diberikan kesehatan dan suka cita setiap harinya.
3. Bapak Rizky Reza Fauzi, D.Phil.Math. selaku dosen pembimbing yang sungguh penuh kejutan serta sangat penulis kagumi. Terima kasih karena selalu berdedikasi dan menjadi penolong di setiap langkah penulis dalam penyusunan skripsi ini. Bapak yang sangat baik dan sabar, dengan wibawanya selalu mampu membantu anak bimbingannya keluar dari kebingungan dan ketidaktahuan. Penulis bersyukur karena bisa menjadi anak bimbingan dari bapak yang selalu sabar menghadapi berbagai sikap penulis. Semoga Bapak dapat terus menjadi terang bagi orang-orang di sekitar.
4. Ibu Farah Kristiani, Ph.D. dan Bapak Robyn Irawan, M.Sc. selaku dosen penguji yang juga merupakan orang-orang hebat. Terima kasih karena ibu dan bapak telah meluangkan waktu dan memberikan kritik serta saran yang sangat berguna bagi penulis. Penulis bersyukur karena skripsi ini dapat diuji serta dinyatakan lulus sidang oleh Ibu dan Bapak. Semoga Ibu dan Bapak selalu dilimpahkan berkat oleh Tuhan.
5. Anggota grup Kardus sebagai teman-teman belajar dari penulis selama berkuliah dari semester awal hingga akhir. Terima kasih sudah menerima penulis ke dalam grup belajar dan menjadi tempat saran ketika penulis mengalami masa-masa sulit selama menjadi mahasiswa. Banyak momen yang sudah kita bentuk selama 4 tahun bersama. Mari kita tetap menjadi teman seterusnya dan penulis doakan yang terbaik selalu untuk kalian di manapun berada.
6. Reynaldi Anggara sebagai teman seperbimbingan yang selalu mendengarkan dan membantu masalah penulis ketika sedang mengerjakan skripsi. Terima kasih atas kebersamaan kita

selama bimbingan skripsi ini. Penulis doakan agar setiap langkah ke depannya selalu menjadi berkat bagi orang lain.

7. Stephanie Stella sebagai teman seperbimbingan dengan Pak Rizky yang membantu penulis ketika ada materi yang penulis tidak miliki. Terima kasih atas bantuannya kepada penulis saat meminjamkan buku catatan pribadi kepada penulis. Kebaikan tersebut sangat penulis hargai karena membantu dalam menyelesaikan masalah yang ada di skripsi. Penulis doakan semoga selalu diberi kesehatan dan kesuksesan ke depannya.
8. Teman-teman Matematika UNPAR angkatan 2019 yang tidak dapat diucapkan satu per satu. Terima kasih atas kerja samanya selama 4 tahun perkuliahan penulis walaupun perjalanan kita sempat terhalang dengan wabah Covid-19. Banyak sekali momen-momen asik, seru, dan lucu saat bersama kalian yang tidak akan penulis lupakan. Mari berjumpa kembali di versi terbaik dari diri kita.
9. Seluruh dosen dan staf Matematika Universitas Katolik Parahyangan yang luar biasa. Terima kasih atas ilmu dan semangatnya dalam mengajar. Penulis banyak belajar hal-hal baru dari Bapak dan Ibu selama perkuliahan. Semoga semangat mengajar bapak dan ibu tidak pernah pudar dan selalu dilimpahkan kesabaran, kebijaksanaan, dan kebaikan meliputi setiap karya-karya bapak dan ibu.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari kesempurnaan karena keterbatasan yang ada. Untuk itu, penulis sangat menghargai adanya dukungan dan sumbangsih pikiran berupa kritik dan saran yang membangun. Selama penulisan skripsi ini, penulis banyak sekali memperoleh manfaat atas bertambahnya wawasan dan kesempatan bertemu dengan orang-orang istimewa. Dengan demikian, penulis berharap agar skripsi ini juga dapat menjadi berkat dan manfaat bagi segala pihak yang membacanya.

Bandung, 19 Agustus 2023

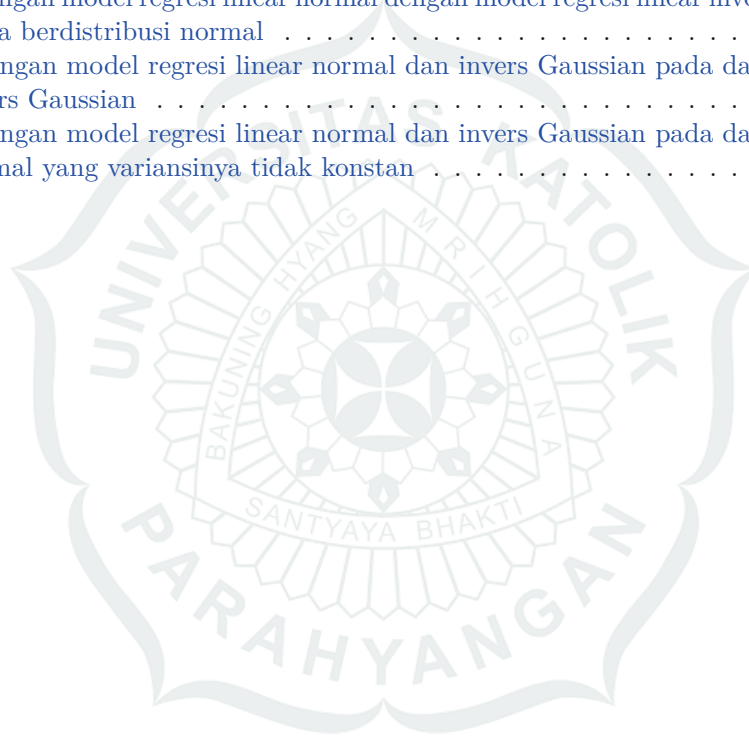
Penulis

DAFTAR ISI

KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xi
DAFTAR TABEL	xii
1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	1
1.3 Tujuan	2
1.4 <i>State of the Art</i>	2
1.5 Batasan Masalah	2
1.6 Sistematika Pembahasan	2
2 LANDASAN TEORI	4
2.1 Distribusi Normal	4
2.2 Distribusi Invers Gaussian	7
2.3 Metode Maksimum <i>Likelihood</i>	10
2.4 Regresi Linear	11
3 REGRESI LINEAR INVERS GAUSSIAN	13
3.1 Taksiran Parameter Regresi Linear Sederhana Distribusi Invers Gaussian	13
4 ANALISIS PERBANDINGAN MODEL REGRESI MENGGUNAKAN DATA SIMULASI	16
4.1 Simulasi Data	16
4.1.1 Data Berdistribusi Normal Dengan Variansi Konstan	16
4.1.2 Data Berdistribusi Invers Gaussian	17
4.1.3 Data Berdistribusi Normal dengan Variansi Tidak Konstan	17
4.2 Hasil dan Analisis Data	18
4.2.1 Hasil dan Analisis Data Berdistribusi Normal	18
4.2.2 Hasil dan Analisis Data Berdistribusi Invers Gaussian	20
4.2.3 Hasil dan Analisis Data Berdistribusi Normal dengan Variansi Tidak Konstan	22
5 KESIMPULAN DAN SARAN	25
5.1 Kesimpulan	25
5.2 Saran	25
DAFTAR REFERENSI	26
A DATA SIMULASI	27

DAFTAR GAMBAR

2.1	Visualisasi data acak dan pdf distribusi normal	6
2.2	Distribusi Normal dengan Parameter μ dan σ berbeda-beda	7
2.3	Grafik Distribusi Invers Gaussian	9
2.4	Parameter μ dan λ berbeda-beda	10
4.1	Perbandingan model regresi linear normal dengan model regresi linear invers Gaussian pada data berdistribusi normal	19
4.2	Perbandingan model regresi linear normal dan invers Gaussian pada data berdistribusi invers Gaussian	21
4.3	Perbandingan model regresi linear normal dan invers Gaussian pada data berdistribusi normal yang variansinya tidak konstan	23



DAFTAR TABEL

4.1	Data Berdistribusi Normal	16
4.2	Data Berdistribusi Invers Gaussian	17
4.3	Data Berdistribusi Normal dengan Variansi yang Tidak Konstan	17



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Dalam penelitian ini, akan dibahas mengenai distribusi invers Gaussian. Pada awalnya, distribusi invers Gaussian bernama distribusi waktu kunjung pertama. Dari invers fungsi pembangkit kumulatif atau invers logaritma dari fungsi pembangkit momen distribusi normal, terbentuk fungsi pembangkit kumulatif distribusi waktu kunjung pertama yang diberi nama distribusi invers Gaussian oleh Tweedie. Distribusi invers Gaussian dapat digunakan dalam memodelkan berbagai pola data kehidupan atau jumlah kejadian dalam berbagai bidang. Sebagai contoh, penggunaan model invers Gaussian untuk kekosongan bendungan, analisis regresi, mengaplikasikan model invers Gaussian ke dalam studi *purchase incidence models*, menggunakan model invers Gaussian untuk model *duration of strikes*, serta menggunakannya dalam distribusi waktu yang dihabiskan untuk zat suntik yang dilabelkan bernama *tracer*, yang berada di dalam sistem biologis [1].

Pada tahun 1977, Chhikara dan Folks mengatakan bahwa teori distribusi sampling dari invers Gaussian muncul sebagai kandidat kuat dalam menyelesaikan beberapa kasus, sehingga lebih disarankan untuk menggunakan distribusi invers Gaussian dibandingkan distribusi log-normal [1]. Distribusi invers Gaussian dapat menampung berbagai jenis bentuk data *skewed-right* hingga hampir berbentuk normal, serta disebut unik di antara distribusi lainnya untuk data *skewed-right* positif seperti Weibull, Gamma, dan log-Normal [2].

Data yang berdistribusi invers Gaussian tersebut akan dicari tahu mengenai hubungan antara variabel bebas dan variabel terikatnya. Penggunaan analisis regresi sebagai teknik statistik yang digunakan untuk mencari tahu dan memodelkan hubungan antar variabel [3]. Berbagai masalah sederhana dalam kehidupan sehari-hari yang dapat diselesaikan dengan analisis regresi. Contoh kasusnya, seperti seorang yang bekerja sebagai karyawan minuman bersoda di dalam bidang teknik industri sedang menganalisis pengiriman barang dan layanan operasi untuk *vending machines* atau mesin penjual otomatis dengan menggunakan waktu antar dan volume produk yang diantar sebagai variabel. Di dalam regresi hanya ada satu fokus, yaitu variabel terikat atau dependen [4].

1.2 Rumusan Masalah

1. Bagaimana karakteristik distribusi invers Gaussian?
2. Bagaimana proses pemodelan regresi linear invers Gaussian?
3. Bagaimana perbandingan model regresi normal dengan invers Gaussian?

1.3 Tujuan

1. Mengevaluasi fungsi distribusi, rata-rata, dan variansi dari distribusi invers Gaussian.
2. Mengonstruksi model regresi linear invers Gaussian.
3. Menganalisis perbandingan model regresi linear normal dengan invers Gaussian.

1.4 *State of the Art*

Model regresi linear normal dapat digunakan jika data berdistribusi normal dan mempunyai variansi yang konstan [5]. Sementara untuk model regresi linear invers Gaussian penggunaannya lebih fleksibel karena variansinya tidak harus konstan [1]. Secara umum, banyak data yang pada praktik tidak simetris atau berdistribusi normal, tetapi *skewed*, sehingga model regresi linear invers Gaussian menjadi langkah yang tepat untuk mengakomodasi permasalahan persebaran data *skewed*.

1.5 Batasan Masalah

1. Pada penelitian ini, model regresi yang digunakan adalah model regresi *no intercept*.
2. Data yang digunakan dalam penelitian ini berupa data buatan atau simulasi sendiri yang dilakukan oleh peneliti.

1.6 Sistematika Pembahasan

Dalam penelitian ini, akan digunakan sistematika pembahasan yang terdiri dari lima bab, yaitu:

Bab I: Pendahuluan

Bab ini menguraikan penjelasan awal mengenai hal-hal yang akan dibahas dalam penelitian ini yang terdiri dari latar belakang, rumusan masalah, tujuan penulisan, batasan masalah, dan sistematika pembahasan.

Bab II: Landasan Teori

Pada bab ini, akan dibahas mengenai distribusi Normal dan distribusi invers Gaussian. Kemudian akan dibahas juga mengenai metode maksimum *likelihood* yang digunakan untuk mencari penaksir parameter, serta regresi linear.

Bab III: Regresi Linear Invers Gaussian

Pada bab ini, akan dibahas mengenai parameter pada regresi linear invers Gaussian. Parameter tersebut akan ditaksir menggunakan metode maksimum *likelihood* agar dapat dibentuk garis regresi linear invers Gaussian.

Bab IV: Analisis Perbandingan Model Regresi Menggunakan Data Simulasi

Pada bab ini, akan dibahas mengenai data yang dibangkitkan dengan menggunakan bahasa program R, di mana data tersebut akan dibuat ke bentuk distribusi dengan parameter yang sudah dipilih. Kemudian akan dibentuk garis regresi sebagai calon model regresi ke dalam simulasi data tersebut.

Bab V: Kesimpulan dan Saran

Pada bab terakhir dari penelitian ini akan ditutup dengan kesimpulan dan saran untuk pengembangan penelitian lebih lanjut ke arah yang lebih baik.

